All Eyes on Them

A Field Experiment on Citizen Oversight and Electoral Integrity*

Natalia Garbiras-Díaz† and Mateo Montenegro‡

October 28, 2021

Abstract

Can Information and Communication Technologies help citizens monitor their elections? We analyze a large-scale field experiment designed to answer this question in Colombia. We leveraged Facebook advertisements sent to over 4 million potential voters to encourage citizen reporting of irregularities and varied whether candidates were informed about the campaign in a subset of municipalities. Total reports, as well as evidence-backed ones, experienced a large increase. Across a wide array of measures, electoral irregularities decreased. Finally, the reporting campaign reduced the vote share of candidates dependent on irregularities. This light-touch intervention is more cost-effective than monitoring efforts traditionally used by policymakers. (JEL D72, C93, P16, O17)

Clientelism, voter intimidation and electoral fraud are part of the long list of electoral irregularities that persistently threaten democratic institutions in the developing world (World Bank 2017). Politicians draw on these different strategies, often combining several of them, as a way of distorting elections to their advantage. Beyond the direct consequences of undermining fair elections and eroding political accountability (Stokes 2005; Hicken 2011), a growing amount of evidence has shown that different types of electoral irregularities also

*We are grateful for the guidance provided by Daron Acemoglu, Esther Duflo and Ben Olken. This paper has benefited greatly from the conversations with Abhijit Banerjee, Thad Dunning, Leopoldo Fergusson, Ray Fisman, Stuti Khemani, Horacio Larreguy, Monica Martinez-Bravo, Tara Slough, and all of the participants at the MIT Development and Political Economy lunches, the CPD working group at UC Berkeley and the BUDIE group at CEMFI. We would also like to thank Laura Pulecio, Juliana Barberena and Diana Velazco at the *Procuraduría General de la Nación*, Esteban Salazar at *PARES* and Marlon Pabón, along with the other members of the MOE that helped us, and without whom this project would have not been possible. Last but not least, we are indebted to César Gutiérrez and Sebastián Cáceres for their amazing help designing the ads used in our interventions, and Estefanía Avedaño for her outstanding research assistance. Funding for this project was generously provided by the J-Pal Governance Initiative, the Center on the Politics of Development (UC Berkeley) and the George and Obie Schultz Fund. The experiment was approved by MIT's IRB (the *Committee on the Use of Humans as Experimental Subjects*) with reference #1904805455. The RCT is registered in the AEA RCT Registry with unique identifying number "AEARCTR-0004678".

†University of California, Berkeley. Contact: nataliagarbirasdiaz@berkeley.edu.

‡CEMFI. Contact: mateomontenegro@gmail.com, Address: Calle Casado del Alisal, 5, Madrid 28014.

harm the economic and political well-being of countries.[1]

Despite systematic efforts by governments and the international community to reduce the prevalence of electoral irregularities, these remain pervasive across the developing world. The most widely used and studied measures to deter electoral irregularities, such as deploying electoral observers or auditors, require high state capacity, coordination, and large investments (including, for instance, the training of several hundreds of *in-situ* observers), which are scarce in these contexts, and particularly in remote regions where these issues are more pronounced. At the same time, both the digitization and the information and communication technology (ICT) revolution have provided a large number of tools that are cost-effective, scalable, and thus able to overcome those hurdles, which have only recently begun to be exploited to curb electoral irregularities.

In this paper, we study whether crowdsourcing the oversight of elections—i.e., outsourcing this task to a large group of people through online platforms—can be effective in the fight against electoral irregularities.[2] We analyze a massive online campaign conducted through Facebook that had the goal to encourage citizen monitoring of elections around the 2019 mayoral elections in Colombia and ask whether it was effective at both generating citizen reporting and reducing electoral irregularities.

This context provides an ideal setting to study this question for three reasons. First, electoral irregularities remain highly pervasive in Colombia despite the presence of domestic and international observers and auditors. Second, in recent years, there have been initiatives by both the Colombian government and NGOs to increase civil society's engagement in overseeing elections through online reporting websites and applications. Finally, internet access as well as social media usage have been on the rise throughout the country, with 69% of the population connected and over 60% of them registered as Facebook users.

Leveraging the features of this setting, we launched a large-scale Facebook advertisement campaign, reaching more than 4.4 million citizens, which represents a third of the targeted population. We promoted the use of an existing online website where citizens can report electoral irregularities, hosted by the *Misión de Observación Electoral* (MOE), a prominent local NGO. We randomized a subset of our sample of 698 municipalities (more than half in

[1]By increasing the political returns of targeted transfers, clientelism leads to the under-provision of public goods and it generates inefficiencies (Khemani 2015; Baland and Robinson 2007; Vicente and Wantchekon 2009). Indirectly, it is also correlated to fiscal corruption (Singer 2009), which in turn might cause inefficiencies for firms and governments alike (Olken and Pande 2012). Furthermore, voter intimidation might also help to perpetuate violence in weak states (Acemoglu et al. 2013; Robinson and Torvik 2014).

[2]Although crowdsourcing has been previously used by governments and NGOs to monitor elections, experimental studies testing for its effectiveness are still scant. A well-known example is the case of *Uchaguzi*, an election monitoring digital platform used in Kenya, which has been credited with reducing electoral violence (World Bank 2016).

the country) to receive advertisements encouraging citizen reports.

In designing the advertisements we sent to citizens, we targeted two underlying reasons why citizens might shy away from reporting. As emphasized in the literature studying persuasion (DellaVigna and Gentzkow 2010), media campaigns make a direct emotional appeal to viewers, which in our setting might prime intrinsic motivations to spur citizens to report. Alternatively, despite being encouraged to report, citizens may face a "last-mile problem" due to the difficulty of reporting and acquiring the information to do so. To understand the role of these different channels, we designed three versions of the advertisements, containing either: (i) a message informing citizens about the MOE's reporting website and providing them with a link to their online form, (ii) a call-to-action message urging citizens to act against irregularities by reporting them (but without any information about how and where to do so), or (iii) a combination of both.

In addition to understanding citizens' responses to the reporting campaign, we also wanted to study how candidates reacted to it. To do so, we further cross-randomized whether candidates running for mayor in certain municipalities, as well as their staff, received a letter informing them about the reporting campaign or not. This second treatment arm allows us to study the general equilibrium effect of these interventions to the extent that, aside from citizens, candidates also became aware of the reporting campaign and had time to react to it. Ours is one of the few studies that experimentally varies candidates' and citizens' involvement separately when examining interventions around elections.

We find that citizens in the municipalities exposed to the reporting campaign were significantly more likely to file reports, and when they did so they were also more likely to include information useful for prosecuting the instigators of electoral irregularities. In particular, being exposed to the reporting campaign increases the proportion of municipalities with at least one report by about 10.6 percentage points, and the proportion of them with evidence-backed reports by 8.3 percentage points (corresponding to an increase of 37% and 52%, respectively, from the control mean). We further find significant differences in exposing citizens to variations in the content of the ads. In particular, providing information about the reporting website generated more reports than the ads containing only messages encouraging citizens to report. This indicates that an important barrier citizens face to reporting in this context is a lack of information about where to do so.

This first set of results suggest that the reporting campaign did spur citizens' monitoring of elections. But did this discourage candidates from engaging in electoral malpractice? To answer this second question, we constructed two mutually complementary proxies to measure electoral irregularities. First, we created an original database based on media coverage of electoral irregularities at the municipal level from local and national media outlets in a wide array of formats, including written, TV and radio news, both online and offline. Second, drawing from the literature on electoral forensics, we constructed a measure of electoral irregularities based on administrative data. In particular, we leverage fine-grain voting record data to proxy electoral manipulation as deviations in the distribution of the tallies' second-digit obtained by each candidate from Benford's 2nd Digit Law, a popular measure in this literature (Hicken and Mebane 2017).

Results using both of these measures suggest that the increase in citizen monitoring significantly deterred irregularities, affecting equally municipalities where candidates were notified about the campaign and where they were not. Municipalities exposed to the reporting campaign experienced a reduction of 5.5 percentage points in the likelihood of having an irregularity mentioned in the media, and it decreased the probability that the voting counts deviated from Benford's 2nd Digit Law by approximately 8 percentage points. In examining the types of irregularities affected, we find that the largest decline comes from instances of vote-buying reported in the media, which is a common type of irregularity in this context.

Next, we study whether this reduction in electoral irregularities might have affected the electoral prospects of candidates whose votes depended on electoral malpractice. To identify such candidates from the large set of those running in the mayoral elections, we conducted a pre-treatment survey three weeks before the intervention. For every candidate running in their municipality, we asked citizens to assess each candidate's perceived likelihood of engaging in electoral irregularities. We then validated this variable against objective measures of candidates' engagement in illegal activities. The results indicate that the reporting campaign reduced the vote share of these candidates substantially. For instance, we find that candidates who rated above average in the survey-based measure of electoral irregularities experienced a 2.5 percentage point decrease in their vote share in municipalities exposed to the reporting campaign. We further provide evidence showing that this effect is not driven by a difference in the popularity of candidates identified as more prone to irregularities.

The detected drop in electoral irregularities due to the reporting campaign could explain the reduction in the vote share of those candidates who would have likely engaged in such irregularities. However, a potential alternative channel could be a change in citizens' social norms about irregularities and their preferences for different types of candidates. To approximate the relative importance of the first mechanism, we perform a set of bounding exercises that indicate that at least 28%, and potentially 97% of the effect on candidates' vote share, is due to the drop in irregularities.

Finally, we show that the reporting campaign was very cost-effective compared to other interventions studied in the literature. A simple back-of-the-envelope calculation suggests that the intervention reduced by one percentage point per polling station the vote share of the candidates most prone to engage in electoral irregularities, at a cost of $0.69 USD. As a benchmark, the more traditional strategy of deploying electoral observers costs several orders of magnitude more (over $500 USD also for a one percentage point change in vote share per polling station, in our most "optimistic" calculations), and this is true even for more recent innovations, which use ICTs to enhance electoral monitoring. Thus, by informing on how to innovate existing practices, this paper promises to contribute to ongoing endeavors of governments and democracy-promotion organizations to curb electoral irregularities.

This paper makes several contributions and builds on at least three strands of literature. First, it speaks to extant work on ways to fight electoral irregularities. Table 1 provides a summary of the interventions and their findings studied in this literature. Most studies have focused on either top-down methods to monitor elections—e.g., the use of domestic and international observers—or on bottom-up approaches to educating civil society about the negative consequences of electoral malpractice. We contribute to this scholarship by analyzing the effectiveness of a citizen monitoring campaign that combines elements from the mobilization and the monitoring strategies, which constitutes a third understudied approach.[3] Furthermore, we show that this approach is more cost-effective than the previously-studied alternatives, and we highlight its potential to be fully scalable.

Second, we also contribute to the growing literature examining the use of bottom-up technologies to improve governments' accountability and the delivery of public goods and services. Spurred by the World Bank's 2004 *World Development Report*, governments and NGOs alike have heeded the call to use these types of technology in areas as diverse as education, health, public works and elections. Most of the papers in this strand of the literature[4] have focused on studying offline interventions that do not exploit the potential of ICTs and social media to create engagement and to facilitate the spread of information to increase citizen oversight effectively. Despite recent calls to study this "second generation" of bottom-up technologies (Peixoto and Fox 2016), our paper is one of the few studies to do

[3] Four papers study interventions that included electoral reporting components, but they do so tangentially and/or their focus does align with ours. (i) Ferree et al. (2017) show that monetary incentives provided through several ICT channels increase different forms of political participation, including volunteering as citizen electoral observers, but do not study the effects on electoral irregularities. (ii) Driscoll and Hidalgo (2014) study an education campaign informing citizens how to file complaints about electoral irregularities in Georgia and find that it increased electoral irregularity reports but depressed turnout. They interpret these results as a consequence of citizens' belief they were being monitored by either the regime or researchers and that retaliation might ensue. (iii) Aker et al. (2017) study the effect of sending SMS messages inviting citizens to report electoral problems to a hotline around elections in Mozambique. Although they find no effects on the occurrence of election problems, they do find increased turnout and knowledge about elections. (iii) Finally, Gonzalez (2021) finds that areas in Afghanistan with access to cellphone coverage present less electoral fraud and argues that this is due to greater use of an electoral irregularity reporting hotline.

[4] See Fox (2015) for a review of this vast literature.

so using field experiments.[5]

Related to this last point, this paper also speaks to the literature on e-governance and ICT interventions designed to increase the accountability and the efficiency of public services. While most of this literature has focused on studying top-down efforts to increase data availability or decrease leakages,[6] this paper focuses on crowdsourcing, which is an understudied technology. Importantly, our findings provide a rare example of how social media can be used to enhance democracy and transparency, in contrast to the preponderant view that has focused on its negative effects (for a review, see Zhuravskaya et al. 2020).

The remainder of this paper is organized as follows. Section 1 provides background information about the intervention's context, including a discussion of electoral irregularities in Colombia, an overview of online electoral reporting, and the elections around which the experiment was conducted. Section 2 provides a description of the experimental design, the study sample and the data used. Section 3 presents and discusses the main results, and Section 4 provides a cost-benefit analysis of the intervention. Finally, Section 5 considers the relevance of the findings from a policy perspective and concludes.

1 Context

1.1 Electoral Irregularities in Colombia

Electoral irregularities take many forms and permeate every election in Colombia's democracy. Most existing studies have explored their extent and their main features drawing largely on qualitative accounts.[7] However, a few studies have documented the incidence of the problem quantitatively. Fergusson et al. (2017) use a list experiment to estimate Colombians' engagement in clientelistic practices—broadly defined as receiving particularistic benefits in exchange for political support—, and find that approximately 18% do so at some point in their lives.[8] Using this same method, Garcia and Pantoja (2015) show that about 7% of voters were intimidated to vote in a particular way in the 2014 presidential elections.

We will be using the term "electoral irregularities" to refer to any conduct affecting elections penalized by Colombian law. Apart from vote-buying and voter intimidation, common irregularities include:[9]

[5] Recent examples include Aker et al. (2017), Aker and Ksoll (2020) and Muralidharan et al. (2021).

[6] See Muralidharan et al. (2016), Lewis-Faupel et al. (2016), Banerjee et al. (2020), and Callen et al. (2020) for recent examples.

[7] See, for instance, Leal and Dávila (1990) and Holland and Palmer-Rubin (2015).

[8] Other studies find similar estimates in other countries in the region (Gonzalez-Ocantos et al. 2012).

[9] We approximate how common these irregularities are by the number of reports gathered about them. Figure A1 displays the number of reports made to the MOE in the 2015 elections, using the same window of time around election day used in our intervention.

Illicit political advertising: Political advertisement is forbidden on election day and on public infrastructure.

Campaigning by public servants: When public servants use their position to interfere in elections by either trying to favor or harm a particular candidate or party.

Fraud in voter registration: When citizens register to vote in a polling station in a location different from their place of residence in order to obtain an illicit profit or to alter electoral results. This is usually done as a way to facilitate vote buying as explained below.

Election disruption: When someone disturbs or prevents the normal development of elections or vote counting by deceit or force. In practice it frequently occurs through riots led by citizens or candidates.

Electoral fraud: It refers to the other illicit forms of altering electoral results, such as ballot stuffing.

The organizational details about how electoral irregularities occur vary according to the type of irregularity. As reported and studied in diverse contexts (Stokes 2005; Stokes et al. 2013), in Colombia vote-buying and other forms of clientelism are carried out via local brokers mediating between political organizations and voters. These brokers play the important role of providing political organizations with the information required to target and recruit potential voters into clientelistic relationships, as well as ensuring that these voters actually vote in the intended way. A common form of monitoring compliance is by registering voters in polling stations outside of their place of residence so that brokers can control their clients' vote choice (Rueda 2017).[10]

Other types of electoral irregularities, such as illicit political advertising and campaigning by public servants, are commonly done by or with the complicity of local politicians, who collude with candidates to return political favors (Arenas 2018). Voter intimidation is commonly performed by armed actors such as guerrillas, paramilitaries, criminal gangs, or even the military, in collusion with local or national politicians (Acemoglu et al. 2013), but also by non-armed actors, such as employers who threaten their employees with job loss if they fail to vote in a particular way.

1.2 Electoral Oversight and Reporting in Colombia

Several governmental agencies and NGOs run online electoral reporting platforms in Colombia. The *Misión de Observación Electoral* (MOE), an independent, non-partisan NGO and

[10] Evidence from other contexts shows that reciprocity (Finan and Schechter 2012) or, alternatively, targeting voters likely to sympathize with the candidate supported by the broker (Nichter 2008) are also prevalent methods to ensure compliance in vote-buying.

one of the largest devoted to promoting electoral integrity,[11] hosts one of the most important of these websites, called *Pilas con el voto* (roughly translated, "keep an eye on your vote"). This website allows users to submit reports anonymously, requiring them to specify only the date and municipality of the reported irregularity. Additionally, it asks the users to describe the reported event in a free-form field, which is then classified by the MOE's staff into one (or several) of the electoral irregularities. Depending on the amount of evidence and facts (places, names, and proof such as videos) provided about the reported irregularities, they also rank the report's "trustworthiness" in three categories: high, medium, and low.

As an NGO, the MOE does not have the power to directly investigate and take legal action about the reports they receive through *Pilas con el Voto*, yet, these are consequential. The MOE prepares official reports based on the information provided by citizens, which it disseminates widely among the general public and the government agencies in charge of investigating and sanctioning electoral irregularities. Furthermore, it shares information with these watchdog entities based on citizen reports and other sources resulting from their oversight activities.[12] This partnership has proven successful. For instance, during the 2019 local elections, after receiving several reports from citizens in El Rosario (Nariño), a municipality in the southwest of the country, the MOE alerted the Attorney-Inspector General's Office about potential electoral irregularities taking place there. Less than a year later (in May 2020), it resulted in the arrest and conviction of a mayoral candidate and his campaign chief of staff, both accused of vote buying. A similar case occurred in Manaure (La Guajira), where the MOE raised concerns about systematic reporting of "identity theft" during the voter registration process and on the day of the election. In 2020, the AG convicted eight people related to this case.

One of the main agencies involved in both investigating and sanctioning the electoral misbehavior reported to the MOE is the Attorney-Inspector General's Office in Colombia (henceforth, AG).[13] The AG is an independent institution that oversees the correct conduct of public servants through both preventive and sanctioning powers. During the electoral process, it ensures candidates abide by the law and are in good standing to run for office. Additionally, it closely monitors the electoral process.

How effective are watchdog intuitions in investigating and prosecuting the irregularities reported? Answering this question is complicated because a significant fraction of reports do

[11]The MOE has offices in more than half of the municipalities in the country, spread across all 32 departments.

[12]This collaboration has been institutionalized, as in the case of the *Eje Temático de Protección a los Mecanismos de Participación Democrática*, a special unit within the Attorney General's Office, created in 2014 to investigate cases related to electoral irregularities.

[13]The Spanish name for this institution is the *Procuraduría General de la Nación*.

not contain enough evidence for the electoral watchdogs to start a judicial case. For instance, only 13% of reports submitted to the MOE were deemed to have high-level "trustworthiness" (meaning they contain hard evidence about the facts reported). Notwithstanding this difficulty, a brief glance at the AG's database of historical reports in the 2010-2018 period suggests two important facts: (1) 2.5% of reports ultimately reached a judicial decision, such as sanction or acquittal; and (2) an additional 12.4% of reports were being investigated by the AG at the end of the period. Since investigations are publicly conducted, and thus might harm the reputation of candidates and parties, these statistics reflect that reporting does entail substantial costs for candidates accused of committing irregularities.

1.3 The 2019 Mayoral Elections

In Colombia, local authorities—such as mayors, governors, and council members—are elected every four years in a single round under a simple majority rule. For this study, we focus on the 2019 local elections and, particularly, on mayoral races.[14] Reelection is not allowed for this office, which implies that there are no incumbents in any of the races. Furthermore, given the country's weakly institutionalized party system, mapping candidates to parties to make a case for party-wise reelection is not straightforward. This lack of party discipline is reflected in the multiplicity of candidates (e.g, the average number of candidates is five in our sample of municipalities).

2 Research Design

2.1 Experimental Design

The main intervention was a large-scale Facebook advertising campaign designed to encourage citizen reporting of electoral irregularities. The campaign lasted for five days and targeted all Facebook users of 18 years or older. It started on October 24, three days before the elections, and it ended on the night of October 28, one day after the elections. This extra day allowed citizens who had witnessed electoral irregularities but had not reported them on election day to report them the next day through the reporting website.

We designed four versions of ads to flesh out the underlying mechanisms through which the campaign could potentially affect reporting. Namely, municipalities were randomized to receive Facebook ads with the following messages:

C. Placebo Control Group: A message reminding viewers about the coming elections, *"Don't forget that local elections will take place on Sunday, October 27"*.

[14] We focus on mayors since this post is assigned at the municipal level, corresponding to the level at which we randomized our intervention. Additionally, the number of mayoral candidates is tractable compared to the council members, which often had several dozen candidates.

TI. Information message: A message informing viewers about the MOE's reporting website, *"The MOE has the following website where you can report electoral irregularities: [LINK]. Don't forget that local elections will take place on Sunday, October 27".*

TC. Call-to-action message: A message with a call-to-action to report irregularities and act against them, *"In these elections let's stop electoral irregularities. Report them! Don't forget that local elections will take place on Sunday, October 27".*

TI+C. Information + call-to-action message: A message containing both TI and TC.

The rationale for each of these experimental groups is the following. We included a placebo message in the control group to net out the effect of politically-oriented advertisements on citizens' behavior. Treatments TI, TC, and TI+C, sought to separate two mechanisms through which the ads could have impacted citizens' propensity to monitor elections: either (1) by reducing the cost of reporting by informing citizens about the reporting website; or (2) by highlighting the urgency to take action against electoral irregularities.

In each of these treatment groups, the corresponding text is featured as the header of the advertisement. Additionally, a short video in a slide-show format was shown beneath the header highlighting the main points of each message. Figure 3 depicts the slides used for the different ads.[15]

By targeting ads to the universe of (adult) Facebook users in a municipality, both citizens and members of the candidates' campaigns might have become aware of the monitoring campaign directly. Moreover, they might have also learned about the campaign indirectly, by hearing about it through others. As such, the effects of treatments TI, TC, and TI+C combine both citizens' and candidates' responses to the campaign. However, it is important to note that candidates' responses were limited by the fact that we sent the ads only three days before the elections.

To fully understand how candidates would react to this campaign with more foresight, we further cross-randomized whether candidates were informed about the intervention two months before its occurrence through letters. In particular, we randomized municipalities in groups TI, TC and TI+C[16] to the following conditions:[17]

[15] The background in these slides was white so that it did not reflect any of the colors associated with particular candidates, and the main image shows a ballot box with a text urging to vote, which was the message transmitted through the placebo message.

[16] We did not allow the ad control group to be included in this second treatment arm to avoid deception —i.e., informing candidates about a campaign that was not taking place in that municipality.

[17] We further randomized two versions of these letters, with a slight change in the text. In the first version, which we call the "Full Knowledge Letter," candidates were informed about the campaign and the website promoted for the ad campaign. In the second version, which we call the "Partial Knowledge Letter," they were told that the campaign would take place, but the specific website was not mentioned. Since the difference between the treatment effects of these two groups are not statistically significant in most analyses,

TL. Letter to Candidates: All of the candidates running for Mayor in the municipalities in this group were informed about the monitoring campaign.

TC. No Letter: None of the candidates running for Mayor in the municipalities in this group were informed about the monitoring campaign.

Figure A2 displays an example of the letters sent. We partnered with the AG in sending these letters, for two main reasons. First, in Colombia, candidates' and their staffs' contact information is not public. Therefore, the AG helped by providing this information. Second, sending the letters on behalf of this entity maximized the chances that candidates would actually read them.

To boost the effectiveness of these interventions, we sent reminder letter three weeks before the elections.[18] We sent both physical letters and emails to maximize the chances of getting the candidates' attention.[19]

Figure 1 summarizes the full factorial design of our experiment, and Figure 2 shows the timeline of the different interventions and electoral milestones.

We intentionally designed all interventions not to affect any particular candidate or party: all ads were non-partisan, and we sent the letter to the universe of the mayoral candidates running in the treated municipalities. Thus, we do not have *a priori* reasons to believe that we intervened in the political outcomes of interest in ways different than potentially dampening the electoral prospects of those politicians likely to engage in electoral irregularities.

2.2 Study Sample

The study sample consisted of 698 municipalities coming from every Colombian department (see Figure A3), containing approximately 19 million inhabitants, almost 40% of the population in the country.[20] Table A1 presents the summary statistics for a selected set of variables for the municipalities included in the study sample and compares them to the uni-

we do not report their results separately in the main text. See footnote 42 for their results.

[18] As discussed in Section 1.1, the planning and execution of electoral irregularities can take place several weeks and even months before the election, so this additional forewarning time would potentially have an important deterrent effect over candidates.

[19] Notably, the deadline to register for candidacy occurred more than a month before we sent the first of these letters, so this treatment did not generate any differential effect on the selection of candidates.

[20] There are two reasons why we chose to use municipalities as the units of randomization. First, because it is the finest geographical level at which the MOE have systematic data on reports. Second, we opted to treat the universe of mayoral candidates running in a municipality to minimize the risk of favoring a particular party or candidate. The sample was chosen following population-based criteria. Namely, we only included municipalities with at least $5,000$ and no more than $97,000$ inhabitants. These cutoffs followed two guiding principles. First, we used a lower bound because Facebook's API does not allow targeting populations with too few users. Second, we chose the upper bound to keep the costs of the ads within our budget, as we wanted to ensure a high-impact treatment able to reach at least 30% of the population. Additionally, we did not consider a few municipalities where we had run a pilot and where ads had not reached more than 5% of users.

verse of municipalities in the country. As expected, given the selection criteria, the average population is smaller in our sample ($\approx 27,000$) than in the whole country ($\approx 37,000$). The municipalities in our sample also have a relatively lower Facebook penetration rate, with an average of approximately 41% of the population reported as active users by Facebook, while the average penetration rate across the country is 61%. However, apart from these differences, the experimental sample is fairly similar to the country's average across other characteristics, such as GDP per capita, rurality rates, poverty rates and previous reporting behavior.

2.3 Randomization in Practice: Stratification and Balance Checks

To increase the balance on potential confounders across treatment conditions, we conducted a stratified randomization. We defined strata by the intersection of bins partitioning the sample in three ways: (1) by the 50% and 85% percentiles of the population over the age of 18; (2) by the 20% and 80% percentiles of voter turnout in the first round of presidential elections in 2018; and (3) by whether the municipalities filed reports through the MOE's website around the congressional elections of 2018 above or below the median.

Table A2 reports balance checks for the different treatment arms using five sets of covariates, and Figure A4 displays the p-values for a subset of the treatment arms. The first set of covariates includes measures of past reporting through the MOE's website. The second one has socioeconomic covariates, such as population or per capita income. The third set includes political covariates, such as past turnout and election outcomes for different parties, as well as past values of the forensic tests we use as outcome variables. The fourth set consists of region dummies. Finally, we also include covariates of interest such as the share of the adult population reached by the intervention ads and the responses to a pre-treatment survey we discuss in the next sections.

The results suggest that municipalities are well balanced across treatment arms. Only eleven differences in means out of 248 comparisons in Table A2 are statistically significant at a 10% level or less. While these imbalances might have arisen by chance, this justifies including covariates in our main specifications (as explained in Section 2.6).

2.4 Ad Campaign Scale and Engagement

Table 2 provides summary statistics of measures reflecting the scale of the ad campaign (i) aggregated across all of the municipalities in the sample, (ii) per municipality, and (iii) per population age 18 years and above (i.e., those eligible to vote).[21]

[21] The information contained in this table comes from Facebook's Marketing API. For ease of interpretation, we altered the original name of some of the variables. In particular, "viewers of the ad" and "times the ad appeared on a screen" correspond to Facebook's "unique reach" and "impressions" variables.

Overall, the ad campaign was successful in reaching a large population in the targeted municipalities. Approximately 4.4 million Facebook users saw their treatment condition ad an average of ≈ 3 times. In each municipality, the ad reached, on average, $6,245$ people, which represents 33% of the adult population. Additionally, the ads generated substantial engagement by users: more than $23,000$ people clicked on the link to the MOE's reporting website—i.e., an average of 33.5 people per municipality—at a cost of approximately $0.23 USD per link click.[22] In terms of other types of engagement, in the average municipality, 14 people 'reacted' to the ads (i.e. by "liking" it, "loving" it, etc.), 6.5 of them shared them, and 0.63 of them left a comment.

While these statistics indicate that the ad campaign was successful in producing engagement and reaching a wide audience across the *average* municipality in our sample, it did not do so in some cases. For instance, in 21 municipalities there were absolutely no viewers. The reasons for these differences in ad viewers might have included many external factors, such as problems with internet connectivity in different areas or errors in Facebook's geolocation of some municipalities. However, these differences do not represent a threat to the identification strategy, given that they are well balanced across treatment arms: Panel E of Table A2 shows that there is balance in terms of municipalities with no viewers, in the number of viewers and also in the percentage of the population who viewed the ads.

2.5 Data

2.5.1 Outcome Variables

Our main outcomes of interest are the reports submitted by citizens to the MOE's reporting website and (proxies) for the actual occurrence of electoral irregularities. Additionally, we also analyze the impact of the interventions on electoral outcomes. We measure these outcomes combining administrative and originally collected data. Below, we describe in detail how we measure each of them.

1. Reporting. Our main outcome variable to assess whether this campaign was successful in getting citizens to report electoral irregularities is the number of reports per municipality collected through the MOE's website. For our main measure, we only consider reports submitted on the dates of the intervention: i.e., from October 24, when the campaign started, to October 28, when it ended. In addition to analyzing the total number of reports, we also disaggregate them by their 'trustworthiness,' assessed and classified by the MOE as either high, medium, or low, depending on the evidence and the information contained in the

[22] The implied "click rate"—i.e. the number of clicks by 10,000 impressions—is 0.4%. This is in the ballpark of related interventions. For instance, Brookman and Green (2014) use Facebook ads to promote campaigns of US politicians and find a click rate of 0.02%. Similarly, Enríquez et al. (2021) use Facebook ads to spread information about the performance of Mexican politicians, and find a click rate of 1-2%.

reports, as explained in Section 1. We define reports as "high quality" if these are classified as either of a medium or high level of trustworthiness.[23]

2. Electoral irregularities. We use two different types of measures to proxy for the actual occurrence of electoral irregularities.

First, we capture irregularities by their mention in the news. There are two main challenges in using this type of measure to proxy for actual electoral irregularities. First, municipalities included in our sample are small and thus do not attract as much attention from major news outlets as bigger ones. Second, some news might originate from citizen reports through the MOE, which could bias our estimates in the direction of finding more news about irregularities in places receiving the treatment.

We overcome these difficulties by constructing an original database of electoral irregularities covering both local and national media from a large set of media types, such as TV, radio, and print and online newspapers. This focus on various news types allows us to maximize the chances of including news about the municipalities in our sample. Furthermore, we purge from our measure any news that originated from citizens' reports to the MOE to avoid confounding both phenomena, although we alternatively use the full set of news as a robustness check. To avoid any type of researcher bias in the coding of these data, we outsourced the collection of this information to third parties. Finally, in using these data, we aggregate all of the irregularities in the news occurring in each municipality, but we also report the effects on specific types of irregularities. Appendix B provides further details about how we constructed this data set.

To construct our second measure of electoral irregularities, we borrow from the electoral forensics literature, which uses data-driven methods to detect electoral irregularities. In Section 3.2.2, we discuss the specific tests we use and their interpretation. In constructing these tests we use the official voting records provided by the *Registraduría Nacional del Estado Civil*, Colombia's electoral office, at the voting booth level.

3. Electoral outcomes. We also examine the effect of our intervention on election results. Specifically, we study whether it affected turnout, the vote share of particular candidates, or electoral competition. We compute all of these outcomes at the municipal level using official voting records from the *Registraduría Nacional del Estado Civil*.

2.5.2 Pre-Treatment Survey

We administered a pre-treatment survey to gather additional information that was not available from existing external sources. We conducted this survey between October 7, three

[23] We do not distinguish medium and high trustworthiness reports, since there are very few reports of the latter type in our sample (under 10% of municipalities in the sample submit such reports).

weeks before election day, until October 21, two days before the advertisement campaign began (see Figure 2). We recruited respondents through Facebook advertisements targeting users in the municipalities in the sample, inviting them to participate in the survey.[24] As with the main campaign, we displayed the survey recruitment ads to all users of age 18 or older in the sample of municipalities. Take-up was incentivized by including those who completed the survey in a raffle for several Samsung tablets (valued at $120 USD).

The survey took approximately 10-15 minutes to answer, and the recruitment ad did not refer to its content or to the upcoming elections to avoid biased responses.[25] The main goal of this survey was to collect data on voters' perceptions about the mayoral candidates running in their municipalities, which we used in the analyses that we discuss later in detail in Section 3.3. The number of complete surveys is $6,121$, coming from 630 municipalities, so the average number of responses per municipality is approximately 10. Our final sample is balanced across treatment conditions in terms of respondents' socio-demographic characteristics (see Table A3), as well as in terms of the characteristics of the municipalities from which we obtained responses (see Table A12).

2.5.3 Other Covariates

We collected a rich set of municipal-level covariates to conduct balance checks and to include as controls in the main specifications. We mentioned these variables in Section 2.3 when we described the balance checks performed. In addition, we also collected candidate-level covariates, such as sex, age, party and type of electoral platform (i.e, single party, coalition of parties, or independent). In Section C of the Appendix we describe all these variables in more detail and indicate their sources.

2.6 Empirical Analysis

Our main specification to estimate the average treatment effect of the different interventions is the following:

$$y_m = \mathbf{T}'_m\beta + X'_m\gamma + \epsilon_m \tag{1}$$

where y_m is the outcome variable for municipality m; $\mathbf{T}_m$ is a vector of indicators for the different treatment arms; X_m is a set of municipal covariates, including a set of fixed effects for the strata used in the randomization; and ϵ_m is the error term.

[24] Using Facebook ads as a survey recruitment strategy has been studied in both developed and developing contexts. It has been shown to be effective at reaching populations that are costly to reach through conventional survey methods (Samuels and Zucco 2013), and to approximate the representativeness of common recruitment methods such as phone surveys (Zhang et al. 2018).

[25] The main header in these ads read "Your opinion counts! Take our survey and participate in the raffle of three Samsung tablets. It will not take more than 10 minutes."

We estimate equation (1) using three different partitions of the treatment arms to study the effects of different subtreatments.[26] First, at the coarser level, we take $\mathbf{T}_m$ to simply include an indicator variable that takes the value of one if municipality m is in any of the treatment arms receiving ads and zero if it is in the placebo control group. Second, we study the differences between the specific messages featured in the ads by including separate indicators for whether municipality m received the information message, the call-to-action message or both, in vector $\mathbf{T}_m$. Third, we analyze the effect of the letter sent to politicians by including indicators for whether municipality m received any of the reporting ads and either (i) we did not send a letter to candidates, or (ii) we did send it.

We report Huber-White standard errors for estimates of equation (1), along with randomization inference p-values to allow for inference that does not depend on distributional assumptions or asymptotic theory (Athey and Imbens 2017; Young 2018).

Instead of specifying the covariates to be included in X_m, we use the double-post-lasso covariate selection method proposed by Chernozhukov et al. (2015) and Belloni et al. (2014) to choose them. This method increases the precision of the estimates without running into overfitting issues. Crucially, it includes covariates for which there is imbalance across treatment arms as well as important predictors of the outcome variables considered. Unless otherwise specified, the set of covariates in Panels A-D in Table A2 along with the set of strata fixed effects is the one used when relying on this method. As detailed in Section 3.5, we also report estimates without control variables as a robustness check.

3 Main Results

3.1 Results on Reporting

We begin by examining whether the reporting campaign was successful in inducing citizens to report irregularities and, in particular, high-quality reports with hard evidence that could be useful to prosecute offenders. Table 3 presents the point estimates from equation (1), while Figure 4 provides a visual representation of a subset of the estimates.

Overall, the evidence indicates that the reporting campaign increased reporting substantially, both on the extensive as well as on the intensive margin. First, we find that receiving any of the treatments increased the probability that a report was filed from a municipality by 10.6 percentage points ($p < 0.01$), which corresponds to an increase of 33% compared to the control mean. Similarly, the number of reports increased in these municipalities by

[26] This approach allows us to study the effect of different treatment conditions without sacrificing the statistical power needed to estimate the full model with nine different treatment arm combinations. However, we also report the results of the 'long model,' including all nine dummies as a robustness check.

about 0.37, a 67% increase compared to the control.[27]

Next, we examine if the intervention affected the subset of higher quality reports. Municipalities receiving any of the treatments were 8 percentage points more likely to submit a report deemed as high-quality by the MOE than control ones ($p < 0.01$), and they increased the number of these types of report by 0.17, which represent, respectively, a 52% and 87% increase compared to the control group.

We further explore *how* the ads' messages might have impacted citizens' incentives to report. To do so, we separately estimate the effect of each version of the ad (Panel B of Table 3). We start by pointing out that across outcomes, the ads containing the information of where individuals could report—i.e., the link to the MOE's website—seem to be the ones driving the positive treatment effects. In particular, the 'call-to-action' message does not change citizen reporting in a statistically significant way, while the municipalities receiving the information message saw an increase in the probability of filing reports by approximately 14 percentage points and the number of reports by 0.45 ($p < 0.01$). Also note that there are no statistically significant differences between the 'information' and the 'information + call-to-action' messages, rejecting the hypothesis that the interaction of both versions could boost the incentives to report even further than the information message alone.

Next, we study whether the letters sent to candidates had differential effects on reporting. Interestingly, the effects are larger when no letter is sent, although this difference is only significant at the extensive but not the intensive margin. We interpret this finding in light of the results reported in the following section, which suggest that the letter sent to candidates might have had an *extra* deterrent effect over candidates' decisions to engage in electoral irregularities, which in turn might have reduced reporting by citizens.

3.2 Effects on Electoral Irregularities

We now examine whether the intervention reduced the occurrence of electoral irregularities. To the extent that the reporting campaign was public, it is possible that campaign staff and candidates were informed about it and changed their behavior in response to the threat of being reported. However, as mentioned in Section 2.1, candidates' knowledge about the reporting campaign and their subsequent reaction to it would have been substantially constrained by the fact that the advertising campaign started only three days before election day, which would have given them relatively little time to adjust accordingly. We thus expect that the letters sent to candidates informing them of the reporting campaign—which were

[27] One potential concern with our main reporting outcomes is that they only capture reports filed during the period of the intervention. Thus, if the treatments shift the timing of reporting from the post-intervention into the intervention period, the main estimates might mechanically overstate the overall effect. However, as reported in Table A4, the opposite is true: the treatments caused a positive (but mostly insignificant) increase in reporting in the post-intervention period, suggesting that effects persist in the short run.

sent almost two months before election day—would generate a larger behavioral response on candidates and, in particular, a larger deterrent effect on irregularities.

Given the illicit nature of electoral irregularities, measuring them has been one of the main challenges in the literature. A popular strategy has been to use citizen reports to infer the occurrence of different types of irregularities. However, in our context, this is not possible given that reporting is directly affected by the intervention in potentially opposite ways to its effect on irregularities: i.e., while the campaign increases reporting, as shown in the last section, it might have reduced their actual occurrence. We develop two strategies to measure electoral irregularities that overcome this difficulty, which we describe next. Additionally, in Appendix F, we describe the construction of irregularity proxies using the responses to a post-intervention online survey. We prefer the outcomes reported in the main text since they rely on external sources, less subject to the biases from citizens' perceptions of irregularities, which might have been affected by the advertisement campaign. The results across different variables are consistent qualitative and quantitatively.

3.2.1 Media-Based Measure of Electoral Irregularities

Our first measure comes from an original database of electoral irregularities covered by the media, described in detail in Section 2.5.1. We construct two outcome measures using this dataset: an indicator for whether any irregularity was mentioned in the news in a given municipality, and the number of such irregularities.

Panel A of Figure 5 depicts the main treatment effects on the first of these variables, while Table 4 shows the estimates of the different treatment arms on both variables. These results indicate that the campaign had a large negative effect on electoral irregularities, at both the extensive and intensive margins. Receiving any of the treatments reduced the probability of irregularities occurring by 5.5 percentage points ($p < 0.05$), and the number of irregularities by 0.08 ($p < 0.05$). These effects are particularly large given that they represent reductions of approximately 34% and 41%, respectively, compared to the control group mean.

When examining the effect of the different ad messages, we observe that the ads containing information about the reporting website are the main driving force behind these treatment effects. The estimates for the call-to-action message are also negative, but they are not statistically significant.[28]

In line with our previous discussion about the short period that candidates had to react to the reporting campaign, we find that the letters sent to candidates generated larger and more precisely estimated reductions in the probability and the number of irregularities (see Panel C of the same table). Yet, in drawing these comparisons, note that the difference

[28]However, the difference between these treatment arms is not statistically significant, as reported at the end of Panel B, Table 4.

between the municipalities receiving the letter and those that did not is not significant.

As explained in Section 2.5.1, to avoid confounding reporting and actual irregularities, we purge our main measure of electoral irregularities from all news that arises from citizen reports to the MOE. One potential problem in excluding report-related news is that we might be mechanically ignoring (actual) irregularities in municipalities treated by the reporting campaign since news about irregularities might be more likely to mention the MOE in those municipalities. As a robustness check, we then consider an alternative measure containing *all* irregularities mentioned in the news, whether they mention reports made to the MOE or not, and we report the results in Table A5. We see that, although the precision and magnitude of some of the estimated effects drop slightly, the main results hold, and the estimates remain statistically significant.

In addition to considering the effect of the interventions on all types of electoral irregularities, in Table A6 we also examine this effect on particular cases. Although the estimates are noisy given that some types of irregularities have a substantially low probability of occurrence, we see that the reduction of irregularities seems to be concentrated on vote-buying, whose likelihood drops by 4 percentage points in municipalities exposed to the reporting campaign. Even though this type of irregularity seems to be the main driver of the effect, most types of irregularities also experience a decrease in their probability of occurrence. The direction of the estimates are preserved when excluding types of irregularities one at a time from our main measure (see Table A7).

3.2.2 Forensic Measure of Electoral Irregularities

Our second measure of electoral irregularities comes from the election forensics literature, which uses "anomalies" in the administrative voting data to infer the occurrence of irregularities (see Hicken and Mebane 2017 for a review). In particular, we test for deviations of the second digit of voting counts from Benford's 2nd Digit Law, a "natural" distribution of second digits.[29]

Although this type of test remains one of the most popular tools in the forensics literature,[30] some have cautioned against interpreting it as a *sufficient* proof of electoral irregularities (Mebane 2011; Deckert et al. 2011). In our setting, these concerns are alleviated by the fact that the we will focus on comparing the relative adherence to Benford's Law of municipalities in our treatment arms as proxies of electoral irregularities, instead of considering

[29] Jara et al. (2011) use this method on a long panel of elections in Colombia, and show that elections that fail Benford's Law coincide with those held as fraudulent by public opinion and academic studies.

[30] A related 1st Digit Law has also been proposed, but the evidence indicates that it is not suited to detect fraud in contexts, such as ours, in which the maximum number of vote counts per voting booth are capped at relatively low numbers (300 in our setting) since these caps distort the natural distribution of the first digit (Pericchi and Torres 2011).

this test as an absolute fail/pass measure of the occurrence of irregularities. Moreover, we use this measure as *complementary* to our news-based measure in assessing the occurrence of electoral irregularities, as advised by Hicken and Mebane (2017). Benford's law complements it by detecting less conspicuous electoral irregularities, which might not be covered by the news but still leave a trace in administrative records. As shown in Table A8, both types of measures are positively correlated ($\rho \approx 0.1$), but this correlation is not perfect, as expected from this discussion. However, when looking at the correlation with the individual types of irregularities reported in the media (Table A9), we find that the forensic measures are strongly correlated to those suggesting fraud, such as registration fraud and electoral fraud.[31]

We use three of the most popular tests to verify compliance with Benford's 2nd Digit Law: Pearson's χ^2, Kolmogorov-Smirnov, and Kuiper's test. We describe the respective test statistics, our construction of the main variables, and discuss their relative strengths and weaknesses in Appendix D. For ease of interpretation, we synthesize the results of these tests into two main outcome variables. First, we construct a standardized index of all three test statistics, so that larger values reflect larger deviations from Benford's 2nd Digit Law. Second, we use an indicator variable which takes the value of one if any of the tests rejects the null hypothesis —i.e., there is statistical evidence of deviations—with p-value< 0.05.

Panel B of Figure 5 presents a visual representation of the estimated effects of the interventions on the latter outcome, while the last two columns of Table 4 report the point estimates for the full set of treatments on both outcomes. Across measures, we find that the municipalities exposed to the reporting campaign experienced fewer deviations from Benford's distribution, suggesting that fewer electoral irregularities took place in these locations. In general, rejection of the null hypothesis was substantial in the study sample, with 51% of the municipalities rejecting it across any of the tests. Being exposed to any of the interventions in the reporting campaign reduced this substantially, by 8 percentage points, while, in the intensive margin, the test statistics decreased by 0.14 standard deviations ($p < 0.05$).

Consistent with our findings using the media-based measures, we also find that (i) municipalities receiving the information message had a larger drop in the deviation from Benford's distribution than those receiving the call-to-action message (although this difference is not statistically significant in the extensive margin), and (ii) that the letter sent to candidates had an "extra" effect in terms of the magnitude of the reduction in the deviation from Benford's 2nd Digit Law (although it is not statistically different from the effect in the group where we did not send letters to candidates).

[31] Interestingly, it is also positively correlated with cases of riots reported in the media, which might occur because many of these riots arise from allegations of fraud.

As a robustness check, in Table A10 we report the results when using each of the three tests separately. Consistent with the fact that the Pearson χ^2 test is less powered to reject the null hypothesis in small samples (see Appendix D for a discussion), only 28% of municipalities in the control group reject the null under this test with a significance level of 5%, while 43% and 33% of them do so when using the Kolmogorov-Smirnov and the Kuiper test, respectively. The main results remain qualitatively unchanged by the choice of tests, although the effects are not statistically significant in the case of the Pearson χ^2 test, while they are significant and larger in the case of the two remaining tests.

Lastly, in Appendix E we report and confirm the robustness of these results to using the tests suggested by Beber and Scacco (2012)—an alternative set of forensic tests commonly used in the literature—as outcome variables.

3.3 Effects on Election Outcomes

The evidence presented in the last sections shows that the reporting campaign not only increased citizen reporting but also had a robust and substantial effect deterring electoral irregularities. In so far as some candidates' electoral prospects might have depended on engaging in irregularities (e.g., the votes they might have bought or the voters they might have intimidated to vote for them), we may expect the reporting campaign to also reduce their vote share. In this section, we begin by studying this possibility and then examine if the intervention affected other electoral outcomes, such as turnout. In the next section, we discuss alternative channels through which the reporting campaign might have affected electoral outcomes and also provide estimates of how much of the changes can be accounted for by the observed reduction in electoral irregularities.

Identifying the Candidates Likely to Engage in Irregularities. While the reporting campaign might have depressed the vote share for candidates whose success depended on electoral irregularities, it might not have affected other candidates. In our setting, the main challenge in studying these differences is identifying which candidates could have been involved in irregularities had our intervention not occurred.[32] To overcome this difficulty, we use citizens' responses in the pre-treatment survey we described in Section 2.5.2 to identify which candidates were more likely to engage in electoral irregularities in each municipality. In particular, we asked respondents to state whether each candidate would engage in different types of electoral irregularities.[33]

[32] This is especially difficult given the large number of candidates running in the mayoral elections in each municipality and the scarce information about most candidates' backgrounds. Upon registering their candidacy, candidates are subject to a legal background check by the government. Thus, there is virtually no variation in candidates' legal history, which could otherwise be used for this purpose.

[33] We did this for the most common irregularities described in Section 1: vote-buying, illicit advertisement,

We then aggregated respondents' answers to these questions to construct three different variables at the candidate level. First, we computed the percentage of respondents that state a candidate was going to engage in any electoral irregularity. Denote this first variable as Z^1_{mc}, for candidate c in municipality m. Second, we created a within-municipality version of this variable by subtracting the municipality's mean from Z^1_{mc}, so that positive values indicate that the candidate is more likely than the average one to engage in electoral irregularities. Thus, if $\bar{Z}_c$ is the municipality's mean, our second variable is defined as $Z^2_{mc} = Z^1_{mc} - \bar{Z}_c$. Finally, we also created an indicator of whether Z^1_{mc} was above average, so that $Z^3_{mc} = \mathbb{1}\left\{Z^1_{mc} \geq \bar{Z}_c\right\}$. While the first variable captures the full candidate-level variation in the likelihood to engage in irregularities that we aim to test, the latter two variables have the advantage of netting out municipal-level factors that could influence responses in ways unrelated to candidates' behavior.[34]

A potential concern in using these survey-based measures is that citizen views about candidates might not be good indicators about the actual behavior of candidates in general, especially when it comes to hard-to-observe behavior such as engagement in electoral misbehavior. In order to validate our measures vis-à-vis this concern, we contrast them with an external measure that does not rely on citizen input for its construction. In particular, we take advantage of the fact that the *Fundación Paz y Reconciliación* (PARES), a well-known NGO in Colombia, conducted a background check of the candidates in 48 of the municipalities in our sample and constructed a data set indicating if each candidate had a history of past malfeasance, such as involvement in corruption or association with violent organizations.[35]

Despite the fact that the sample of municipalities investigated by PARES is relatively small and is not representative of the municipalities in our sample,[36] this outcome allows us to study how this "objective" measure compares to our survey-based method. As reported in Table A13 in the Appendix, having a history of malfeasance according to PARES is strongly and significantly correlated with all of our survey-based measures and, in particular, with the within-municipality measures, which have correlations of ≈ 0.4.

Candidate-Level Regressions. We use our candidate-specific measures of the likelihood of

campaigning by public servants, voter intimidation, fraud in voter registration, and electoral fraud.

[34]For instance, in "pessimistic" municipalities, most respondents might say that candidates will engage in irregularities regardless of whether they indeed will. The second and third measures will alleviate this concern by focusing on the within-municipality variation.

[35]The full report and dataset can be accessed here.

[36]In conversations with representatives of PARES they explained that the main selection criteria were to choose municipalities where there were high threats to electoral integrity due to a history of violence, corruption or mining interests.

engaging in electoral irregularities to estimate the following candidate-level regressions:

$$\text{Vote Share}_{cm} = \mathbf{T}'_m\beta + \psi \text{Z}_{cm} + \text{Z}_{cm} \times \mathbf{T}'_m\delta + X'_{cm}\gamma + \epsilon_{cm} \tag{2}$$

where Vote Share_{cm} is the vote share obtained by candidate c running in municipality m; $\mathbf{T}_m$ is a vector including different indicators for partitions of the treatment groups as before; Z_{cm} is one of the candidate-specific measures of the likelihood of engagement in electoral irregularities; and X_{cm} is a set of municipal and candidate covariates. Given our previous discussion, we expect that $\delta < 0$, which would suggest that candidates more engaged in electoral irregularities would experience a greater drop in their vote share.

Following our randomization strategy, we cluster the standard errors at the municipal level. However, these standard errors fail to incorporate the uncertainty generated by the fact that our survey-based measures (Z_{cm}) are estimates themselves. To account for this extra uncertainty, we report the p-values from a two-step clustered bootstrap procedure, which incorporates the variation coming from estimating Z_{cm} and then the equation (2). We give a detailed description of this procedure in Appendix G.

Given that the pre-treatment survey responses came from a subset of municipalities in the original sample, we recalculate the balance checks for these regressions and report the results in Table A12. Results in this table show that the municipalities continue to be well balanced in this sub-sample across both municipal and candidate covariates. As before, we use the double-post-lasso covariate selection method to choose the municipal and candidate level covariates to be included in the regressions.

Results from the Candidate-Level Regressions. Table 5 presents the results of estimating equation (2) using the three different survey-based measures of candidates' likelihood to engage in irregularities. Across measures, we find that candidates more likely to engage in electoral irregularities experienced a large decrease in their vote share when they were running in municipalities included in the reporting campaign interventions. The reporting campaign reduced the vote share of candidates for which 100% of respondents say they will be engaged in electoral irregularities by 3.3 percentage points ($p < 0.05$), by 2.9 for the demeaned version of this variable ($p < 0.1$), and by 2.5 for candidates above the municipal average according to this variable ($p < 0.01$). Notice that, although large, these estimates are in the "ballpark" of other interventions to curb electoral irregularities.[37]

[37] For instance, in the context of Russia, Enikolopov et al. (2013) find that polling stations where electoral observers were assigned experienced a decrease of 11 p.p. in the vote share for the incumbent party. Similarly, Blattman et al. (2019) find that an anti-vote-buying campaign in Uganda decreased the vote share of incumbents by approximately 3 p.p. Finally, Callen and Long (2015) find that announcing the implementation of a photo quick count to detect aggregation fraud in Afghanistan reduced the vote share of the most

Mimicking the results for electoral irregularities, we see that the effect on these candidates' vote share is largest for municipalities receiving the information message, yet, the difference is not always statistically significant. Likewise, the effect difference between municipalities where the candidates received a letter or not depends on the outcome used.

In the last column of Table 5 we alternatively estimate equation (2) using an indicator variable that takes the value of one if the candidate was found to have a history of past malfeasance according to the investigation conducted by PARES. Despite this regression being limited to the municipalities they investigated, we find estimates consistent with our results using the survey-based measures.

As a robustness check, we verify whether, rather than measuring the propensity to engage in electoral irregularities, our measures capture candidates' popularity.[38] To test this hypothesis, we estimated equation (2) interacting different measures of the popularity of candidates with the treatment indicators. In particular, we use two different proxies of candidates' popularity: (i) the proportion of respondents in the pre-treatment survey who say that the candidate will win the election, and (ii) an indicator for whether the candidate is running with an incumbent party or not. The results, presented in Table A15, show that the heterogeneity with respect to popularity is not significant, which suggests that this is not driving previous results.

Additional Electoral Outcomes. We now examine if the reporting campaign also impacted other electoral outcomes, such as turnout or the margin of victory. As seen in Table A16, the intervention did not affect either of these outcomes. In particular, the fact that turnout was not significantly affected suggests that the decrease in the vote share for candidates more likely to engage in irregularities must have come with a parallel increase in the vote share for "cleaner" candidates.

3.4 Interpreting the Effects of the Intervention

The evidence presented in the previous sections shows that the reporting campaign reduced electoral irregularities as well as the vote share of candidates likely to engage in them. In this section, we discuss the possible mechanisms underlying these results. We begin by analyzing the decrease in irregularities. This could have occurred through several mechanisms, including (i) a reduction in candidates' willingness to engage in irregularities and (ii) a reduction in voters' willingness to sell their votes (e.g., out of the fear of being reported or a change in social norms linked to irregularities). Importantly, both mechanisms operate in the same direction. While we cannot fully disentangle these mechanisms, the letter sent

politically connected candidates by 5.5 p.p.

[38] As reported in Table A14, the tested measures are indeed correlated with the popularity of candidates.

to candidates helps us shed light on the first one since it elicits candidates' reactions to the reporting campaign. The fact that municipalities receiving these letters experienced a larger reduction than those without them (30% or more) suggests that candidate reactions are an important factor driving the effects. Yet, the lack of statistical significance in some differences undermines our ability to draw strong conclusions from this evidence.

We now turn to the effects on election outcomes. These effects might have occurred because of the decline in electoral irregularities that would, in tandem, reduce the vote share of candidates relying on them to boost their electoral prospects. Alternatively, it is plausible that the reporting campaign could have raised the salience of electoral irregularities to voters by exposing them to this issue through Facebook advertisements. In turn, this might have changed their social norms about electoral malpractice, which might affect their decision to vote against the candidates they perceived to be at the source of irregularities, an interpretation that would be consistent with our findings. Moreover, these channels have distinct practical implications. According to the "salience channel," voters' (potentially inaccurate) perceptions about different candidates determine their votes whereas, under the "electoral irregularity channel," actual irregularities are what drive the reduction in the electoral support for candidates likely to engage in irregularities.

Observationally, both the drop in electoral irregularities and the increased issue salience might have determined the overall effect reported in the previous section.[39] One feature of our experimental design that allows us to take a first step in figuring the relative contribution of both channels in determining the overall effect is the presence of the "call-to-action" message treatment condition. Municipalities in this group were exposed to the increased salience about irregularities, but, as we have seen in the previous sections, they experienced no significant changes in reporting, or in the occurrence of electoral irregularities. Thus, they provide a group in which most (if not all) of the changes in electoral outcomes would have come from the "salience channel" and not the "electoral irregularity channel." As seen in Table 5, the effect of the call-to-action message on the vote share of candidates more likely to engage in electoral irregularities is not statistically significant and its magnitude corresponds to 3%–72% of the effect estimated for the group receiving only the information message (depending on the specification).[40] Assuming the effect of the call-to-action treatment captures the full extent of the salience channel, this would imply that between 28% and 97% of the

[39]Another channel that might have affected candidates' vote share is substituting the resources devoted to electoral irregularities for legal campaigning efforts. However, we do not emphasize this channel since it is likely that it might have dampened the drop in the vote share of candidates more likely to engage in irregularities rather than explain the effect we find.

[40]Notice that the difference between both treatments is only significant at conventional levels in the specification in column (2), although it is close to being significant in columns (1) and (4).

total decrease in the vote share of candidates likely to engage in irregularities is due to the electoral irregularity channel.

As an alternative way of quantifying the contribution of this channel, we formalize the previous discussion with a simple model. Suppose the vote share obtained by candidate c, conditional on her likelihood of engaging in irregularities (Z_c), is given by a function $V_c = V(I(T), O(T); Z_c)$, where $I(T)$ are the electoral irregularities that candidate c engages in, given the treatment status T; $O(T)$ are other factors influencing the vote share of the candidate c—including the salience of irregularities—which depend on T. Given this expression, the treatment effect on the vote share of c can be decomposed as:

$$\frac{\mathrm{d}V_c}{\mathrm{d}T} = \frac{\mathrm{d}V_c}{\mathrm{d}I}\frac{\mathrm{d}I}{\mathrm{d}T} + \frac{\mathrm{d}V_c}{\mathrm{d}O}\frac{\mathrm{d}O}{\mathrm{d}T} \tag{3}$$

We are interested in the first term of equation (3): the effect of irregularities on candidate c's vote share, which depends on treatment assignment T. We have direct estimates of $\frac{\mathrm{d}I}{\mathrm{d}T}$ from the results presented in the previous sections. On the other hand, we do not have causal estimates for $\frac{\mathrm{d}V_c}{\mathrm{d}I}$, but we can approximate this by using correlations in the control group.[41] Using these estimates, we can then approximate the percentage of the total effect on vote share due to the change in electoral irregularities by computing $\left(\frac{\mathrm{d}V_c}{\mathrm{d}I}\frac{\mathrm{d}I}{\mathrm{d}T}\right) \times 100 / \frac{\mathrm{d}V_c}{\mathrm{d}T}$.

In Table A17, we report the results of performing this exercise using different combinations of the proxies used for electoral irregularities (I_m) and Z_c in the previous analysis. The results range between 11% and 52%, with a mean of 27%, which is similar to estimates comparing the effects of the call-to-action and information message treatment arms. While this exercise is far from perfect, it provides complementary evidence of the relative importance of the "electoral irregularity channel" in determining the drop in the vote share of candidates likely to engage in electoral irregularities.

3.5 Additional Robustness

In this section, we show that our results are robust to different methodological decisions in estimating the treatment effects of the interventions. First, we study the robustness with respect to excluding controls in the estimation, except for fixed effects for the strata used in randomization. We report the redoing of the main results using these specifications in Tables A18-A20. As expected, across outcomes, the magnitude of the treatment effects is virtually unaffected, but these are less precisely estimated. Despite this, we see that most

[41] In practice, we get an estimate for $\frac{\mathrm{d}V_c}{\mathrm{d}I}$ by estimating coefficient α_3 in the following regression:

$$\text{Vote Share}_{cm} = \alpha_1 I_m + \alpha_2 \mathrm{Z}_{cm} + \alpha_3 \mathrm{Z}_{cm} \times I_m + u_{cm}$$

using different proxies for irregularities (I_m) and candidates' propensity to engage in these (Z_{mc}).

of the treatment effects discussed in the previous sections remain significant, and the main conclusions remain unaltered.

Second, we study the sensitivity of the results to using a "long model"—i.e., including all possible interactions of the treatment arms—as recommended by Muralidharan et al. (2019). We report these alternative specifications in Tables A21-A23. Despite the sizeable loss in power when using the long specifications, we see that virtually all of the estimated treatment effects preserve the same sign as the short models, and that many continue to be significant. Additionally, we see that in most of the estimations, we cannot reject the null that effects are equal within the aggregated treatment arms used in the short models.[42]

4 Cost-Benefit Analysis

To better appreciate the effectiveness of the reporting campaign, we compare its cost-effectiveness to other similar strategies to reduce electoral irregularities studied in the literature. In the following, we limit ourselves to comparisons to the few papers that report costs and we compare their cost-effectiveness as measured by their impact on votes for candidates benefiting from irregularities, which is a measure provided by most papers. We focus on interventions studying (i) electoral observers, (ii) top-down ICT technologies, and (iii) voter-education campaigns against irregularities, which are the most common interventions studied in the literature. In Appendix H, we provide the details of the following computations and provide a summary of the estimates depicted in Table A27.

We use two main benchmark metrics for the cost-benefit analysis of our intervention. First, a back-of-the-envelope calculation suggests that our intervention reduced one percentage point the vote share of candidates above average in the proportion of people who say they will engage in electoral irregularities at the cost of \$0.69 USD per polling station. Second, a similar calculation suggests that a single dollar spent in the reporting campaign reduced by 21 the votes for these same candidates.

1. Comparison to electoral observers. Despite being the strategy most used and studied to curb electoral irregularities, virtually no paper examining the effects of electoral observers reports their cost. In a discussion of their cost-effectiveness, Callen et al. (2016) estimate that the European Union spends about $\$6,000 - \$20,000$ USD per polling station deploying electoral observers in developing countries. These large costs undermine the cost-effectiveness of this strategy, even if we consider the most "optimistic" evaluations of their impact. For

[42] In Tables A24–A26, we further present the estimates of equations (1) and (2) using separate indicators for the two versions of the letters sent to candidates described in footnote 17. The results indicate that, in most cases, there is no significant difference between these treatment arms, except when analyzing the candidate-level regressions, when there seems to be an extra effect of the partial knowledge letter.

instance, Enikolopov et al. (2013) find that the random deployment of electoral observers around elections in Russia generated a drop of 11 percentage points per polling station of the vote share of the allegedly corrupt incumbent party, which is the largest effect reported for these types of interventions that we know of. Taking the range of costs per polling station mentioned before, this would imply that a one percentage point decrease in the vote share costs between $545 and $1818 USD per polling station, which is several orders of magnitude higher than the costs of our intervention.[43]

2. Comparison to top-down ICT monitoring. Two papers report the cost-effectiveness of ICT monitoring interventions. Both of them study the impact of announcing the adoption of a technology that takes pictures of electoral tallies before aggregating vote counts, as a way of reducing aggregation fraud. In the first paper, Callen and Long (2015) report that deploying this intervention in Afghanistan costs approximately $210 USD per polling station and reduces by 6 the votes for candidates "connected" to electoral authorities (which are the ones most able to benefit from fraud). This would imply that a single dollar reduced by 0.03 the votes for these candidates. In the second paper, Callen et al. (2016) report a cost of $40 USD per polling station in Uganda. Their intervention decreased the vote share of the incumbent candidate by 3 percentage points, which implies that reducing a single percentage point costs approximately $13 USD. Although these strategies are substantially more cost-effective than deploying electoral observers, they are considerably less so than the reporting campaign we study, mainly because they require sending staff to take pictures of voting tallies. Crowd-sourcing this task to decrease these costs, in a spirit similar to our intervention, might be a promising avenue for future research.

3. Comparison to voter-education interventions. The lion's share of the voter education campaigns studied in the literature involves the training and mobilization of staff to convey messages and develop activities related to reducing electoral irregularities (Collier and Vicente 2014; Vicente 2014; Hicken et al. 2018; Blattman et al. 2019; Schechter and Vasudevan 2021). Although we can legitimately suspect such efforts probably involve large costs relative to their reported effects on irregularities, virtually none of these papers report the costs of their studied interventions. One exception is Schechter and Vasudevan (2021), who, instead of relying on in-person campaigns, explore the effect of a radio campaign informing citizens of the economic consequences of vote-buying in India. They find that the intervention was extremely cost-effective in reducing the votes of candidates running with vote-buying parties, with a one-dollar investment translating into 109 fewer votes for these

[43] These estimates might be overstated given that the costs reported by Callen et al. (2016) refer to less developed countries than Russia. However, even if real costs were 10% of the considered ones, this intervention would still be more than two orders of magnitude costlier than the one considered in this paper.

candidates, approximately five times as cost-effective as the reporting campaign we deployed. Importantly, however, the experimental design of this intervention ruled out possible reactions of candidates, which might have muted its effect.[44] This implies that the practical implementation of this intervention, which would entail the full knowledge and reaction of candidates, could presumably lead to smaller cost-effectiveness estimates. However, since this intervention produces estimates in the same order of magnitudes as ours, it highlights the great promise of using media to enhance electoral integrity.

As these comparisons show, the intervention considered in this paper leads to substantial cost reductions compared to other alternatives studied in the literature. However, we note some of the prerequisites necessary to make this intervention work: (i) electoral watchdogs willing and capable to punish authors of electoral irregularities; (ii) a well-functioning (user-friendly) reporting platform that guarantees reporter anonymity; (iii) simple and attractive social media advertisements to popularize the use of the reporting platform; and (iv) widespread internet access. Before adopting this strategy in different contexts, it is thus necessary to check whether these conditions are in place in the particular setting.

5 Conclusion

Despite substantial efforts by both governments and international agencies to fight electoral irregularities, these remain an important issue that hampers accountability and development in a large part of the world. In this paper, we provide evidence that ICTs incorporating civil society in the oversight of elections are an effective way to promote electoral integrity in the context of widespread irregularities.

The policy implications of these findings are three-fold. First, we find that using a social media campaign to induce citizens to report is an effective way to spur high-quality, evidence-backed reports that can be used by competent authorities to prosecute the authors of electoral irregularities. Second, we find that this intervention deters electoral irregularities in a more cost-effective way than other strategies traditionally used by governments, NGOs, and international organizations—e.g., election observers and different education campaigns. The reason for this is that hiring, recruiting, and mobilizing specialized personnel to oversee elections involves substantial costs compared to the simple idea of recruiting citizens to watch over their own elections through inexpensive ICT channels, such as Facebook advertisements. These cost-saving strategies should inform current endeavors to increase electoral integrity, and be adapted to local conditions to exploit the relative advantages of

[44] Schechter and Vasudevan (2021) argue that since the radio campaign occurred three days before the elections when electioneering is prohibited and most vote-buying occurs in this context, candidates would not have the chance to react to it.

this approach. Finally, by relying on social media to disseminate our reporting campaign, we believe this type of strategy can easily be scaled up in contexts where there is enough internet and social media penetration. Moreover, given the rapid expansion of the internet throughout the developing world, these types of strategies will be increasingly more suitable for the most remote regions, where alternative options may be forbiddingly more costly, and government accountability is worse (World Bank 2017).

The findings in this paper also open several important avenues for future research. To begin with, how do these types of interventions affect citizens' views and trust about democracy and the government more generally?[45] Although we do not find effects on turnout, access to voice channels and increased accountability might improve citizen attitudes toward democratic institutions. Yet, this probably requires sustained exposure for the mechanisms to emerge. Second, what are the downstream effects of this type of intervention on the accountability of governments and the provision of public goods and services? By decreasing the popular support of candidates more engaged in electoral irregularities, these interventions might generate positive effects on the selection and agency of candidates. Third, we have shown the potential of a social network like Facebook to amplify the scope and effect of campaigns that encourage citizen monitoring of elections. However, our experimental design does not allow us to fully test for their efficacy compared to other dissemination vehicles, including traditional media or other technologies.[46]

References

Acemoglu, Daron, Ali Cheema, Asim I. Khwaja, and James A. Robinson. 2020. "Trust in State and Nonstate Actors: Evidence from Dispute Resolution in Pakistan." *Journal of Political Economy* 128 (8): 3090–30147.

Acemoglu, Daron, James Robinson, and Rafael J. Santos-Villagran. 2013. "The Monopoly of Violence: Evidence from Colombia." *Journal of the European Economic Association* 11:5–44.

Aker, Jenny C, and Christopher Ksoll. 2020. "Can ABC Lead to Sustained 123? The Medium-Term Effects of a Technology-Enhanced Adult Education Program." *Economic Development and Cultural Change* 68 (3): 1081–1102.

Aker, Jenny C., Paul Collier, and Pedro C. Vicente. 2017. "Is Information Power? Using Mobile Phones and Free Newspapers during an Election in Mozambique." *The Review of Economics and Statistics* 99 (2): 185–200.

Arenas, Natalia. 2018. "El primer eslabón: La compra de los ediles." In *El dulce poder: Así funciona la política en Colombia,* 51–58. Penguin Random House.

[45] Recent accounts studying this question are Ofosu (2019), Berman et al. (2019), and Acemoglu et al. (2020).

[46] See Erlich et al. (2018) for a discussion of the practical advantages and disadvantages of using different communication channels in a similar setup.

Asunka, Joseph, Sarah Brierley, Miriam Golden, Eric Kramon, and George Ofosu. 2019. “Electoral Fraud or Violence: The Effect of Observers on Party Manipulation Strategies.” *British Journal of Political Science* 49 (1): 129–151.

Athey, Susan, and Guido Imbens. 2017. “The Econometrics of Randomized Experiments.” In *Handbook of Economic Field Experiments,* edited by Abhijit Banerjee and Esther Duflo, 1:73–140. North-Holland.

Baland, Jean-Marie, and James A. Robinson. 2007. “How Does Vote Buying Shape the Economy.” In *Elections for Sale: The Causes and Consequences of Vote Buying,* edited by Frederic Charles and Schaffer Andreas. Lynne Rienner Publishers.

Banerjee, Abhijit, Esther Duflo, Clement Imbert, Santhosh Mathew, and Rohini Pande. 2020. “E-Governance, Accountability, and Leakage in Public Programs: Experimental Evidence froma Financial Management Reform in India.” *American Economic Journal: Applied Economics* 12 (4): 39–72.

Beber, Bernd, and Alexandra Scacco. 2012. “What the Numbers Say: A Digit-Based Test for Election Fraud.” *Political Analysis* 20:211–234.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. “Inference on Treatment Effects after Selection among High-Dimensional Controls.” *The Review of Economic Studies* 81 (2): 608–50.

Berman, Eli, Michael Callen, Clark C Gibson, James D Long, and Arman Rezaee. 2019. “Election Fairness and Government Legitimacy in Afghanistan.” *Journal of Economic Behavior & Organization* 168:292–317.

Blattman, Christopher, Horacio Larreguy, Benjamin Marx, and Otis Reid. 2019. “Eat Widely, Vote Wisely? Lessons from a Campaign Against Vote Buying in Uganda.” Working paper.

Brookman, David, and Donald Green. 2014. “Do Online Advertisements Increase Political Candidates’ Name Recognition or Favorability? Evidence from Randomized Field Experiments.” *Political Behaviour* 36:263–289.

Buzin, Andrei, Kevin Brondum, and Graeme Robertson. 2016. “Election Observer Effects: A Field Experiment in the Russian Duma Election of 2011.” *Electoral Studies* 44:184–191.

Callen, Michael, Clark C. Gibson, Danielle F. Jung, and James D. Long. 2016. “Improving Electoral Integrity with Information and Communications Technology.” *Journal of Experimental Political Science* 3:4–17.

Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaeek. 2020. “Data and Policy Decisions: Experimental Evidence from Pakistan.” *Journal of Development Economics* 146:1–10.

Callen, Michael, and James D. Long. 2015. “Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan.” *American Economic Review* 105 (1): 354–381.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. “Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments.” *American Economic Review: Papers and Proceedings* 105 (5): 486–490.

Collier, Paul, and Pedro C Vicente. 2014. “Votes and Violence: Evidence from a Field Experiment in Nigeria.” *The economic journal* 124 (574): 327–355.

Deckert, Joseph, Mikhail Myagkov, and Peter C Ordeshook. 2011. "Benford's Law and the Detection of Election Fraud." *Political Analysis* 19 (3): 245–268.

DellaVigna, Stefano, and Matthew Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics* 2 (1): 643–669.

Driscoll, Jesse, and Daniel Hidalgo. 2014. "Intended and Unintended Consequences of Democracy Promotion Assistance to Georgia After the Rose Revolution." *Research and Politics,* 1–13.

Enikolopov, Ruben, Vasily Korovkin, Maria Petrova, Konstantin Sonin, and Alexei Zakharov. 2013. "Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections." *Proceedings of the National Academy of Sciences* 110 (2): 448–452.

Enríquez, José Ramón, Horacio Larreguy, John Marshall, and Alberto Simpser. 2021. "Online Political Information: Facebook Ads, Electorate Saturation and Electoral Accountability in Mexico." Working paper.

Erlich, Aaron, Danielle F. Jung, James D. Long, and Craig McIntosh. 2018. "The Double-Edged Sword of Mobilizing Citizens Via Mobile Phone in Developing Countries." *Development Engineering* 3:34–46.

Fergusson, Leopoldo, Carlos Molina, and Juan Felipe Riaño. 2017. "I Sell My Vote, And So What? A New Database and Evidence From Colombia." *Documentos CEDE,* no. 20, 1–13.

Ferree, Karen E., Clark C. Gibson, Danielle F. Jung, James D. Long, and Craig McIntosh. 2017. "How Technology Shapes the Crowd: Participation in the 2014 South African Election." CEGA Working paper n. 67.

Finan, Frederico, and Laura Schechter. 2012. "Vote-Buying and Reciprocity." *Econometrica* 80 (2): 863–881.

Fox, Johnathan A. 2015. "Social Accountability: What Does the Evidence Really Say?" *World Development* 72:346–361.

Garcia, Miguel, and Sebastian Pantoja. 2015. "Incidencia del clientelismo segun riesgo electoral y de violencia: Un análisis de las elecciones presidenciales de 2014 en municipios de consolidación territorial." In *Mapas y factores de riesgo electoral. Elecciones de autoridades locales,* 291–313. Mision de Observación Electoral.

Gonzalez, Robert. 2021. "Cell Phone Access and Election Fraud: Evidence from a Spatial Regression Discontinuity Design in Afghanistan." *American Economic Journal: Applied Economics* 13 (2): 1–51.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David W. Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science* 56 (1): 202–217.

Hicken, Allen. 2011. "Clientelism." *Annual Review of Political Science* 14:289–310.

Hicken, Allen, Stephen Leider, Nico Ravanilla, and Dean Yang. 2018. "Temptation in Vote-Selling: Evidence from a Field Experiment in the Philippines." *Journal of Development Economics* 131:1–14.

Hicken, Allen, and Walter R. Mebane. 2017. "A Guide to Elections Forensics." *Research and Innovation Grants Working Papers Series: USAID,* 1–38.

Holland, Alisha C., and Brian Palmer-Rubin. 2015. "Beyond the Machine: Clientelist Brokers and Interest Organizations in Latin America." *Comparative Political Studies* 48 (9): 1186–1223.

Hyde, Susan D. 2007. "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60 (1): 37–63.

———. 2010. "Experimenting in Democracy Promotion: International Observers and the 2004 Presidential Elections in Indonesia." *Perspectives on Politics* 8 (2): 511–527.

Ichino, Nahomi, and Matthias Schündeln. 2012. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *The Journal of Politics* 74 (1): 292–307.

Jara, Diego, Luis Parra, Álvaro Riascos, Mauricio Romero, and Santiago Saavedra. 2011. "Análisis digital y detección de elecciones atípicas en Colombia." *Documentos CEDE, n.2011-40,* 1–32.

Khemani, Stuti. 2015. "Buying Votes Versus Supplying Public Services: Political Incentives to Under-Invest in Pro-Poor Policies." *Journal of Development Economics* 117:84–93.

Leal, Francisco, and Andrés Dávila. 1990. *Clientelismo: El sistema politico de Colombia y su expresión regional.* Bogotá: Universidad Nacional/Tercer Mundo Editores.

Leeffers, Stefan, and Pedro C. Vicente. 2019. "Does Electoral Observation Influence Electoral Results? Experimental Evidence for Domestic and International Observers in Mozambique." *World Development* 114:42–58.

Lewis-Faupel, Sean, Yusuf Neggers, Benjamin A. Olken, and Rohini Pande. 2016. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia." *American Economic Journal: Economic Policy 2016, 8(3): 258–283* 8 (3): 258–283.

Mebane, Walter R. 2011. "Comment on "Benford's Law and the Detection of Election Fraud"." *Political Analysis* 19 (3): 269–272.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review* 106 (10): 2895–2929.

Muralidharan, Karthik, Paul Niehaus, Sandip Sukhtankar, and Jeffrey Weaver. 2021. "Improving Last-Mile Service Delivery Using Phone-Based Monitoring." *American Economic Journal: Applied Economics* 13 (2): 52–82.

Muralidharan, Karthik, Mauricio Romero, and Kaspar Wüthrich. 2019. "Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments." NBER Working paper N.26562.

Nichter, Simeon. 2008. "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102 (1): 19–31.

Ofosu, George Kwaku. 2019. "Do Fairer Elections Increase the Responsiveness of Politicians?" *American Political Science Review* 113 (4): 963–979.

Olken, Benjamin, and Rohini Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4:479–505.

Peixoto, Tiago, and Jonathan Fox. 2016. "When Does ICT- Enabled Citizen Voice Lead to Government Responsiveness?" *Background paper for the World Development Report 2016,* 1–26.

Pericchi, Luis, and David Torres. 2011. "Quick Anomaly Detection by the Newcomb–Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela." *Statistical Science* 26 (4): 502–516.

Robinson, James A., and Ragnar Torvik. 2014. "The Real Swing Voter's Curse." *American Economic Review: Papers and Proceedings* 99 (2): 310–315.

Rueda, Miguel R. 2017. "Small Aggregates, Big Manipulation: Vote Buying Enforcement and Collective Monitoring." *American Journal of Political Science* 61 (1): 163–177.

Samuels, David, and Cesar Zucco. 2013. "Using Facebook as a Subject Recruitment Tool for Survey-Experimental Research." SSRN Working paper.

Schechter, Laura, and Srinivasan Vasudevan. 2021. "Persuading Voters to Punish Corrupt Vote Buyers: Experimental Evidence from a Large-Scale Radio Campaign in India." Working paper.

Singer, Matthew M. 2009. "Buying Voters with Dirty Money: The Relationship between Clientelism and Corruption." Presented at the annual American Political Science Association Meeting.

Stokes, Susan C. 2005. "Perverse Accountability: A Formal Model of Machine Politics with Evidence from Argentina." *American Political Science Review* 99 (3): 315–325.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics.* Cambridge University Press.

Vicente, Pedro C. 2014. "Is Vote-buying Effective? Evidence from a Field Experiment in West Africa." *The Economic Journal* 124 (574): 356–387.

Vicente, Pedro C., and Leonard Wantchekon. 2009. "Clientelism and Vote Buying: Lessons from Field Experiments in African Elections." *Oxford Review of Economic Policy* 25 (2): 292–305.

World Bank. 2016. *World Development Report: Digital Dividends.* Washington, DC: World Bank.

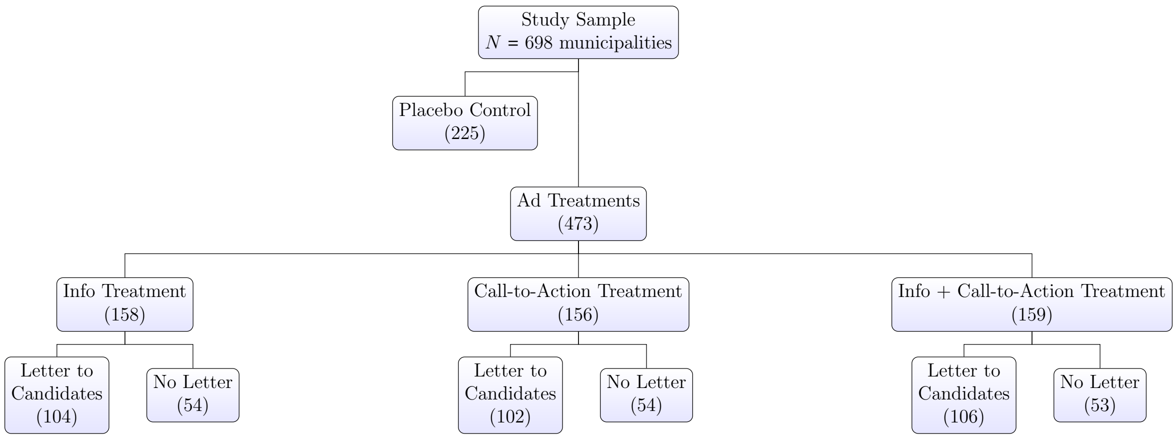
———. 2017. *World Development Report: Governance and the Law.* Washington, DC: World Bank.

Young, Alwyn. 2018. "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *he Quarterly Journal of Economics* 134 (2): 557–598.

Zhang, Baobao, Matto Mildenberger, Peter D. Howe, Jennifer Marlon, Seth A. Rosenthal, and Anthony Leiserowitz. 2018. "Quota Sampling Using Facebook Advertisements." *Political Science Research and Methods,* 1–7.

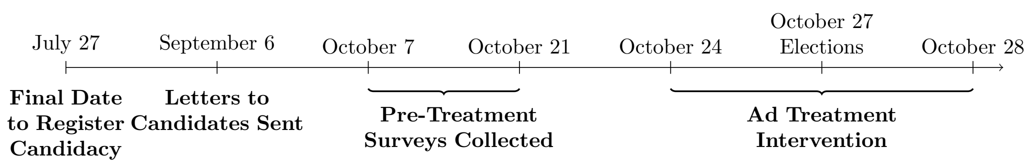
Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political Effects of the Internet and Social Media." *Annual Review of Economics* 12:415–438.

Figure 1: Randomization Design



Notes: This figure illustrates the experimental design of the experiment. The sample size within each treatment group is shown in parenthesis.

Figure 2: Timeline of the Intervention



Notes: This figures shows the timeline of the interventions performed in the study. Note that the timeline is not drawn to scale.

Figure 3: Ad Slideshow

(a) Slide A: *"Report Electoral Irregularities!"*

(b) Slide B: *"Reporting Website: Pilas con el voto"*

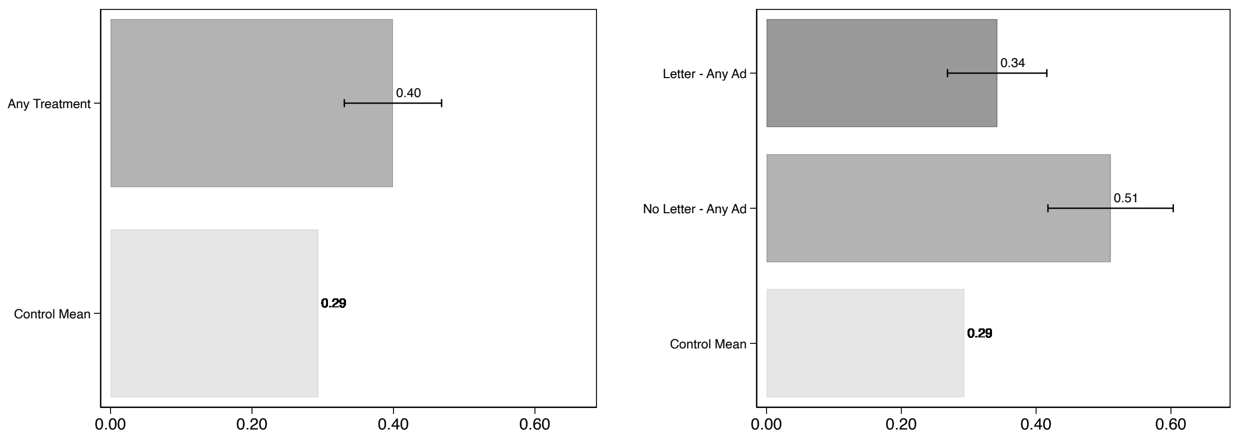
(c) Slide C: *"Sunday October 27"*

(d) Slide D: *"Next local elections"*

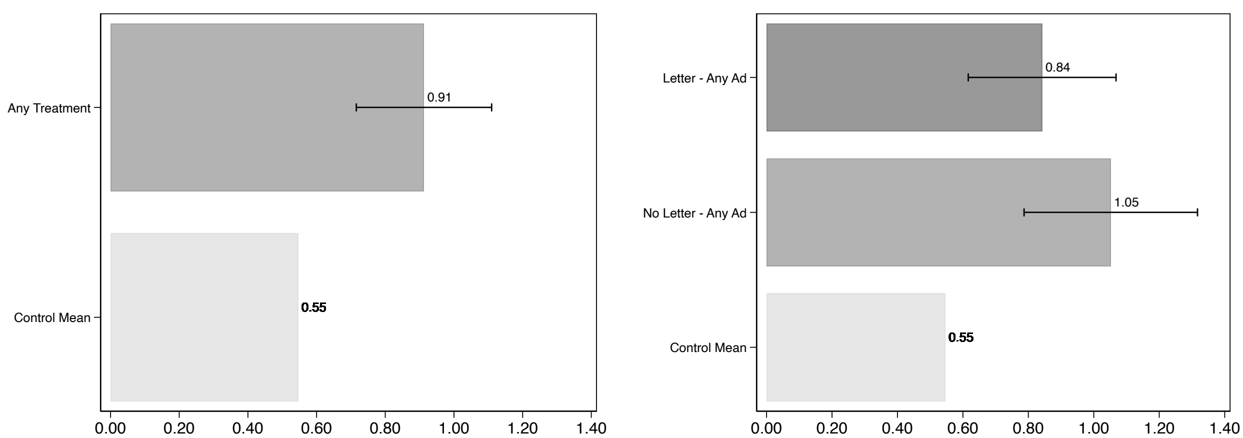
Notes: The four possible slides shown on the ad interventions are shown in this figure. Below each slide is a translation to English of the text contained in the slides. The Placebo Control group was shown only Slides C and D. The Information message group was shown slides B, C and D. The Call-to-Action message group was shown slides A, C and D. Finally, the group with both the Call-to-Action and the Information message was shown all of the slides, A-D.

Figure 4: Impacts on Reports

(a) Reports (=1)



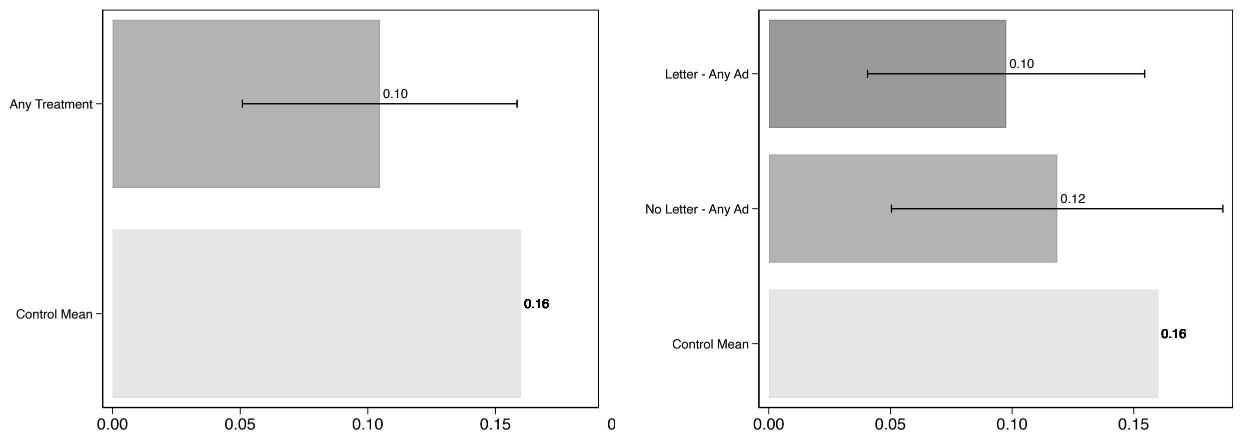
(b) Number of Reports



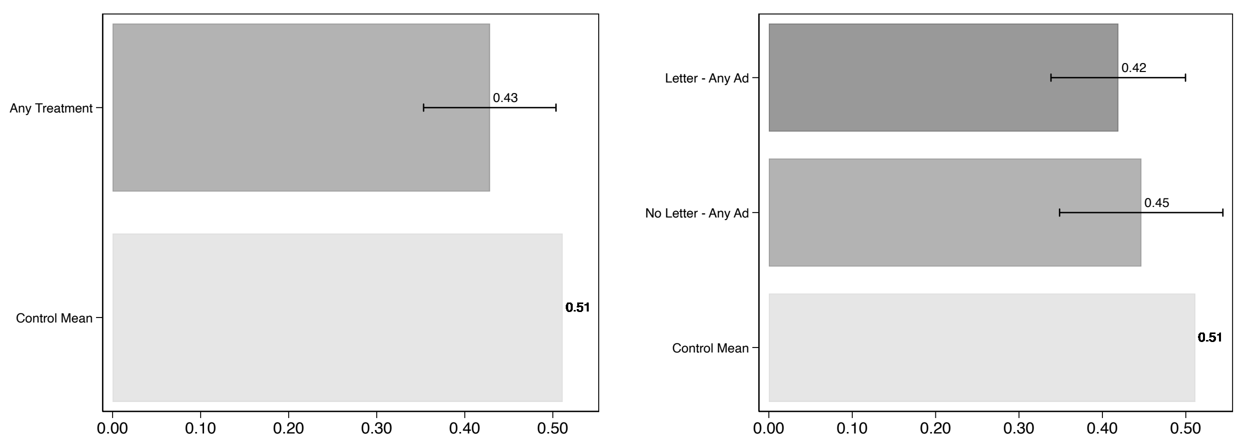
Notes: This figure reports the effects of the intervention on report outcomes. The outcome in panel (a) is an indicator for whether any report was issued to MOE from each municipality. In panel (b) it is the number of such reports. The control group mean is shown in each plot, while the remaining bars report the control group mean plus the effect for each treatment arm estimated using specification (1). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). 95% confidence intervals are reported.

Figure 5: Impacts on Irregularity Measures

(a) Media Irregularities (=1)



(b) Any P-value From Forensic Tests < 0.05 (=1)



Notes: This figure reports the effects of the intervention on two different electoral irregularity measures. The outcome in panel (a) is an indicator for whether any irregularity was reported in the media in a particular municipality. In panel (b) it is an indicator that takes the value of one if the p-value of any of the tests described in Section 3.2.2 leads to rejection of the null hypothesis with less than a 5% significance level. The control group mean is shown in each plot, while the remaining bars report the control group mean plus the effect for each treatment arm estimated using specification (1). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). 95% confidence intervals are reported.

Table 1: Experimental and Quasi-experimental Interventions to Curb Electoral Irregularities

Paper	Context	Type of intervention	N. Observations/Treated	Direct measures of irregularities	Other related outcomes
Panel A. Election observers					
Asunka et al. (2019)	Ghana/ 2012	Domestic observers	1,794 polling stations/1,230	•Reports of violence/intimidation (from surveys): -7pp	•Turnout: -4.5pp •Abnormal Turnout[a] (=1): -3.3pp – -4.7pp
Buzin et al. (2016)	Russia/ 2011	Domestic observers	7,123 polling stations/768	NA	•Turnout, Incumbent's and challengers' vote share: Null[†]
Enikolopov et al. (2013)	Russia/ 2011	Domestic observers	3,164 polling stations/156	NA	•Turnout: -6.5pp •Incumbent's vote share:-10.8pp •Challengers' vote share: 1.7pp – 3.5pp
Hyde (2007)	Armenia/ 2003	International observers	1,763 poling stations/1,008	NA	•Incumbent's vote share: -2pp–-5.9pp
Hyde (2010)	Indonesia/ 2004	International observers	1,822 villages/482	NA	•Incumbent's vote share: 6.5pp
Ichino and Schündeln (2012)	Ghana/ 2008	Domestic observers	868 electoral areas /276	• % Change in registered voters: -3.5	NA
Leeffers and Vicente (2019)	Mozambique/ 2009	Domestic observers	8,394 polling stations/989	NA	•Turnout: -2pp •Blank votes: 1.2pp •Incumbent's vote share: Null[†] •Challenger's vote share: Null[†]
		International observers	1,812 polling stations/ 67[b]	NA	•Turnout: Null •Blank votes: Null •Incumbent's vote share: Null •Challenger's vote share: Null[†]
Panel B. Top-down ICT monitoring					
Callen and Long (2015)	Afghanistan/ 2010	Announcement of photo quick count	471 polling centers/238	•Aggregation fraud (number of votes):[‡] -18.790 (politicians connected to aggregators)	•Votes for most politically connected candidate: -5.5pp •Damaged election material (= 1): -10.8pp
Callen et al. (2016)	Uganda/ 2011	Announcement of photo quick count[c]	1,001 polling centers/681	•Missing voting tallies (=1): -5.6pp •Adjacent two last digits in winning candidate's voting tally (=1): -7.9pp	•Incumbent's votes: Null
Panel C. Voter education campaigns					
Blattman et al. (2019)	Uganda/ 2016	Anti-vote-buying in-person campaign	2,796 villages/1,427	Survey data: •Perception about others' likelihood of selling vote: -0.061SD •Vote-buying: Null for incumbents and 0.062SD for challengers	•Turnout: Null[†] •Incumbent's vote share: -0.063SD (identical but opposite sign for challenger's vote share)
Collier and Vicente (2014)	Nigeria/ 2007	Anti-violence campaign	24 enumeration areas (50 inds. per area)/12	•Violence events reported by journalists (=1): -46.8pp •Perceptions of politically motivated violence: -0.23SD	•Turnout: 11.1pp •Incumbent's vote share: 8–12.8pp •Challengers' vote share: -7pp
Hicken et al. (2018)	Philippines/ 2013	T1: Promise not to sell vote, T2: Promise to vote in good consciousness	883 voters/592	Vote switching as proxy for vote-selling: •T1: -9.53 •T2: Null	NA
Schechter and Vasudevan (2021)	India/ 2014	Anti-vote-buying radio campaign	60 radio stations/30	•Vote share for clientelistic parties: -7.14 pp	•Turnout: Null •Incumbent's vote share: Null
Vicente (2014)	São Tomé and Príncipe/ 2006	Anti-vote-buying in-person campaign	50 enumeration areas/40	•Perception that voting is affected by vote-buying: -0.42– -0.46SD •Perception that voting was conducted in good conscience: 0.32–0.48SD •Frequency of vote-buying: -0.17 – -0.22SD	•Turnout: -2.5– -6.4pp •Incumbent's vote share: 3.4pp •Challengers' vote share: -3.5pp
Panel D. Bottom-up monitoring					
Aker et al. (2017)	Mozambique/ 2009	T1: Education campaign; T2: SMS w/info. to reporting hotline; T3: newspapers w/info. to hotline + T1	161 polling stations/120	•N. of electoral problems reported by election observers (T1/T2/T3): Null/ Null/ -0.588	•Turnout: 5.3pp/5.3pp/5.4pp •Incumbent's vote share: 4.6/Null/4.1pp •Challengers' vote share: -3.2pp/Null/-1.4
Driscoll and Hidalgo (2014)	Georgia/ 2008	Education campaign to file electoral complaints	84 precincts/42	NA	•Complaints (=1): 12pp •Number of complaints: 0.26 •Turnout: -5.44pp •Incumbent's vote share: Null
Gonzalez (2021)	Afghanistan/ 2009	Cellphone access	1,074-2,039 polling stations	•Abnormal votes:[d] -4pp •Abnormal polling centers: -7.7pp	•Complaints to reporting hotline: 0.144 - 0.257 •Number of insurgent attacks: Null

Notes: This table summarizes the literature on experimental and quasi experimental interventions to curb electoral irregularities. In the interest of space, we chose the most relevant findings related to electoral irregularities and related outcomes for each paper. In some cases this meant leaving out findings from some papers that do not completely align with the objectives of this table. [†] Effects are statistically significant for some subgroups. [‡] Only for estimated lower Lee bound. [a] Abnormal turnout is measured using different indicators of high turnout relative to the constituency mean or median. [b] The authors indicate that 67 international observers were sent to several polling stations, but do not indicate the exact number. [c] The authors study different types of announcement letter, but we only report the overall effect. [d] The author's measure of abnormal votes combine outliers in turnout, votes for the winning candidate and complaints about irregularities.

Table 2: Scale of Ad Campaigns

	Total	Per municipality	Per Population > 18 years
Viewers of the Ad	4,358,870	6244.80	0.33
Times the Ad Appeared on a Screen	12,886,427	18,461.93	0.99
People Clicking on the MOE's Link*	23,418	33.55	1.79 (per 1000)
Users Reacting to Ad	9623	13.79	0.74 (per 1000)
Post Shares	4531	6.49	0.35 (per 1000)
Comments on Ad	437	0.63	0.03 (per 1000)

Notes: This table reports several measures of the scale of the Facebook advertisement campaign, as well as metrics of the engagement of users with the ads. The variables reported in this table are defined as follows. 'Viewers of the Ad' are the number of distinct individuals who saw the ads at least once. 'Times the Ad Appeared on a Screen' are the number of times the ads appeared on any screen. 'People Clicking on MOE's Link' are the number of distinct individuals who clicked on the link landing on MOE's reporting website. $*$: For this metric, we only considered the municipalities actually receiving the link to MOE's website (i.e. the ones receiving the Information message) when computing the measures by municipality and population 18 or older. 'Users Reacting to Ad' are the number of distinct individuals who reacted to the ad by clicking on one of the available Facebook reactions (i.e. like, love, laugh, etc...). 'Post Shares' are the number of times people shared the ad in their own timeline, in other friends' timelines or in groups. 'Comments on Ad' are the number of comments made on the ads.

Table 3: Impacts on Reports

	(1) Reports (=1)	(2) N. Reports	(3) High Quality Reports (=1)	(4) High Quality N. Reports
Panel A. Pooled Treatment				
$[T]$ Any treatment	0.106***	0.366***	0.083***	0.173***
	(0.035)	(0.101)	(0.030)	(0.054)
	[0.004]	[0.002]	[0.012]	[0.000]
Panel B. Subtreatments by Types of Ads				
$[IA]$ Information Ad	0.144***	0.452***	0.069*	0.114*
	(0.048)	(0.147)	(0.040)	(0.064)
	[0.000]	[0.000]	[0.056]	[0.062]
$[CA]$ Call-to-Action Ad	0.018	0.180	0.069*	0.144*
	(0.045)	(0.137)	(0.041)	(0.083)
	[0.700]	[0.170]	[0.070]	[0.058]
$[I+CA]$ Info + Call-to-Action Ad	0.154***	0.462***	0.109***	0.259***
	(0.046)	(0.148)	(0.041)	(0.086)
	[0.000]	[0.002]	[0.016]	[0.002]
Test $IA = CA$, p-value	0.02	0.12	0.99	0.74
Test $IA = I+CA$, p-value	0.85	0.96	0.39	0.13
Test $CA = I+CA$, p-value	0.01	0.11	0.39	0.29
Panel C. Subtreatments by Letter - No Letter				
$[NL]$ No Letter - Any Ad	0.217***	0.505***	0.138***	0.246***
	(0.047)	(0.135)	(0.043)	(0.084)
	[0.000]	[0.000]	[0.002]	[0.000]
$[L]$ Letter - Any Ad	0.049	0.295**	0.054*	0.136**
	(0.038)	(0.115)	(0.032)	(0.058)
	[0.210]	[0.020]	[0.098]	[0.024]
Test $NL = L$, p-value	0.00	0.16	0.04	0.21
Control Mean	0.29	0.55	0.16	0.20
Sample Size	698	698	698	698

Notes: The outcome in column (1) is an indicator for whether any report was issued to MOE from each municipality. In column (2) it is the number of such reports. In columns (3)-(4) the same definitions are used on the subset of reports of a high quality (see Section 1 for a discussion about how quality of reports is assessed by MOE). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table 4: Impacts on Irregularity Measures

	(1)	(2)	(3)	(4)
	Media-Based Irregularities		Deviations from Benford's 2nd Digit Law	
	Media Irregularities (=1)	Number of Media Irregularities	Index of all Forensic Test Stats (z-score)	Any P-value From Forensic Tests < 0.05 (=1)
Panel A. Pooled Treatment				
$[T]$ Any treatment	-0.055**	-0.082**	-0.144**	-0.083**
	(0.027)	(0.037)	(0.068)	(0.038)
	[0.042]	[0.018]	[0.044]	[0.024]
Panel B. Subtreatments by Types of Ads				
$[IA]$ Information Ad	-0.060*	-0.092**	-0.299***	-0.120**
	(0.035)	(0.044)	(0.084)	(0.048)
	[0.100]	[0.046]	[0.000]	[0.018]
$[CA]$ Call-to-Action Ad	-0.039	-0.063	-0.095	-0.076
	(0.035)	(0.046)	(0.085)	(0.051)
	[0.250]	[0.172]	[0.276]	[0.160]
$[I + CA]$ Info + Call-to-Action Ad	-0.067**	-0.089**	-0.039	-0.053
	(0.033)	(0.044)	(0.091)	(0.049)
	[0.038]	[0.058]	[0.668]	[0.296]
Test $IA = CA$, p-value	0.56	0.50	0.02	0.42
Test $IA = I + CA$, p-value	0.83	0.94	0.01	0.20
Test $CA = I + CA$, p-value	0.42	0.56	0.56	0.68
Panel C. Subtreatments by Letter - No Letter				
$[NL]$ No Letter - Any Ad	-0.041	-0.058	-0.080	-0.065
	(0.035)	(0.046)	(0.088)	(0.050)
	[0.228]	[0.214]	[0.366]	[0.186]
$[L]$ Letter - Any Ad	-0.062**	-0.094**	-0.176**	-0.092**
	(0.029)	(0.039)	(0.073)	(0.041)
	[0.022]	[0.014]	[0.020]	[0.028]
Test $NL = L$, p-value	0.50	0.35	0.23	0.56
Control Mean	0.16	0.20	0.00	0.51
Sample Size	698	698	698	698

Notes: The outcome in column (1) is an indicator for whether any irregularity was reported in the media in a particular municipality. In column (2) it is the number of different irregularities. In column (3) it is the index of the χ^2, Kolmogorov-Smirnov and Kuiper test statistics testing for Benford's 2nd digit law, described in Section 3.2.2. In column (4) it is an indicator that takes the value of one if the p-value of any of these tests leads to rejection of the null hypothesis with less than a 5% significance level. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table 5: Impacts on Vote Share of Candidates Likely to Engage in Irregularities

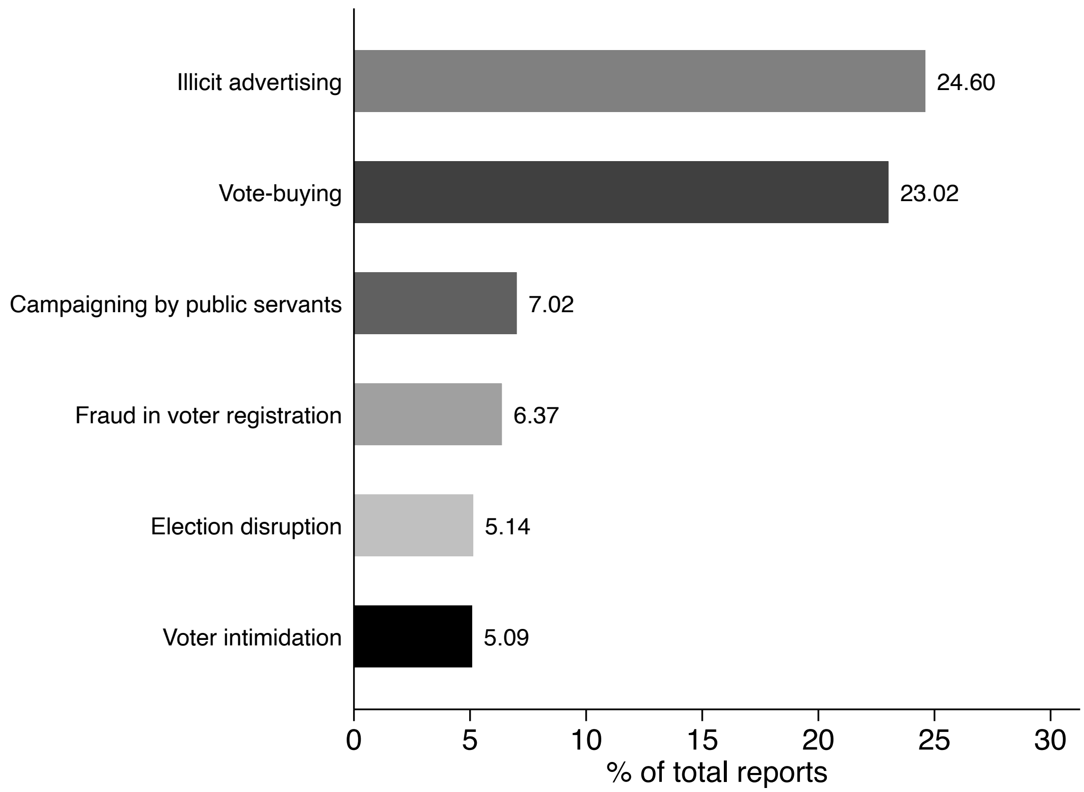
	(1)	(2)	(3)	(4)
	Vote Share (%)			
Interaction term Z:	Candidate will engage in irregularities (fraction of respondents)	Demeaned Candidate will engage in irregularities (fraction of respondents)	Above Average Candidate will engage in irregularities (=1)	Past Malfeasance (=1)
Panel A. Pooled Treatment				
$[T \times Z]$ Any treatment $\times Z$	-3.271**	-2.949*	-2.482***	-5.272*
	(1.513)	(1.715)	(0.832)	(3.056)
	[0.018]	[0.080]	[0.000]	[0.088]
	{0.030}	{0.084}	{0.003}	
Panel B. Subtreatments by Types of Ads				
$[IA \times Z]$ Information Ad$\times Z$	-4.719**	-4.812**	-2.510**	-9.093***
	(1.872)	(2.056)	(1.053)	(2.886)
	[0.018]	[0.036]	[0.028]	[0.018]
	{0.010}	{0.019}	{0.017}	
$[CA \times Z]$Call-to-Action Ad$\times Z$	-1.608	0.132	-1.799*	-2.847
	(1.922)	(2.329)	(1.075)	(4.802)
	[0.444]	[0.954]	[0.120]	[0.550]
	{0.395}	{0.953}	{0.085}	
$[I + CA \times Z]$Info + Call-to-Action Ad$\times Z$	-3.128	-3.854	-3.149***	-1.860
	(1.995)	(2.464)	(1.127)	(4.258)
	[0.086]	[0.086]	[0.002]	[0.672]
	{0.118}	{0.121}	{0.003}	
Test $IA \times Z = CA \times Z$, p-value	0.12	0.05	0.54	0.17
Test $IA \times Z = I + CA \times Z$, p-value	0.44	0.71	0.60	0.06
Test $CA \times Z = I + CA \times Z$, p-value	0.47	0.16	0.28	0.86
Panel C. Subtreatments by Letter - No Letter				
$[NL \times Z]$ No Letter - Any Ad $\times Z$	-4.406**	-3.241	-2.427**	-4.478
	(1.892)	(2.229)	(1.066)	(3.570)
	[0.032]	[0.162]	[0.014]	[0.216]
	{0.020}	{0.147}	{0.023}	
$[L \times Z]$ Letter - Any Ad $\times Z$	-2.637	-2.784	-2.501***	-5.914
	(1.637)	(1.907)	(0.907)	(3.574)
	[0.104]	[0.152]	[0.008]	[0.130]
	{0.108}	{0.153}	{0.007}	
Test $NL \times Z = L \times Z$, p-value	0.32	0.84	0.94	0.70
Control Mean	18.91	18.91	18.91	18.91
Sample Size	2989	2989	2989	263
N. Municipalities	630	630	630	48

Notes: The outcome in all columns is the vote share of each candidate, expressed as a percentage of total valid votes. In each of these columns, a different measure of the likelihood that a candidate commits irregularities is used to compute the candidate-level heterogeneous effects. In column (1) it is the proportion of respondents from the pre-treatment survey that say the candidate might commit at least one type of electoral irregularity. In column (2) the outcome is this same variable, demeaned using the municipality-level mean. In column (3) it is an indicator that takes the value of one if this variable is above the municipal-level mean. Finally, in column (4) it is an indicator for whether a candidate was found to be involved in malfeasance in the past according to the investigation by the NGO PARES. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Clustered standard errors at the municipal-level are shown in parentheses, randomization inference p-values are shown in square brackets, and clustered wild-bootstrap p-values correcting for the variance in estimating Z are shown in curly brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Supplementary Data and Appendix

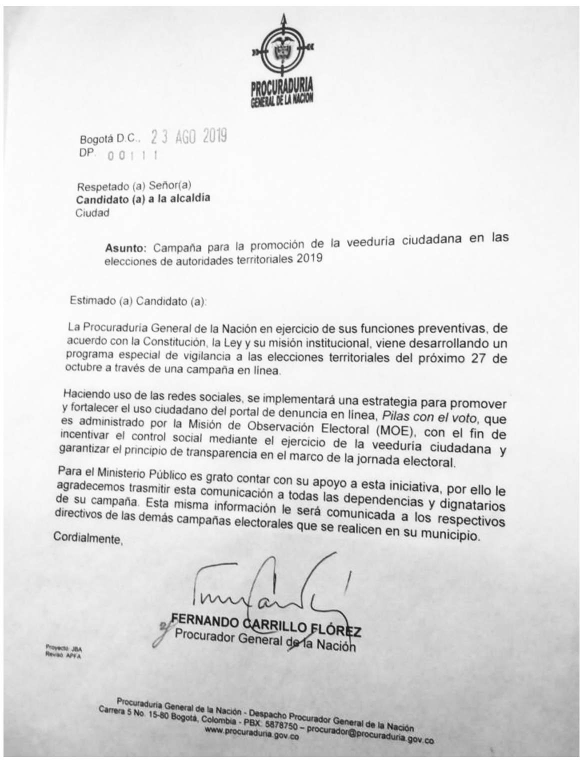
A Additional Figures and Tables

Figure A1: Electoral irregularities Reported in 2015 Elections to the MOE



Notes: This figure displays the proportion of electoral irregularities of different types as a percentage of total irregularities reported through the MOE's *Pilas con el voto* in the context of the 2015 mayoral elections. The reports are restricted to those received between October 22-26, 2015 (election day was on October 25). The definitions for each type of electoral irregularity are presented in Section 1.

Figure A2: Letter sent to candidates - Full Knowledge



PROCURADURIA
GENERAL DE LA NACION

Bogotá D.C., 23 AGO 2019
DP. 00111

Respetado (a) Señor(a)
Candidato (a) a la alcaldía
Ciudad

Asunto: Campaña para la promoción de la veeduría ciudadana en las elecciones de autoridades territoriales 2019

Estimado (a) Candidato (a):

La Procuraduría General de la Nación en ejercicio de sus funciones preventivas, de acuerdo con la Constitución, la Ley y su misión institucional, viene desarrollando un programa especial de vigilancia a las elecciones territoriales del próximo 27 de octubre a través de una campaña en línea.

Haciendo uso de las redes sociales, se implementará una estrategia para promover y fortalecer el uso ciudadano del portal de denuncia en línea, *Pilas con el voto*, que es administrado por la Misión de Observación Electoral (MOE), con el fin de incentivar el control social mediante el ejercicio de la veeduría ciudadana y garantizar el principio de transparencia en el marco de la jornada electoral.

Para el Ministerio Público es grato contar con su apoyo a esta iniciativa, por ello le agradecemos trasmitir esta comunicación a todas las dependencias y dignatarios de su campaña. Esta misma información le será comunicada a los respectivos directivos de las demás campañas electorales que se realicen en su municipio.

Cordialmente,

FERNANDO CARRILLO FLÓREZ
Procurador General de la Nación

Proyectó: JBA
Revisó: APFA

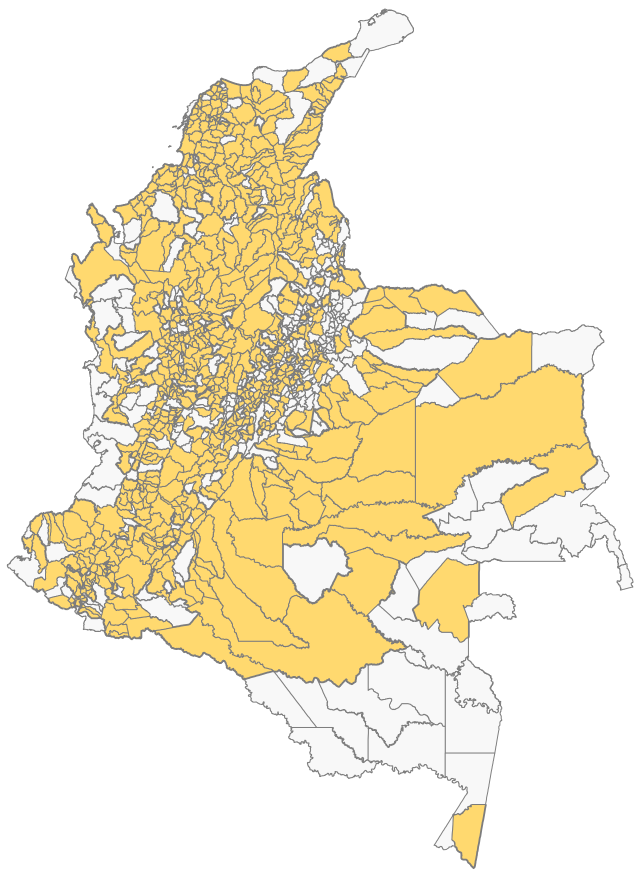
Procuraduría General de la Nación - Despacho Procurador General de la Nación
Carrera 5 No. 15-80 Bogotá, Colombia - PBX: 5878750 – procurador@procuraduria.gov.co
www.procuraduria.gov.co

Translation:

Respected Sir/Madam, Candidate to the Mayor's Office
Subject: *Campaign to promote citizens' oversight in the 2019 local elections*
The Attorney General of the Nation, in the exercise of its preventive functions, the Constitution, the Law and its Institutional Mission, is implementing a special program to watch over the forthcoming local elections of October 27 through an online campaign.
A strategy to promote and strengthen citizens' use of an online reporting website, Pilas con el voto, administered by the Misión de Observación Electoral will be set in place making use of social media. The goal of this strategy is to incentivize social control through citizen oversight and to guarantee transparency in the context of election day.
The Public Ministry welcomes your support, and thus we ask you to spread this information to your campaigns' offices and members. This same information will be communicated to the leaders of the other campaigns held in your municipality.[...]

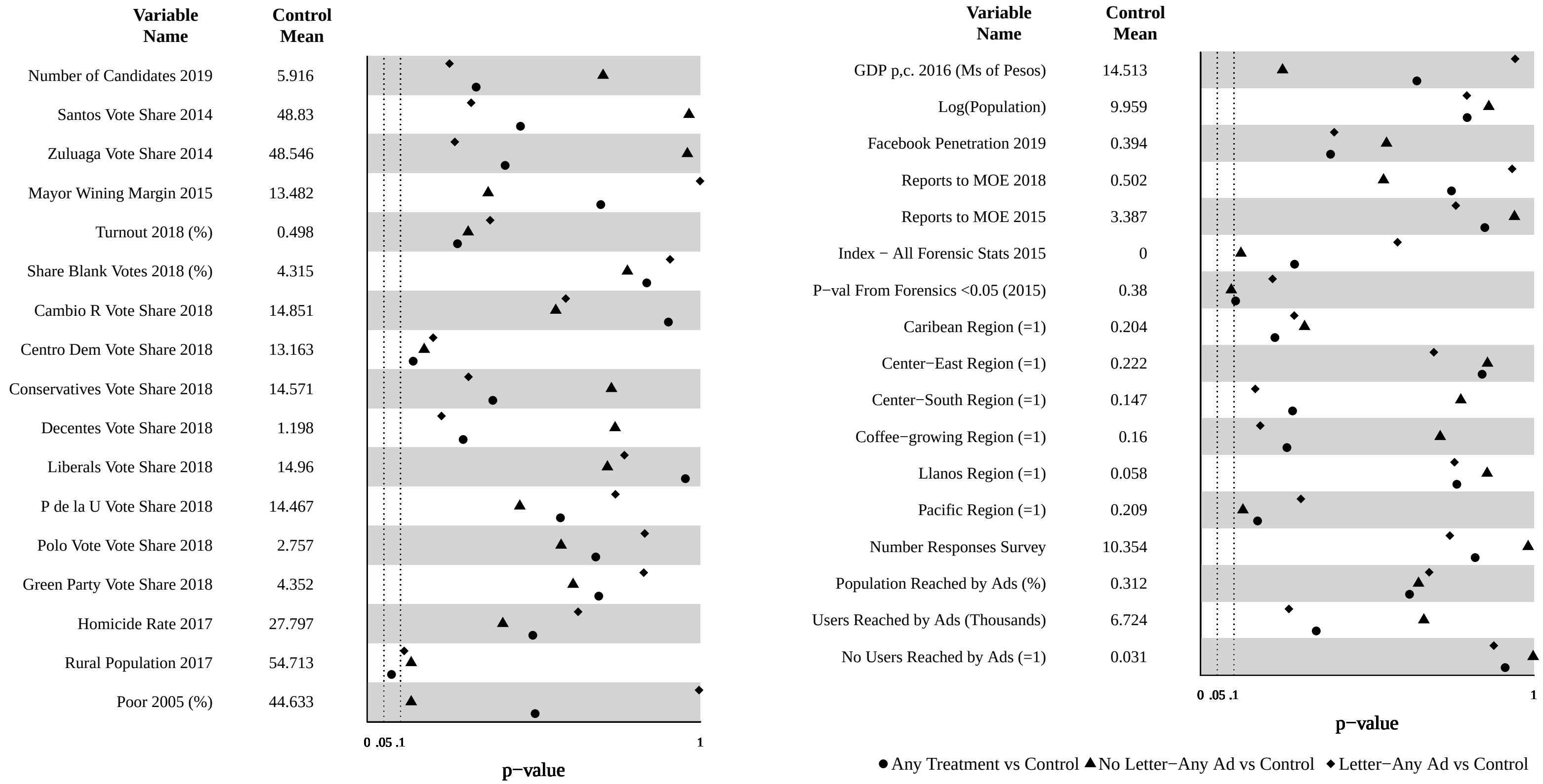
Notes: On the left, we show the actual letter sent to candidates in the full knowledge condition. On the right is a translation to English of the text contained in the letter.

Figure A3: Municipalities in the Study Sample



Notes: This figure shows a map of Colombia with the administrative boundaries of municipalities. Municipalities in yellow were part of the study sample; those in light grey are not in the sample.

Figure A4: Covariate Balance



Notes: This figure presents the p-values of the balance checks across three main treatment arms, using five sets of selected covariates: *Previous Reports:* reports to MOE 2018 and 2015; *Socioeconomic:* log of the population, Facebook penetration in 2019, GDP per capita 2016 (in Colombian pesos), percentage of poor in 2005, rural population in 2017, and homicide rate in 2017; *Political:* number of candidates registered in the 2019 mayoral elections, turnout in the 2015 mayoral elections, percentage of blank votes in 2015, the elected mayor's margin of victory in 2015, Santos' and Zuluaga's vote shares in the 2014 presidential elections, and the vote shares obtained by each party in the 2018 congressional elections; *Geographic:* regional dummies; and *Other:* number of responses in pre-treatment survey, percentage of population reached by the Facebook Ad, number of users reached by the Facebook Ad, and whether there were no users reached by the ad. The tests correspond to difference in means, in which observations are weighted by the percentage of the population older than 18 that was reached by a Facebook ad, except when reporting the difference in means of variables referring to the reach of the Facebook ads. The control group mean of each variable is presented.

Table A1: Summary Statistics Comparing Study Sample to Average Municipality

	Study Sample			All Municipalities in Country		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	Min	Max	Mean	Minimum	Maximum
Population 2018 (Thousands)	27.22	7.20	139.36	37.16	0.41	2529.40
Facebook Penetration 2018	0.41	0.02	2.21	0.61	-	-
Per Capita GDP 2016 (Millions of Pesos)	13.92	2.55	349.12	13.72	2.49	349.12
% Rural Population 2017	52.36	1.65	97.99	55.88	0.09	100.00
% Poor 2005	45.39	6.84	100.00	45.03	5.36	100.00
Reports to MOE 2018	0.52	0.00	8.00	0.80	0.00	45.00
Reports to MOE 2015	3.43	0.00	33.00	3.83	0.00	152.00
Sample size		698			1121	

Notes: This table displays summary statistics for the municipalities in the study sample (columns 1-3) and the full set of municipalities in Colombia (columns 4-6) on a selected group of variables. Data for the average Facebook penetration rate for the entire country is averaged across the whole population, not across municipalities.

Table A2: Covariate Balance

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Panel A. Previous Reports Covariates									
Reports to MOE 2018	0.502	0.026	0.080	-0.047	0.045	-0.056	-0.133	0.021	0.154
		(0.084)	(0.110)	(0.104)	(0.109)	(0.100)	(0.111)	(0.122)	(0.120)
Reports to MOE 2015	3.387	0.068	0.310	-0.284	0.173	0.152	0.227	0.079	-0.148
		(0.367)	(0.476)	(0.442)	(0.485)	(0.419)	(0.501)	(0.489)	(0.525)
Panel B. Socioeconomic Covariates									
Log(Population)	9.959	-0.014	0.027	-0.052	-0.019	-0.003	0.001	-0.006	-0.007
		(0.056)	(0.073)	(0.071)	(0.071)	(0.067)	(0.078)	(0.076)	(0.075)
Facebook Penetration 2019	0.394	0.020	0.030	0.029	0.002	0.005	-0.003	0.013	0.016
		(0.023)	(0.032)	(0.031)	(0.029)	(0.028)	(0.033)	(0.034)	(0.035)
GDP p.c. 2016 (Ms of Pesos)	14.513	-0.876	-1.092	-2.188	0.625	2.147	1.690	2.599	0.909
		(1.923)	(2.125)	(1.936)	(2.547)	(1.416)	(1.585)	(1.973)	(2.189)
Poor 2005 (%)	44.633	1.128	1.301	1.718	0.379	-3.344	-2.893	-3.789	-0.896
		(1.687)	(2.237)	(2.214)	(2.109)	(2.096)	(2.442)	(2.381)	(2.378)
Rural Population 2017	54.713	-3.471*	-6.156**	-3.238	-1.033	0.510	1.299	-0.269	-1.568
		(1.935)	(2.560)	(2.376)	(2.440)	(2.330)	(2.680)	(2.654)	(2.589)
Homicide Rate 2017	27.797	-1.734	-3.485	0.160	-1.851	1.167	-0.205	2.522	2.727
		(2.551)	(3.061)	(3.493)	(3.126)	(2.799)	(3.352)	(3.323)	(3.629)
Panel C. Political Covariates									
Number of Candidates 2019	5.916	-0.154	-0.093	-0.037	-0.331*	-0.117	-0.084	-0.151	-0.067
		(0.157)	(0.199)	(0.209)	(0.195)	(0.191)	(0.223)	(0.213)	(0.212)
Index - All Forensic Stats 2015	0.000	0.087	0.281**	-0.004	-0.016	-0.119	-0.126	-0.113	0.013
		(0.081)	(0.113)	(0.099)	(0.098)	(0.100)	(0.117)	(0.112)	(0.110)
P-val From Forensics < 0.05 2015	0.380	0.064	0.114**	0.024	0.054	-0.033	-0.034	-0.033	0.001

Continued on next page

Table A2 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
		(0.040)	(0.051)	(0.051)	(0.051)	(0.048)	(0.056)	(0.056)	(0.056)
Turnout 2018 (%)	0.498	0.008	0.017*	0.000	0.007	-0.003	-0.005	-0.001	0.004
		(0.008)	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)	(0.011)
Share Blank Votes 2018 (%)	4.315	0.069	0.341	-0.227	0.089	-0.075	-0.266	0.113	0.380
		(0.337)	(0.477)	(0.439)	(0.437)	(0.431)	(0.492)	(0.526)	(0.543)
Mayor Margin of Victory 2015	13.482	-0.365	0.128	-0.705	-0.522	1.068	0.329	1.798	1.468
		(0.950)	(1.311)	(1.119)	(1.166)	(1.080)	(1.187)	(1.326)	(1.294)
Santos Vote Share 2014	48.830	1.327	2.599	1.528	-0.136	1.868	1.586	2.147	0.561
		(1.795)	(2.243)	(2.249)	(2.283)	(2.006)	(2.417)	(2.278)	(2.433)
Zuluaga Vote Share 2014	48.546	-1.427	-2.684	-1.670	0.059	-2.007	-1.661	-2.349	-0.688
		(1.745)	(2.174)	(2.180)	(2.220)	(1.942)	(2.339)	(2.204)	(2.351)
Liberals Vote Share 2018	14.960	0.051	-0.780	0.407	0.525	0.683	-0.053	1.410	1.464
		(0.896)	(1.123)	(1.196)	(1.140)	(1.060)	(1.201)	(1.278)	(1.286)
Cambio R Vote Share 2018	14.851	-0.126	0.862	-1.163	-0.091	-1.362	-1.284	-1.439	-0.156
		(1.047)	(1.347)	(1.321)	(1.285)	(1.221)	(1.416)	(1.381)	(1.359)
Centro Dem Vote Share 2018	13.163	-1.292	-1.500	-1.675	-0.709	0.208	1.218	-0.788	-2.005*
		(0.870)	(1.081)	(1.083)	(1.116)	(0.951)	(1.189)	(1.051)	(1.193)
P de la U Vote Share 2018	14.467	0.506	1.105	-0.133	0.538	-0.542	-0.589	-0.496	0.092
		(0.913)	(1.173)	(1.198)	(1.163)	(1.103)	(1.251)	(1.311)	(1.301)
Green Party Vote Share 2018	4.352	0.183	0.309	0.544	-0.296	-0.223	-0.337	-0.110	0.226
		(0.466)	(0.599)	(0.720)	(0.539)	(0.648)	(0.759)	(0.705)	(0.680)
Polo Vote Share 2018	2.757	-0.100	-0.222	0.206	-0.280	0.135	0.016	0.254	0.238
		(0.248)	(0.269)	(0.394)	(0.260)	(0.311)	(0.322)	(0.366)	(0.296)
Decentes Vote Share 2018	1.198	0.108	0.127	0.020	0.176	0.099	-0.005	0.201	0.206
		(0.102)	(0.123)	(0.153)	(0.148)	(0.141)	(0.147)	(0.184)	(0.179)
Panel D. Geographic Covariates									

Continued on next page

Table A2 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Caribean Region (=1)	0.204	0.041	0.093**	0.014	0.016	-0.005	-0.023	0.013	0.035
		(0.033)	(0.045)	(0.043)	(0.043)	(0.042)	(0.048)	(0.049)	(0.049)
Center-East Region (=1)	0.222	-0.007	0.006	-0.024	-0.002	-0.021	-0.030	-0.013	0.017
		(0.034)	(0.043)	(0.042)	(0.043)	(0.040)	(0.046)	(0.047)	(0.046)
Center-South Region (=1)	0.147	-0.030	-0.039	-0.044	-0.008	-0.031	-0.027	-0.035	-0.008
		(0.028)	(0.034)	(0.034)	(0.036)	(0.032)	(0.037)	(0.036)	(0.035)
Coffee-growing Region (=1)	0.160	0.035	0.011	0.071*	0.022	0.031	0.033	0.030	-0.003
		(0.031)	(0.039)	(0.042)	(0.039)	(0.038)	(0.044)	(0.044)	(0.046)
Llanos Region (=1)	0.058	0.006	0.006	-0.019	0.030	0.002	0.009	-0.005	-0.014
		(0.019)	(0.025)	(0.022)	(0.027)	(0.024)	(0.028)	(0.027)	(0.028)
Pacific Region (=1)	0.209	-0.044	-0.076**	0.003	-0.058	0.024	0.038	0.010	-0.028
		(0.032)	(0.038)	(0.043)	(0.039)	(0.035)	(0.042)	(0.041)	(0.043)
Panel E. Other Covariates									
Number Responses Survey	10.354	-0.130	-0.313	1.232	-1.251*	-0.220	-0.130	-0.307	-0.177
		(0.581)	(0.757)	(0.847)	(0.657)	(0.722)	(0.805)	(0.875)	(0.858)
Population Reached by Ads (%)	0.312	0.007	0.009	0.011	0.002	-0.002	0.001	-0.005	-0.007
		(0.015)	(0.020)	(0.020)	(0.020)	(0.019)	(0.023)	(0.022)	(0.024)
Users Reached by Ads (Thousands)	6.724	-0.707	-0.172	-1.053	-0.900	-0.491	-0.456	-0.525	-0.069
		(0.751)	(0.949)	(0.826)	(0.882)	(0.717)	(0.833)	(0.812)	(0.804)
No Users Reached by Ads (=1)	0.031	-0.002	-0.012	0.014	-0.006	-0.002	0.001	-0.006	-0.007
		(0.014)	(0.016)	(0.020)	(0.017)	(0.017)	(0.020)	(0.019)	(0.019)

Notes: This table presents the balance checks for a selected set of covariates. The control group mean of each variable is presented in column (1). In each of the remaining columns the difference in means is reported for the shown treatment groups. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A3: Balance on Pre-Treat Survey Respondent Characteristics

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Female(=1)	0.507	0.008	-0.008	0.010	0.022	-0.014	-0.005	-0.022	-0.017
		(0.014)	(0.017)	(0.018)	(0.020)	(0.017)	(0.020)	(0.019)	(0.020)
Age	34.496	-0.206	-0.543	-0.185	0.142	0.979*	0.884	1.072*	0.188
		(0.379)	(0.531)	(0.494)	(0.477)	(0.500)	(0.567)	(0.585)	(0.571)
High School or Less (=1)	0.488	0.019	0.018	0.016	0.025	0.031*	0.032	0.030	-0.002
		(0.016)	(0.021)	(0.020)	(0.020)	(0.018)	(0.020)	(0.022)	(0.022)

Notes: This table presents the balance checks for a set of survey respondent characteristics. The control group mean of each variable is presented in column (1). In each of the remaining columns the difference in means is reported for the shown treatment groups. Clustered standard errors at the municipal-level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A4: Impacts on Reports After the Intervention

	(1) Reports After Intervention(=1)	(2) N. Reports After Intervention	(3) High Quality Reports After Intervention (=1)	(4) High Quality N. Reports After Intervention
Panel A. Pooled Treatment				
[T] Any treatment	0.032	0.062	0.016	0.021
	(0.024)	(0.042)	(0.020)	(0.035)
	[0.224]	[0.174]	[0.536]	[0.588]
Panel B. Subtreatments by Types of Ad				
[IA] Information Ad	0.049	0.096	0.014	0.031
	(0.033)	(0.060)	(0.027)	(0.048)
	[0.136]	[0.108]	[0.670]	[0.574]
[CA] Call-to-Action Ad	0.006	-0.017	-0.005	-0.031
	(0.030)	(0.043)	(0.025)	(0.034)
	[0.842]	[0.676]	[1.000]	[0.506]
[$I + CA$] Info + Call-to-Action Ad	0.042	0.106*	0.038	0.062
	(0.033)	(0.064)	(0.029)	(0.051)
	[0.236]	[0.060]	[0.170]	[0.232]
Test IA = CA, p-value	0.23	0.05	0.52	0.14
Test IA = $I + CA$, p-value	0.84	0.89	0.44	0.59
Test CA = $I + CA$, p-value	0.32	0.05	0.16	0.04
Panel C. Subtreatments by Letter - No Letter				
[NL] No Letter - Any Ad	0.003	0.047	0.006	0.010
	(0.030)	(0.058)	(0.026)	(0.044)
	[1.000]	[0.414]	[0.834]	[0.916]
[L] Letter - Any Ad	0.048*	0.070	0.021	0.026
	(0.027)	(0.045)	(0.023)	(0.038)
	[0.096]	[0.148]	[0.396]	[0.568]
Test NL = L, p-value	0.13	0.70	0.55	0.70
Control Mean	0.09	0.12	0.06	0.09
Sample Size	698	698	698	698

Notes: The outcome in column (1) is an indicator of whether any report was issued to the MOE from each municipality in the month after the intervention. In column (2) it is the number of such reports. In columns (3)-(4) the same definitions are used on the subset of reports of a high quality (see Section 1 for a discussion about how quality of reports is assessed by the MOE). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A5: Robustness of the Impacts on the Media-Based Irregularity Measures: Including News Coming From MOE Reports

	(1) Media Irregularities (=1)	(2) Number of Media Irregularities
Panel A. Pooled Treatment		
$[T]$ Any treatment	-0.051*	-0.070*
	(0.028)	(0.038)
	[0.046]	[0.054]
Panel B. Subtreatments by Types of Ad		
$[IA]$ Information Ad	-0.058	-0.081*
	(0.035)	(0.048)
	[0.100]	[0.104]
$[CA]$ Call-to-Action Ad	-0.035	-0.048
	(0.035)	(0.046)
	[0.310]	[0.324]
$[I+CA]$ Info + Call-to-Action Ad	-0.060*	-0.079*
	(0.034)	(0.045)
	[0.104]	[0.098]
Test $IA = CA$, p-value	0.54	0.49
Test $IA = I+CA$, p-value	0.95	0.97
Test $CA = I+CA$, p-value	0.48	0.48
Panel C. Subtreatments by Letter - No Letter		
$[NL]$ No Letter - Any Ad	-0.029	-0.055
	(0.036)	(0.046)
	[0.392]	[0.242]
$[L]$ Letter - Any Ad	-0.062**	-0.077*
	(0.029)	(0.040)
	[0.044]	[0.032]
Test $NL = L$, p-value	0.31	0.57
Control Mean	0.16	0.20
Sample Size	698	698

Notes: The outcome in column (1) is an indicator of whether any irregularity was reported in the media in a particular municipality. In column (2) it is the number of different irregularities. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A6: Impacts on Media-Based Irregularity Measures - By Type of Irregularity

	(1) Vote buying (=1)	(2) Riot (=1)	(3) Candidate intimidation (=1)	(4) Voter intimidation (=1)	(5) Registration fraud (=1)	(6) Public servant campaigning (=1)	(7) Electoral fraud (=1)	(8) Others (=1)
Panel A. Pooled Treatment								
[T] Any treatment	-0.041***	-0.011	-0.012	-0.003	0.012	-0.007	0.004	-0.001
	(0.015)	(0.014)	(0.013)	(0.010)	(0.009)	(0.008)	(0.008)	(0.011)
	[0.002]	[0.458]	[0.420]	[1.000]	[0.356]	[0.372]	[0.708]	[1.000]
Panel B. Subtreatments by Types of Ad								
[IA] Information Ad	-0.042***	-0.004	-0.018	0.001	0.016	-0.013*	-0.003	-0.005
	(0.016)	(0.019)	(0.015)	(0.014)	(0.014)	(0.008)	(0.009)	(0.013)
	[0.030]	[0.914]	[0.340]	[1.000]	[0.234]	[0.272]	[1.000]	[1.000]
[CA] Call-to-Action Ad	-0.030*	-0.014	-0.012	-0.005	0.010	0.006	0.004	0.001
	(0.018)	(0.017)	(0.016)	(0.013)	(0.013)	(0.013)	(0.011)	(0.014)
	[0.114]	[0.524]	[0.512]	[1.000]	[0.680]	[0.682]	[1.000]	[1.000]
[$I + CA$] Info + Call-to-Action Ad	-0.049***	-0.016	-0.006	-0.005	0.010	-0.013*	0.010	0.001
	(0.014)	(0.016)	(0.017)	(0.013)	(0.013)	(0.008)	(0.013)	(0.014)
	[0.002]	[0.380]	[0.796]	[1.000]	[0.620]	[0.272]	[0.646]	[1.000]
Test $IA = CA$, p-value	0.34	0.55	0.64	0.66	0.72	0.08	0.56	0.64
Test $IA = I + CA$, p-value	0.31	0.48	0.41	0.65	0.70	.	0.32	0.66
Test $CA = I + CA$, p-value	0.09	0.90	0.72	0.98	0.98	0.08	0.67	0.98
Panel C. Subtreatments by Letter - No Letter								
[NL] No Letter - Any Ad	-0.042***	-0.004	-0.019	0.001	0.022	-0.007	0.004	0.007
	(0.016)	(0.018)	(0.015)	(0.014)	(0.015)	(0.010)	(0.011)	(0.015)
	[0.022]	[0.892]	[0.282]	[1.000]	[0.138]	[0.666]	[1.000]	[0.716]
[L] Letter - Any Ad	-0.040**	-0.015	-0.009	-0.005	0.007	-0.007	0.004	-0.005
	(0.015)	(0.015)	(0.014)	(0.011)	(0.009)	(0.009)	(0.009)	(0.011)
	[0.008]	[0.318]	[0.592]	[0.750]	[0.702]	[0.676]	[0.704]	[0.734]
Test $NL = L$, p-value	0.74	0.49	0.41	0.64	0.33	0.98	0.97	0.39
Control Mean	0.05	0.04	0.03	0.02	0.01	0.01	0.01	0.02
Sample Size	698	698	698	698	698	698	698	698

Notes: The outcomes in columns (1)-(8) are indicators of whether each of the types of irregularities displayed were reported in the media. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A7: Robustness of the Media-Based Irregularity Measures: Leave-one-out Estimates

	(1) Any irregularity except vote buying (=1)	(2) Any irregularity except riots (=1)	(3) Any irregularity except candidate intimidation (=1)	(4) Any irregularity except voter intimidation (=1)	(5) Any irregularity except registration fraud (=1)	(6) Any irregularity except public servant campaigning (=1)	(7) Any irregularity except electoral fraud (=1)	(8) Any irregularity except others (=1)
Panel A. Pooled Treatment								
[T] Any treatment	-0.024	-0.039	-0.042*	-0.058**	-0.066**	-0.049*	-0.052*	-0.054**
	(0.026)	(0.026)	(0.025)	(0.027)	(0.027)	(0.027)	(0.027)	(0.026)
	[0.336]	[0.122]	[0.110]	[0.022]	[0.014]	[0.048]	[0.048]	[0.036]
Panel B. Subtreatments by Types of Ad								
[IA] Information Ad	-0.021	-0.045	-0.042	-0.068**	-0.067**	-0.051	-0.056*	-0.055*
	(0.033)	(0.032)	(0.032)	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)
	[0.648]	[0.164]	[0.212]	[0.040]	[0.072]	[0.126]	[0.108]	[0.080]
[CA] Call-to-Action Ad	-0.019	-0.029	-0.021	-0.043	-0.052	-0.037	-0.036	-0.037
	(0.033)	(0.033)	(0.032)	(0.034)	(0.034)	(0.034)	(0.034)	(0.033)
	[0.642]	[0.424]	[0.508]	[0.202]	[0.168]	[0.306]	[0.288]	[0.252]
[$I+CA$] Info + Call-to-Action Ad	-0.033	-0.044	-0.061**	-0.064**	-0.079**	-0.058*	-0.064**	-0.069**
	(0.032)	(0.032)	(0.029)	(0.032)	(0.032)	(0.033)	(0.032)	(0.031)
	[0.360]	[0.184]	[0.058]	[0.050]	[0.026]	[0.078]	[0.060]	[0.046]
Test $IA = CA$, p-value	0.94	0.64	0.54	0.45	0.64	0.68	0.58	0.59
Test $IA = I+CA$, p-value	0.72	0.99	0.53	0.90	0.70	0.84	0.80	0.68
Test $CA = I+CA$, p-value	0.66	0.66	0.20	0.52	0.40	0.53	0.41	0.33
Panel C. Subtreatments by Letter - No Letter								
[NL] No Letter - Any Ad	-0.010	-0.027	-0.017	-0.044	-0.062*	-0.039	-0.038	-0.042
	(0.034)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)
	[0.876]	[0.434]	[0.616]	[0.190]	[0.104]	[0.246]	[0.296]	[0.206]
[L] Letter - Any Ad	-0.032	-0.046*	-0.054**	-0.066**	-0.068**	-0.054*	-0.060**	-0.060**
	(0.028)	(0.027)	(0.026)	(0.028)	(0.029)	(0.029)	(0.029)	(0.028)
	[0.244]	[0.102]	[0.042]	[0.018]	[0.016]	[0.066]	[0.024]	[0.040]
Test $NL = L$, p-value	0.47	0.51	0.21	0.46	0.81	0.63	0.48	0.54
Control Mean	0.13	0.13	0.13	0.15	0.16	0.15	0.15	0.14
Sample Size	698	698	698	698	698	698	698	698

Notes: The outcome in each column is an indicator of whether any irregularity was reported in the media in a particular municipality, when one leaves out each of the types of irregularities shown in columns (1)-(8). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A8: Correlation Between Forensic and Media-Based Electoral Irregularity Measures

	(1)	(2)	(3)	(4)
	Index of all Forensic Test Stats (z-score)		Any P-value From Forensic Tests < 0.05 (=1) (z-score)	
Media Irregularities (=1)	0.102		0.099*	
	(0.073)		(0.058)	
Number of Media Irregularities		0.133		0.130*
		(0.095)		(0.075)
Sample Size	225	225	225	225

Notes: This table presents the OLS results of regressing z-scores of the forensic variables on z-scores of the media-based measure of irregularities detailed in Table 4. Each estimate comes from a separate regression of the shown variables. The sample is restricted to municipalities in the control group. Since all of these variables are normalized, the estimates reported can be interpreted as correlations. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A9: Correlation Between Forensic and Media-Based Electoral Irregularity Measures by Type

	(1)	(2)
	Index of all Forensic Test Stats (z-score)	Any P-value From Forensic Tests < 0.05 (=1) (z-score)
Vote Buying in Media (=1) (z-score)	0.009	0.012
	(0.067)	(0.050)
Riot in Media (=1) (z-score)	0.232***	0.161***
	(0.078)	(0.011)
Candidate Intimidation in Media (=1) (z-score)	-0.024	0.018
	(0.053)	(0.056)
Voter Intimidation in Media (=1) (z-score)	0.079	0.054
	(0.089)	(0.049)
Registration Fraud in Media (=1) (z-score)	0.089***	0.128***
	(0.026)	(0.009)
Public Servant Campaigning in Media (=1) (z-score)	-0.071***	-0.082***
	(0.010)	(0.005)
Electoral Fraud in Media (=1) (z-score)	0.093***	0.102***
	(0.024)	(0.007)
Sample Size	225	225

Notes: This table presents the OLS results of regressing z-scores of the forensic variables on z-scores of the media-based measure of irregularities by type, shown in Table A6. The sample is restricted to municipalities in the control group. Since all of these variables are normalized, the estimates reported can be interpreted as correlations. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A10: Impacts on Deviations from Benford's Second Digit Law - By Test

	(1)	(2)	(3)	(4)	(5)	(6)
Test:	Pearson χ^2		Kolmogorov-Smirnov D		Kuiper V	
	Stat (z-score)	P-value < 0.05 (=1)	Stat (z-score)	P-value < 0.05 (=1)	Stat (z-score)	P-value < 0.05 (=1)
Panel A. Pooled Treatment						
[T] Any treatment	-0.090	-0.018	-0.115*	-0.072*	-0.174***	-0.059*
	(0.074)	(0.035)	(0.068)	(0.038)	(0.060)	(0.035)
	[0.206]	[0.588]	[0.090]	[0.066]	[0.002]	[0.082]
Panel B. Subtreatments by Types of Ad						
[IA] Information Ad	-0.168*	-0.059	-0.296***	-0.104**	-0.278***	-0.091**
	(0.092)	(0.045)	(0.082)	(0.048)	(0.075)	(0.044)
	[0.080]	[0.188]	[0.000]	[0.038]	[0.000]	[0.036]
[CA] Call-to-Action Ad	-0.083	-0.023	-0.058	-0.062	-0.160**	-0.039
	(0.090)	(0.046)	(0.085)	(0.050)	(0.073)	(0.046)
	[0.366]	[0.614]	[0.490]	[0.208]	[0.040]	[0.372]
[$I+CA$] Info + Call-to-Action Ad	-0.021	0.025	0.006	-0.050	-0.088	-0.047
	(0.089)	(0.045)	(0.090)	(0.048)	(0.077)	(0.045)
	[0.856]	[0.600]	[0.956]	[0.286]	[0.260]	[0.316]
Test $IA = CA$, p-value	0.35	0.45	0.01	0.42	0.13	0.29
Test $IA = I+CA$, p-value	0.10	0.08	0.00	0.28	0.02	0.36
Test $CA = I+CA$, p-value	0.49	0.32	0.51	0.82	0.36	0.87
Panel C. Subtreatments by Letter - No Letter						
[NL] No Letter - Any Ad	-0.074	-0.015	-0.025	-0.036	-0.137*	-0.059
	(0.093)	(0.045)	(0.088)	(0.049)	(0.077)	(0.044)
	[0.414]	[0.730]	[0.758]	[0.426]	[0.070]	[0.196]
[L] Letter - Any Ad	-0.098	-0.020	-0.162**	-0.090**	-0.194***	-0.059
	(0.077)	(0.038)	(0.073)	(0.041)	(0.064)	(0.038)
	[0.204]	[0.606]	[0.022]	[0.030]	[0.002]	[0.130]
Test $NL = L$, p-value	0.77	0.91	0.09	0.23	0.42	0.99
Control Mean	0.00	0.28	0.00	0.43	0.00	0.33
Sample Size	698	698	698	698	698	698

Notes: This table reports the effects of the intervention on the χ^2, Kolmogorov-Smirnov and Kuiper test statistics testing for Benford's 2nd digit law. For each test, the outcome is the test statistic, as well as indicators that take the value of one if the p-value of each test leads to rejection of the null hypothesis with less than a 10% or 5% significance level. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A11: Impacts on Forensic Tests Suggested by Beber and Scacco (2012)

	(1) Index of all Last Digit Forensic Test Stats (z-score)	(2) Any P-value From Last Digit Forensic Tests < 0.05 (=1)	(3) Repeated Digits Less than Expected (=1)	(4) Adjacent Pairs of Digits More than Expected (=1)
Panel A. Pooled Treatment				
[T] Any treatment	-0.117	0.002	-0.049*	-0.037
	(0.073)	(0.032)	(0.029)	(0.031)
	[0.106]	[0.958]	[0.114]	[0.234]
Panel B. Subtreatments by Types of Ad				
[IA] Information Ad	-0.278***	-0.075*	-0.086**	0.012
	(0.085)	(0.039)	(0.039)	(0.040)
	[0.000]	[0.078]	[0.014]	[0.734]
[CA] Call-to-Action Ad	-0.028	0.048	-0.002	-0.080**
	(0.091)	(0.043)	(0.036)	(0.039)
	[0.764]	[0.282]	[0.974]	[0.056]
[$I+CA$] Info + Call-to-Action Ad	-0.050	0.031	-0.060	-0.044
	(0.093)	(0.042)	(0.040)	(0.041)
	[0.586]	[0.476]	[0.112]	[0.282]
Test $IA = CA$, p-value	0.00	0.01	0.05	0.03
Test $IA = I + CA$, p-value	0.01	0.02	0.57	0.21
Test $CA = I + CA$, p-value	0.82	0.71	0.18	0.40
Panel C. Subtreatments by Letter - No Letter				
[NL] No Letter - Any Ad	-0.079	0.025	-0.078**	-0.044
	(0.095)	(0.041)	(0.039)	(0.039)
	[0.392]	[0.516]	[0.040]	[0.268]
[L] Letter - Any Ad	-0.137*	-0.011	-0.034	-0.034
	(0.076)	(0.035)	(0.032)	(0.034)
	[0.044]	[0.734]	[0.284]	[0.330]
Test $NL = L$, p-value	0.49	0.36	0.25	0.79
Control Mean	-0.00	0.52	0.84	0.25
Sample Size	698	698	698	698

Notes: The outcome in column (1) is the index of the χ^2, Kolmogorov-Smirnov and Kuiper test statistics for deviations of the last digit from a uniform distribution, described in Appendix E . In column (2) it is an indicator that takes the value of one if the p-value of any of these tests leads to rejection of the null hypothesis with less than a 5% significance level. In column (3) it is an indicator that takes the value of one if there is a smaller than expected proportion of repeated digits, according to the distributions derived by Beber and Scacco (2012). In column (4) it is an indicator that takes the value of one if there is a larger than expected proportion of pairs of adjacent digits, according to the distributions derived by Beber and Scacco (2012). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A12: Covariate Balance - Candidate Level Data

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Panel A. Previous Reports Covariates									
Reports to MOE 2018	0.534	0.089	0.182	-0.043	0.128	-0.061	-0.099	-0.025	0.074
		(0.105)	(0.145)	(0.126)	(0.136)	(0.128)	(0.153)	(0.147)	(0.156)
Reports to MOE 2015	3.541	0.491	0.915	0.175	0.371	-0.017	0.101	-0.133	-0.235
		(0.440)	(0.605)	(0.593)	(0.565)	(0.563)	(0.661)	(0.670)	(0.710)
Panel B. Socioeconomic Covariates									
Log(Population)	10.107	-0.021	0.013	-0.063	-0.015	-0.027	-0.006	-0.047	-0.042
		(0.066)	(0.089)	(0.083)	(0.083)	(0.081)	(0.094)	(0.092)	(0.090)
Facebook Penetration 2019	0.448	0.028	0.036	0.045	0.003	0.000	0.007	-0.006	-0.012
		(0.028)	(0.038)	(0.035)	(0.035)	(0.034)	(0.039)	(0.039)	(0.040)
GDP p.c. 2016 (Ms of Pesos)	15.236	-0.458	-1.054	-1.497	1.217	1.561	1.243	1.873	0.630
		(2.042)	(2.237)	(2.108)	(2.729)	(1.658)	(1.841)	(2.153)	(2.251)
Poor 2005 (%)	43.006	0.044	0.127	0.436	-0.442	-2.773	-4.419*	-1.166	3.253
		(1.880)	(2.541)	(2.431)	(2.299)	(2.371)	(2.660)	(2.707)	(2.533)
Rural Population 2017	49.379	-1.289	-3.580	-1.459	1.262	1.990	0.658	3.292	2.634
		(2.195)	(2.916)	(2.666)	(2.748)	(2.609)	(2.973)	(2.978)	(2.869)
Homicide Rate 2017	28.994	-1.823	-2.340	-1.015	-2.110	1.158	0.106	2.186	2.079
		(2.617)	(3.291)	(3.315)	(3.217)	(2.915)	(3.307)	(3.462)	(3.444)
Panel C. Political Covariates									
Number of Candidates 2019	6.744	-0.206	-0.137	-0.147	-0.335	-0.349	-0.317	-0.381	-0.064
		(0.199)	(0.257)	(0.279)	(0.246)	(0.260)	(0.303)	(0.287)	(0.281)
Index - All Forensic Stats 2015	0.000	0.087	0.281**	-0.004	-0.016	-0.119	-0.126	-0.113	0.013
		(0.081)	(0.113)	(0.099)	(0.098)	(0.100)	(0.117)	(0.112)	(0.110)
P-val From Forensics < 0.05 2015	0.380	0.064	0.114**	0.024	0.054	-0.033	-0.034	-0.033	0.001

Continued on next page

Table A12 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
		(0.040)	(0.051)	(0.051)	(0.051)	(0.048)	(0.056)	(0.056)	(0.056)
Turnout 2018 (%)	0.494	0.003	0.009	-0.001	0.001	-0.009	-0.010	-0.007	0.003
		(0.007)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.012)	(0.011)
Share Blank Votes 2018 (%)	4.684	0.141	0.322	-0.174	0.275	-0.048	-0.106	0.009	0.116
		(0.421)	(0.557)	(0.554)	(0.560)	(0.521)	(0.612)	(0.626)	(0.668)
Mayor Margin of Victory 2015	13.455	0.076	0.293	-0.256	0.187	0.399	0.125	0.667	0.542
		(1.064)	(1.495)	(1.280)	(1.268)	(1.246)	(1.381)	(1.505)	(1.462)
Santos Vote Share 2014	48.549	-0.302	0.029	-0.197	-0.751	2.061	-0.107	4.180*	4.287*
		(1.870)	(2.310)	(2.360)	(2.408)	(2.088)	(2.498)	(2.363)	(2.524)
Zuluaga Vote Share 2014	48.734	0.051	-0.316	-0.022	0.505	-2.136	-0.012	-4.212*	-4.199*
		(1.819)	(2.238)	(2.287)	(2.344)	(2.016)	(2.420)	(2.277)	(2.443)
Liberals Vote Share 2018	15.064	-0.377	-1.270	-0.156	0.325	0.821	-0.311	1.926	2.236
		(0.935)	(1.113)	(1.229)	(1.292)	(1.111)	(1.254)	(1.358)	(1.384)
Cambio R Vote Share 2018	15.696	-0.802	-0.321	-1.173	-0.924	-1.088	-1.451	-0.734	0.717
		(1.161)	(1.435)	(1.501)	(1.398)	(1.312)	(1.432)	(1.560)	(1.448)
Centro Dem Vote Share 2018	13.710	-0.340	-0.255	-0.787	0.027	-0.341	1.643	-2.280*	-3.924***
		(0.955)	(1.263)	(1.221)	(1.249)	(1.231)	(1.440)	(1.320)	(1.290)
P de la U Vote Share 2018	14.859	-0.149	0.014	0.404	-0.880	-0.474	-0.785	-0.169	0.616
		(0.937)	(1.184)	(1.297)	(1.139)	(1.110)	(1.251)	(1.359)	(1.381)
Green Party Vote Share 2018	4.365	0.351	0.795	0.129	0.115	-0.557	-0.597	-0.519	0.078
		(0.440)	(0.634)	(0.596)	(0.536)	(0.616)	(0.704)	(0.688)	(0.646)
Polo Vote Share 2018	2.988	-0.110	-0.206	0.157	-0.282	0.391	0.249	0.531	0.283
		(0.295)	(0.326)	(0.392)	(0.315)	(0.250)	(0.262)	(0.335)	(0.336)
Decentes Vote Share 2018	1.289	0.166	0.118	0.141	0.243	0.113	0.007	0.217	0.210
		(0.117)	(0.127)	(0.211)	(0.167)	(0.171)	(0.165)	(0.242)	(0.235)
Panel D. Geographic Covariates									

Continued on next page

Table A12 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Caribean Region (=1)	0.227	-0.013	0.020	-0.006	-0.053	-0.003	-0.041	0.033	0.074
		(0.039)	(0.050)	(0.051)	(0.047)	(0.045)	(0.049)	(0.055)	(0.052)
Center-East Region (=1)	0.222	0.001	-0.006	-0.020	0.030	0.002	0.002	0.002	-0.001
		(0.038)	(0.048)	(0.047)	(0.051)	(0.045)	(0.053)	(0.052)	(0.053)
Center-South Region (=1)	0.169	-0.039	-0.042	-0.066*	-0.008	-0.027	-0.008	-0.045	-0.037
		(0.033)	(0.042)	(0.039)	(0.045)	(0.040)	(0.047)	(0.044)	(0.043)
Coffee-growing Region (=1)	0.153	0.059*	0.044	0.101**	0.031	0.037	0.052	0.022	-0.030
		(0.034)	(0.045)	(0.048)	(0.044)	(0.045)	(0.054)	(0.051)	(0.054)
Llanos Region (=1)	0.071	0.012	0.016	-0.012	0.033	-0.006	0.013	-0.025	-0.038
		(0.025)	(0.035)	(0.031)	(0.034)	(0.034)	(0.042)	(0.036)	(0.037)
Pacific Region (=1)	0.158	-0.021	-0.032	0.004	-0.034	-0.003	-0.018	0.013	0.031
		(0.032)	(0.040)	(0.041)	(0.039)	(0.036)	(0.040)	(0.043)	(0.041)
Panel E. Other Covariates									
Number Responses Survey	10.945	-0.221	-0.461	1.177	-1.394**	-0.401	0.020	-0.813	-0.833
		(0.607)	(0.760)	(0.919)	(0.683)	(0.775)	(0.823)	(0.962)	(0.901)
Population Reached by Ads (%)	0.346	0.011	0.009	0.020	0.003	-0.001	0.007	-0.009	-0.016
		(0.016)	(0.020)	(0.021)	(0.020)	(0.020)	(0.023)	(0.023)	(0.024)
Users Reached by Ads (Thousands)	8.758	-0.950	-0.239	-1.377	-1.254	-0.506	-0.158	-0.846	-0.689
		(1.054)	(1.376)	(1.132)	(1.229)	(1.009)	(1.234)	(1.101)	(1.179)
No Users Reached by Ads (=1)	0.000	0.002	0.000***	0.003	0.002	0.000	-0.001	0.002	0.003
		(0.001)	(0.000)	(0.003)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)
Panel F. Candidate Level Covariates									
Candidate Will Engage in Irregularities (% respondents)	0.511	0.001	0.007	-0.024	0.019	0.008	0.022	-0.007	-0.029
		(0.016)	(0.021)	(0.021)	(0.021)	(0.020)	(0.023)	(0.023)	(0.023)

Continued on next page

Table A12 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Demeaned - Candidate Will Engage in Irregularities (%)	-0.000	0.000	0.000	0.000**	0.000	0.000*	0.000*	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Above Average - Candidate Will Engage in Irregularities (%)	0.473	-0.006	0.002	-0.006	-0.014	0.015	-0.006	0.036*	0.042**
		(0.015)	(0.019)	(0.019)	(0.018)	(0.017)	(0.021)	(0.019)	(0.021)
Past Malfeasance (=1)	0.263	-0.002	0.002	-0.009	0.000	0.019	-0.021	0.061	0.082
		(0.045)	(0.064)	(0.059)	(0.042)	(0.058)	(0.069)	(0.060)	(0.064)
Candidate Will Win the Election (% respondents)	0.255	0.005	0.004	0.003	0.009	0.015	0.019	0.011	-0.008
		(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.013)	(0.012)	(0.013)
Log(Age)	3.791	0.009	0.002	0.016	0.010	0.003	0.013	-0.006	-0.018
		(0.009)	(0.011)	(0.011)	(0.011)	(0.010)	(0.012)	(0.012)	(0.012)
Female(=1)	0.172	-0.029**	-0.039**	-0.019	-0.029*	-0.007	-0.006	-0.008	-0.002
		(0.013)	(0.017)	(0.018)	(0.016)	(0.015)	(0.018)	(0.017)	(0.019)
Incumbent Party - Lax (=1)	0.184	0.007	0.002	0.003	0.014	0.024*	0.038**	0.011	-0.028*
		(0.012)	(0.016)	(0.014)	(0.015)	(0.014)	(0.016)	(0.016)	(0.016)
Incumbent Party - Strict (=1)	0.044	0.009	-0.007	0.010	0.025**	0.009	0.001	0.017	0.016
		(0.007)	(0.009)	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)	(0.011)
Party Coalition (=1)	0.245	0.003	-0.011	-0.000	0.021	0.013	0.015	0.012	-0.003
		(0.019)	(0.026)	(0.023)	(0.025)	(0.022)	(0.025)	(0.027)	(0.027)
Independent Candidate (=1)	0.027	-0.009	-0.008	0.002	-0.021***	0.008	0.014*	0.002	-0.011
		(0.006)	(0.007)	(0.008)	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)

Notes: This table presents the balance checks for a selected set of covariates in the candidate-level data. The control group mean of each variable is presented in column (1). In each of the remaining columns the difference in means is reported for the shown treatment groups. Clustered standard errors at the municipal-level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A13: Correlation Between Past Malfeasance and Survey Measures of Likelihood to Engage in Irregularities

	(1)	(2)	(3)
	Past Malfeasance (z-score)		
Candidate will engage in irregularities (z-score)	0.272***		
	(0.057)		
Demeaned Candidate will engage in irregularities (z-score)		0.412***	
		(0.056)	
Above Average Candidate will engage in irregularities (z-score)			0.383***
			(0.060)
Sample Size	256	256	256
N. Municipalities	48	48	48

Notes: This table presents the OLS results of regressing a z-score of an indicator of whether a candidate was found to be involved in malfeasance in the past according to the investigation by the NGO PARES on different survey-based variables about the likelihood that a candidate commits irregularities. All of these variables are normalized so the estimates reported can be interpreted as correlations. In column (1) the dependent variable is the proportion of respondents from the pre-treatment survey that say that the candidate might commit at least one type of electoral irregularity. In column (2) it is this same variable, demeaned using the municipal-level mean. In column (3) it is an indicator that takes the value of one if this variable is above the municipal-level mean. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A14: Correlation Between Measures of Candidate Engagement in Irregularities and Other Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Candidate Will Win the Election (fraction of respondents)			Incumbent Party Candidate Strict Measure (=1)			Incumbent Party Candidate Lax Measure (=1)		
Candidate will engage in irregularities (z-score)	0.086^{***}			0.102^{***}			0.172^{***}		
	(0.019)			(0.018)			(0.018)		
Demeaned Candidate will engage in irregularities (z-score)		0.111^{***}			0.133^{***}			0.206^{***}	
		(0.019)			(0.022)			(0.018)	
Above Average Candidate will engage in irregularities (z-score)			0.116^{***}			0.122^{***}			0.181^{***}
			(0.018)			(0.019)			(0.018)
Sample Size	2989	2989	2989	2989	2989	2989	2989	2989	2989
N. Municipalities	630	630	630	630	630	630	630	630	630

Notes: This table presents the OLS results of regressing the z-scores of the measures indicating each candidate's likelihood of engagement in electoral irregularities from Table 5 on z-scores of different candidate covariates. In columns (1)-(3) the examined covariate is the proportion of respondents from the pre-treatment survey that say the candidate is going to win the election in their municipality. In columns (4)-(6) it is an indicator that takes the value of one if the candidate belongs to exactly the same party or coalition of parties as the incumbent mayor. In columns (7)-(9) it is an indicator that takes the value of one if the candidate belongs to a party or coalition of parties that shares at least one party with the incumbent mayor. All of these variables are normalized so the estimates reported can be interpreted as correlations. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A15: Impacts on Vote Share of More Popular Candidates

	(1)	(2)	(3)
		Vote Share (%)	
Interaction term Z:	Candidate Will Win the Election (fraction of respondents)	Incumbent Party Candidate Strict Measure (=1)	Incumbent Party Candidate Lax Measure (=1)
Panel A. Pooled Treatment			
$[T \times Z]$ Any treatment $\times Z$	-1.033	-2.001	1.249
	(2.113)	(1.878)	(1.187)
	[0.528]	[0.316]	[0.324]
Panel B. Subtreatments by Types of Ad			
$[IA \times Z]$ Information Ad $\times Z$	-0.469	-6.630***	1.449
	(2.541)	(2.361)	(1.466)
	[0.858]	[0.014]	[0.360]
$[CA \times Z]$ Call-to-Action Ad $\times Z$	-0.839	-1.720	1.167
	(2.765)	(2.409)	(1.574)
	[0.780]	[0.504]	[0.482]
$[I + CA \times Z]$ Info + Call-to-Action Ad $\times Z$	-1.795	0.479	1.136
	(2.737)	(2.449)	(1.598)
	[0.528]	[0.850]	[0.500]
Test $IA \times Z = CA \times Z$, p-value	0.90	0.06	0.87
Test $IA \times Z = I + CA \times Z$, p-value	0.63	0.01	0.85
Test $CA \times Z = I + CA \times Z$, p-value	0.75	0.42	0.99
Panel C. Subtreatments by Letter - No Letter			
$[NL \times Z]$ No Letter - Any Ad $\times Z$	-0.150	-2.323	1.032
	(2.880)	(2.636)	(1.587)
	[0.954]	[0.378]	[0.542]
$[L \times Z]$ Letter - Any Ad $\times Z$	-1.446	-1.832	1.377
	(2.207)	(2.016)	(1.272)
	[0.514]	[0.348]	[0.304]
Test $NL \times Z = L \times Z$, p-value	0.62	0.85	0.82
Control Mean	19.05	19.05	19.05
Sample Size	2989	2989	2989
N. Municipalities	630	630	630

Notes: The outcome in all columns is the vote share of each candidate, expressed as a percentage of total valid votes. In each of these columns, different proxies for candidate popularity are used to compute candidate-level heterogeneous effects. In column (1) it is the proportion of respondents from the pre-treatment survey that say the candidate is going to win the election in their municipality. In column (2) it is an indicator that takes the value of one if the candidate belongs to exactly the same party or coalition of parties as the incumbent mayor. In column (3) it is an indicator that takes the value of one if the candidate belongs to a party or coalition of parties that shares at least one party with the incumbent mayor. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Clustered standard errors at the municipal-level are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A16: Impacts on Additional Electoral Outcomes

	(1) Turnout (%)	(2) Margin of Victory (%)
Panel A. Pooled Treatment		
$[T]$ Any treatment	0.333	0.826
	(0.528)	(0.835)
	[0.410]	[0.306]
Panel B. Subtreatments by Types of Ad		
$[IA]$ Information Ad	0.763	1.542
	(0.926)	(1.215)
	[0.404]	[0.202]
$[CA]$ Call-to-Action Ad	0.502	1.093
	(0.828)	(1.052)
	[0.528]	[0.310]
$[I + CA]$ Info + Call-to-Action Ad	0.471	0.165
	(0.940)	(1.082)
	[0.592]	[0.894]
Test $IA = CA$, p-value	0.78	0.73
Test $IA = I + CA$, p-value	0.78	0.30
Test $CA = I + CA$, p-value	0.97	0.43
Panel C. Subtreatments by Letter - No Letter		
$[NL]$ No Letter - Any Ad	0.476	0.532
	(0.890)	(1.110)
	[0.650]	[0.618]
$[L]$ Letter - Any Ad	0.632	1.138
	(0.758)	(0.923)
	[0.416]	[0.194]
Test $NL = L$, p-value	0.85	0.58

Notes: The outcome in column (1) is turnout, as a percentage of the people registered to vote. In column (2) it is the margin of the winning candidate over the runner-up, expressed as a percentage of total votes. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A17: Estimates of the Percentage of the Effect on Candidate Vote Share Due to Decreasing Electoral Irregularities

Electoral Irregularity Measure / Measure of Z_c:	Candidate will engage in irregularities (fraction of respondents)	Demeaned Candidate will engage in irregularities (fraction of respondents)	Above Average Candidate will engage in irregularities (=1)
Media Irregularities (=1)	11.42	24.07	16.20
Number of Media Irregularities	13.70	21.56	16.45
Index of all Forensic Test Stats (z-score)	37.54	38.26	32.06
Any P-value From Forensic Tests < 0.05 (=1)	29.87	51.90	28.98

Notes: This table presents estimates of the percentage of the effect of the reporting campaign interventions on the vote share of the candidates more likely to engage in electoral irregularities accounted for by the decrease in electoral irregularities using the method described in Section 3.4, using different combinations of variables proxying for electoral irregularities and for the proxies for the likelihood that a candidate engages in electoral irregularities, Z_c.

Table A18: Robustness: Impacts on Reports - No Controls

	(1) Reports (=1)	(2) N. Reports	(3) High Quality Reports (=1)	(4) High Quality N. Reports
Panel A. Pooled Treatment				
$[T]$ Any treatment	0.116***	0.372***	0.089***	0.181***
	(0.036)	(0.103)	(0.031)	(0.053)
	[0.002]	[0.002]	[0.008]	[0.002]
Panel B. Subtreatments by Types of Ad				
$[IA]$ Information Ad	0.167***	0.484***	0.083**	0.127**
	(0.048)	(0.148)	(0.041)	(0.063)
	[0.000]	[0.000]	[0.028]	[0.040]
$[CA]$ Call-to-Action Ad	0.019	0.174	0.071*	0.155*
	(0.046)	(0.141)	(0.040)	(0.082)
	[0.720]	[0.180]	[0.058]	[0.042]
$[I+CA]$ Info + Call-to-Action Ad	0.160***	0.454***	0.113***	0.259***
	(0.048)	(0.148)	(0.042)	(0.085)
	[0.000]	[0.002]	[0.016]	[0.002]
Test $IA = CA$, p-value	0.01	0.08	0.79	0.75
Test $IA = I + CA$, p-value	0.89	0.87	0.53	0.16
Test $CA = I + CA$, p-value	0.01	0.11	0.37	0.33
Panel C. Subtreatments by Letter - No Letter				
$[NL]$ No Letter - Any Ad	0.230***	0.491***	0.148***	0.243***
	(0.048)	(0.138)	(0.043)	(0.084)
	[0.000]	[0.000]	[0.002]	[0.000]
$[L]$ Letter - Any Ad	0.057	0.311***	0.059*	0.149**
	(0.039)	(0.118)	(0.033)	(0.058)
	[0.168]	[0.014]	[0.088]	[0.018]
Test $NL = L$, p-value	0.00	0.23	0.04	0.29
Control Mean	0.29	0.55	0.16	0.20
Sample Size	698	698	698	698

Notes: This table reports the same estimates as in Table 3 without any control variables except for strata fixed effects. The outcome in column (1) is an indicator of whether any report was issued to the MOE from each municipality. In column (2) it is the number of such reports. In columns (3)-(4) the same definitions are used on the subset of reports of a high quality (see Section 1 for a discussion about how the quality of reports is assessed by the MOE). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A19: Impacts on Irregularity Measures - No Controls

	(1)	(2)	(3)	(4)
	Media-Based Irregularities		Deviations from Benford's 2nd Digit Law	
	Media Irregularities (=1)	Number of Media Irregularities	Index of all Forensic Test Stats (z-score)	Any P-value From Forensic Tests < 0.05 (=1)
Panel A. Pooled Treatment				
[T] Any treatment	-0.048*	-0.071*	-0.077	-0.067*
	(0.028)	(0.038)	(0.078)	(0.041)
	[0.096]	[0.070]	[0.348]	[0.092]
Panel B. Subtreatments by Types of Ad				
[IA] Information Ad	-0.045	-0.073	-0.169*	-0.089*
	(0.035)	(0.044)	(0.097)	(0.051)
	[0.232]	[0.142]	[0.106]	[0.092]
[CA] Call-to-Action Ad	-0.036	-0.056	-0.081	-0.067
	(0.036)	(0.047)	(0.097)	(0.052)
	[0.360]	[0.236]	[0.414]	[0.226]
[$I + CA$] Info + Call-to-Action Ad	-0.063*	-0.085*	0.020	-0.045
	(0.034)	(0.045)	(0.100)	(0.052)
	[0.072]	[0.078]	[0.844]	[0.428]
Test $IA = CA$, p-value	0.79	0.70	0.39	0.70
Test $IA = I + CA$, p-value	0.60	0.78	0.07	0.44
Test $CA = I + CA$, p-value	0.43	0.52	0.33	0.70
Panel C. Subtreatments by Letter - No Letter				
[NL] No Letter - Any Ad	-0.034	-0.049	-0.010	-0.050
	(0.036)	(0.047)	(0.101)	(0.052)
	[0.346]	[0.334]	[0.924]	[0.342]
[L] Letter - Any Ad	-0.056*	-0.083**	-0.111	-0.075*
	(0.030)	(0.039)	(0.082)	(0.044)
	[0.048]	[0.036]	[0.186]	[0.100]
Test $NL = L$, p-value	0.48	0.38	0.28	0.60
Control Mean	0.16	0.20	0.00	0.51
Sample Size	698	698	698	698

Notes: This table reports the same estimates as in Table 4 without any control variables except for strata fixed effects. The outcome in column (1) is an indicator of whether any irregularity was reported in the media in a particular municipality. In column (2) it is the number of different irregularities. In column (3) it is the index of the χ^2, Kolmogorov-Smirnov and Kuiper test statistics testing for Benford's 2nd digit law, described in Section 3.2.2 . In column (4) it is an indicator that takes the value of one if the p-value of any of these tests leads to rejection of the null hypothesis with less than a 5% significance level. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A20: Robustness: Impacts on Vote Share of Candidates Likely to Engage in Irregularities - No Controls

	(1)	(2)	(3)	(4)
	Vote Share (%)			
Interaction term Z:	Candidate will engage in irregularities (fraction of respondents)	Demeaned Candidate will engage in irregularities (fraction of respondents)	Above Average Candidate will engage in irregularities (=1)	Past Malfeasance (=1)
Panel A. Pooled Treatment				
$[T \times Z]$ Any treatment $\times Z$	-2.887	-3.245	-2.779**	-3.541
	(2.244)	(2.518)	(1.380)	(4.011)
	[0.146]	[0.190]	[0.012]	[0.404]
	{0.204}	{0.184}	{ 0.044}	
Panel B. Subtreatments by Types of Ad				
$[IA \times Z]$ Information Ad $\times Z$	-3.229	-4.539	-1.772	-11.860***
	(2.869)	(3.277)	(1.785)	(3.461)
	[0.258]	[0.156]	[0.312]	[0.030]
	{0.271}	{0.153}	{ 0.319}	
$[CA \times Z]$ Call-to-Action Ad $\times Z$	-0.913	-0.102	-1.399	4.866
	(3.105)	(3.356)	(1.865)	(6.167)
	[0.756]	[0.988]	[0.446]	[0.422]
	{0.769}	{0.981}	{ 0.444}	
$[I + CA \times Z]$ Info + Call-to-Action Ad $\times Z$	-4.334	-4.856	-5.200***	-1.085
	(2.970)	(3.293)	(1.778)	(5.731)
	[0.164]	[0.128]	[0.008]	[0.846]
	{0.158}	{0.142}	{ 0.002}	
Test $IA \times Z = CA \times Z$, p-value	0.50	0.24	0.86	0.01
Test $IA \times Z = I + CA \times Z$, p-value	0.74	0.93	0.08	0.06
Test $CA \times Z = I + CA \times Z$, p-value	0.33	0.21	0.06	0.43
Panel C. Subtreatments by Letter - No Letter				
$[NL \times Z]$ No Letter - Any Ad $\times Z$	-3.106	-2.107	-2.124	-2.139
	(2.874)	(3.039)	(1.845)	(5.049)
	[0.280]	[0.502]	[0.260]	[0.668]
	{0.275}	{0.460}	{ 0.230}	
$[L \times Z]$ Letter - Any Ad $\times Z$	-2.795	-3.892	-3.158**	-4.739
	(2.477)	(2.850)	(1.492)	(4.986)
	[0.212]	[0.172]	[0.038]	[0.352]
	{0.273}	{0.182}	{ 0.033}	
Test $NL \times Z = L \times Z$, p-value	0.91	0.56	0.56	0.67
Control Mean	18.91	18.91	18.91	18.91
Sample Size	2989	2989	2989	263
N. Municipalities	630	630	630	48

Notes: This table reports the same estimates as in Table 5 without any control variables except for strata fixed effects. The outcome in all columns is the vote share of each candidate, expressed as a percentage of total valid votes. In each of these columns, a different measure of the likelihood that a candidate commits irregularities is used to compute the candidate-level heterogeneous effects. In column (1) it is the proportion of respondents from the pre-treatment survey that say that the candidate might commit at least one type of electoral irregularity. In column (2) the outcome is this same variable, demeaned using the municipal-level mean. In column (3) it is an indicator that takes the value of one if this variable is above the municipal-level mean. Finally, in column (4) it is an indicator of whether a candidate was found to be involved in malfeasance in the past according to the investigation by the NGO PARES. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Clustered standard errors at the municipal-level are shown in parentheses, random inference p-values are shown in square brackets, and clustered wild-bootstrap p-values correcting for the variance in estimating Z are shown in curly brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A21: Long regression - Impacts on Reports

	(1) Reports (=1)	(2) N. Reports	(3) High Quality Reports (=1)	(4) High Quality N. Reports
Info + No Letter	0.217***	0.415**	0.122*	0.152*
	(0.075)	(0.174)	(0.067)	(0.088)
CtA + No Letter	0.132*	0.492*	0.115*	0.324*
	(0.071)	(0.270)	(0.064)	(0.188)
Info + CtA + No Letter	0.303***	0.608***	0.177***	0.263**
	(0.070)	(0.181)	(0.068)	(0.109)
Info + Partial K. Letter	0.078	0.504**	0.038	0.063
	(0.067)	(0.239)	(0.055)	(0.089)
CtA + Partial K. Letter	-0.081	-0.164	0.030	0.019
	(0.065)	(0.157)	(0.060)	(0.084)
Info + CtA + Partial K. Letter	0.103	0.146	0.028	0.057
	(0.071)	(0.174)	(0.055)	(0.092)
Info + Full K. Letter	0.136*	0.441	0.047	0.130
	(0.074)	(0.298)	(0.060)	(0.119)
CtA + Full K. Letter	-0.004	0.197	0.059	0.082
	(0.065)	(0.184)	(0.061)	(0.094)
Info + CtA + Full K. Letter	0.058	0.615**	0.119**	0.443**
	(0.067)	(0.310)	(0.060)	(0.189)
Test: Any treatment is equal, p-value	0.00	0.02	0.57	0.32
Test: Info treatments are equal, p-value	0.32	0.95	0.53	0.73
Test: CtA treatments are equal, p-value	0.05	0.05	0.56	0.31
Test: Info + CtA treatments are equal, p-value	0.01	0.10	0.16	0.09
Test: No Letter treatments are equal, p-value	0.17	0.71	0.75	0.55
Test: Letter treatments are equal, p-value	0.15	0.05	0.86	0.44
Test: Partial K. Letter treatments are equal, p-value	0.06	0.04	0.99	0.91
Test: Full K. Letter treatments are equal, p-value	0.30	0.44	0.61	0.21
Control Mean	0.29	0.55	0.16	0.20
Sample Size	698	698	698	698

Notes: This table reports the effects of the intervention using the long regression and the same outcomes as in Table 3. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). The p-values of F-statistic tests of equality between different groups of coefficients are shown below each regression. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A22: Long regression - Impacts on Irregularities

	(1)	(2)	(3)	(4)
	Media-Based Irregularities		Deviations from Benford's 2nd Digit Law	
	Media Irregularities (=1)	Number of Media Irregularities	Index of all Forensic Test Stats (z-score)	Any P-value From Forensic Tests < 0.05 (=1)
Info + No Letter	-0.049	-0.079	-0.234**	-0.153**
	(0.051)	(0.063)	(0.114)	(0.067)
CtA + No Letter	-0.046	-0.067	-0.097	-0.035
	(0.048)	(0.062)	(0.129)	(0.078)
Info + CtA + No Letter	-0.029	-0.028	0.090	-0.004
	(0.052)	(0.072)	(0.144)	(0.078)
Info + Partial K. Letter	-0.085*	-0.128**	-0.274**	-0.053
	(0.049)	(0.055)	(0.125)	(0.073)
CtA + Partial K. Letter	-0.074*	-0.101*	-0.172*	-0.117
	(0.044)	(0.060)	(0.102)	(0.079)
Info + CtA + Partial K. Letter	-0.098**	-0.138***	-0.118	-0.106
	(0.041)	(0.047)	(0.109)	(0.071)
Info + Full K. Letter	-0.045	-0.068	-0.391***	-0.153**
	(0.049)	(0.062)	(0.125)	(0.074)
CtA + Full K. Letter	0.003	-0.023	-0.016	-0.078
	(0.058)	(0.070)	(0.134)	(0.073)
Info + CtA + Full K. Letter	-0.075	-0.102*	-0.091	-0.051
	(0.047)	(0.061)	(0.144)	(0.067)
Test: Any treatment is equal, p-value	0.85	0.70	0.16	0.75
Test: Info treatments are equal, p-value	0.77	0.63	0.55	0.46
Test: CtA treatments are equal, p-value	0.50	0.61	0.57	0.72
Test: Info + CtA treatments are equal, p-value	0.49	0.30	0.42	0.56
Test: No Letter treatments are equal, p-value	0.95	0.82	0.14	0.21
Test: Letter treatments are equal, p-value	0.68	0.60	0.21	0.87
Test: Partial K. Letter treatments are equal, p-value	0.89	0.83	0.55	0.78
Test: Full K. Letter treatments are equal, p-value	0.50	0.60	0.05	0.51
Control Mean	0.16	0.20	0.00	0.51
Sample Size	698	698	698	698

Notes: This table reports the effects of the intervention using the long regression and the same outcomes as in Table 4. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). The p-values of F-statistic tests of equality between different groups of coefficients are shown below each regression. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A23: Long regression - Impacts on Vote Share of Candidates Likely to Engage in Irregularities

	(1)	(2)	(3)
		Vote Share (%)	
Interaction term Z:	Candidate will engage in irregularities (fraction of respondents)	Demeaned Candidate will engage in irregularities (fraction of respondents)	Above Average Candidate will engage in irregularities (=1)
Info + No Letter $\times Z$	-4.816**	-4.319*	-1.225
	(2.444)	(2.568)	(1.394)
CtA + No Letter $\times Z$	-3.053	-0.024	-2.868*
	(2.994)	(3.730)	(1.660)
Info + CtA + No Letter $\times Z$	-4.258	-4.294	-3.227*
	(2.953)	(3.935)	(1.689)
Info + Partial K. Letter $\times Z$	-5.905**	-6.080	-3.715**
	(2.708)	(3.863)	(1.598)
CtA + Partial K. Letter $\times Z$	-4.520	-6.303*	-3.880**
	(2.780)	(3.775)	(1.560)
Info + CtA + Partial K. Letter $\times Z$	-2.093	-4.507	-3.061*
	(3.019)	(3.621)	(1.720)
Info + Full K. Letter $\times Z$	-2.251	-4.386	-2.518
	(2.792)	(2.902)	(1.640)
CtA + Full K. Letter $\times Z$	2.250	4.568	1.246
	(2.487)	(3.055)	(1.487)
Info + CtA + Full K. Letter $\times Z$	-2.870	-2.611	-3.140*
	(3.052)	(3.870)	(1.711)
Test: Any treatment is equal, p-value	0.35	0.21	0.19
Test: Info treatments are equal, p-value	0.55	0.91	0.42
Test: CtA treatments are equal, p-value	0.09	0.05	0.02
Test: Info + CtA treatments are equal, p-value	0.85	0.92	1.00
Test: No Letter treatments are equal, p-value	0.88	0.56	0.53
Test: Letter treatments are equal, p-value	0.18	0.09	0.08
Test: Partial K. Letter treatments are equal, p-value	0.58	0.92	0.92
Test: Full K. Letter treatments are equal, p-value	0.24	0.05	0.06
Control Mean	18.91	18.91	18.91
Sample Size	2989	2989	2989
N. Municipalities	630	630	630

Notes: This table reports the effects of the intervention using the long regression, including indicators for all possible treatment interactions, on the vote share of each candidate, expressed as a percentage of total valid votes. In each column, the different measure of the likelihood that a candidate commits irregularities is used to compute the candidate-level heterogeneous effects. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). The p-values of F-statistic tests of equality between different groups of coefficients are shown below each regression. Clustered standard errors at the municipal-level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A24: Impacts on Reports - By Type of Letter

	(1) Reports (=1)	(2) N. Reports	(3) High Quality Reports (=1)	(4) High Quality N. Reports
[NL] No Letter - Any Ad	0.217***	0.506***	0.138***	0.247***
	(0.047)	(0.135)	(0.043)	(0.084)
	[0.000]	[0.000]	[0.002]	[0.000]
[PL] Partial Knowledge Letter - Any Ad	0.034	0.167	0.032	0.047
	(0.045)	(0.128)	(0.038)	(0.058)
	[0.472]	[0.176]	[0.392]	[0.414]
[FL] Full Knowledge Letter - Any Ad	0.064	0.423**	0.076*	0.225**
	(0.046)	(0.165)	(0.040)	(0.087)
	[0.170]	[0.004]	[0.058]	[0.002]
Test NL = PL, p-value	0.00	0.03	0.02	0.02
Test NL = FL, p-value	0.00	0.67	0.19	0.84
Test PL = FL, p-value	0.56	0.17	0.32	0.05
Control Mean	0.29	0.55	0.16	0.20
Sample Size	698	698	698	698

Notes: The outcome in column (1) is an indicator of whether any report was issued to the MOE from each municipality. In column (2) it is the number of such reports. In columns (3)-(4) the same definitions are used on the subset of reports of a high quality (see Section 1 for a discussion about how the quality of reports is assessed by the MOE). All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A25: Impacts on Irregularity Measures - By Type of Letter

	(1) Media-Based Irregularities	(2)	(3) Deviations from Benford's 2nd Digit Law	(4)
	Media Irregularities (=1)	Number of Media Irregularities	Index of all Forensic Test Stats (z-score)	Any P-value From Forensic Tests < 0.05 (=1)
[NL] No Letter - Any Ad	-0.041	-0.058	-0.080	-0.065
	(0.035)	(0.046)	(0.088)	(0.050)
	[0.250]	[0.240]	[0.354]	[0.216]
[PL] Partial Knowledge Letter - Any Ad	-0.085***	-0.122***	-0.189**	-0.092*
	(0.032)	(0.041)	(0.080)	(0.050)
	[0.008]	[0.006]	[0.020]	[0.074]
[FL] Full Knowledge Letter - Any Ad	-0.040	-0.065	-0.164*	-0.093*
	(0.035)	(0.046)	(0.092)	(0.048)
	[0.282]	[0.168]	[0.054]	[0.066]
Test NL = PL, p-value	0.19	0.12	0.21	0.62
Test NL = FL, p-value	0.96	0.87	0.39	0.60
Test PL = FL, p-value	0.18	0.15	0.78	0.98
Control Mean	0.16	0.20	0.00	0.51
Sample Size	698	698	698	698

Notes: The outcome in column (1) is an indicator of whether any irregularity was reported in the media in a particular municipality. In column (2) it is the number of different irregularities. In column (3) it is the index of the χ^2, Kolmogorov-Smirnov and Kuiper test statistics testing for Benford's 2nd digit law, described in Section 3.2.2 . In column (4) it is an indicator that takes the value of one if the p-value of any of these tests leads to rejection of the null hypothesis with less than a 5% significance level. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and random inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A26: Impacts on Vote Share of Candidates Likely to Engage in Irregularities - By Type of Letter

	(1)	(2)	(3)	(4)
	Vote Share (%)			
Interaction term Z:	Candidate will engage in irregularities (fraction of respondents)	Demeaned Candidate will engage in irregularities (fraction of respondents)	Above Average Candidate will engage in irregularities (=1)	Past Malfeasance (=1)
$[NL \times Z]$ No Letter - Any Ad $\times Z$	-4.405**	-3.245	-2.427**	-4.427
	(1.892)	(2.230)	(1.067)	(3.585)
	[0.022]	[0.168]	[0.020]	[0.296]
	{0.020}	{0.148}	{0.023}	
$[FL \times Z]$ Full Knowledge Letter - Any Ad $\times Z$	-1.034	-0.418	-1.450	-3.129
	(1.907)	(2.219)	(1.090)	(4.868)
	[0.604]	[0.870]	[0.198]	[0.482]
	{0.588}	{0.852}	{0.181}	
$[PL \times Z]$ Partial Knowledge Letter - Any Ad $\times Z$	-4.245**	-5.587**	-3.584***	-8.830**
	(1.945)	(2.433)	(1.089)	(3.794)
	[0.040]	[0.018]	[0.000]	[0.050]
	{0.032}	{0.026}	{0.002}	
Test $NL \times Z = PL \times Z$, p-value	0.94	0.39	0.33	0.27
Test $NL \times Z = FL \times Z$, p-value	0.09	0.26	0.41	0.80
Test $PL \times Z = FL \times Z$, p-value	0.12	0.06	0.08	0.28
Control Mean	18.91	18.91	18.91	18.91
Sample Size	2989	2989	2989	263
N. Municipalities	630	630	630	48

Notes: The outcome in all columns is the vote share of each candidate, expressed as a percentage of total valid votes. In each of these columns, a different measure of the likelihood that a candidate commits irregularities is used to compute the candidate-level heterogeneous effects. In column (1) it is the proportion of respondents from the pre-treatment survey that say that the candidate might commit at least one type of electoral irregularity. In column (2) the outcome is this same variable, demeaned using the municipal-level mean. In column (3) it is an indicator that takes the value of one if this variable is above the municipal-level mean. Finally, in column (4) it is an indicator of whether a candidate was found to be involved in malfeasance in the past according to the investigation by the NGO PARES. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Clustered standard errors at the municipal-level are shown in parentheses, random inference p-values are shown in square brackets, and clustered wild-bootstrap p-values correcting for the variance in estimating Z are shown in curly brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A27: Cost-Benefit Comparisons

Type of Intervention	Paper	Cost of reducing by 1 p.p. the vote share of candidates likely to engage in irregularities per polling station	Votes reduced for candidates likely to engage in irregularities per dollar
Citizen monitoring	This paper	\$0.69	21
Electoral Observers	Enikolopov et al. (2013)	\$545-\$1818	–
Top-down ICT monitoring	Callen and Long (2015)	–	0.03
Top-down ICT monitoring	Callen et al. (2016)	\$13	–
Voter-education	Schechter and Vasudevan (2021)	–	109

Notes: This table displays cost-benefit estimates of different interventions on two metrics: (1) the USD cost of reducing the vote share of candidates likely to engage in electoral irregularities by one percentage point per polling station; (2) the number of votes for these same candidates that would be reduced by a single dollar investment in the intervention. See Appendix H for further details about these estimates.

Table A28: Balance on Post-Treat Survey Respondent Characteristics

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Panel A. Respondent Covariates									
Female(=1)	0.263	-0.008	-0.018	0.002	-0.009	0.018	0.020	0.016	-0.005
		(0.012)	(0.017)	(0.015)	(0.017)	(0.015)	(0.019)	(0.017)	(0.018)
Age	33.352	-0.254	-0.845	0.124	-0.065	0.598	1.201*	-0.013	-1.214*
		(0.494)	(0.616)	(0.694)	(0.591)	(0.623)	(0.706)	(0.699)	(0.658)
High School or Less (=1)	0.582	0.010	-0.007	0.007	0.029	0.018	0.017	0.019	0.002
		(0.020)	(0.027)	(0.026)	(0.026)	(0.026)	(0.029)	(0.029)	(0.028)
Panel B. Non-Response to Questions About Irregularities									
Vote Buying - No Response (=1)	0.468	0.003	0.012	-0.004	0.002	0.005	-0.006	0.016	0.022
		(0.015)	(0.019)	(0.019)	(0.020)	(0.017)	(0.021)	(0.019)	(0.022)
V. Intimidation - No Response (=1)	0.485	0.006	0.016	0.004	-0.001	0.013	-0.000	0.026	0.026
		(0.016)	(0.020)	(0.019)	(0.020)	(0.017)	(0.021)	(0.020)	(0.022)
Registr. Fraud - No Response (=1)	0.485	0.013	0.016	0.010	0.013	0.010	-0.001	0.021	0.022
		(0.015)	(0.019)	(0.019)	(0.020)	(0.017)	(0.021)	(0.020)	(0.022)
Public Campaign. - No Response (=1)	0.477	-0.003	0.010	-0.005	-0.012	0.006	-0.007	0.019	0.026
		(0.015)	(0.019)	(0.019)	(0.020)	(0.017)	(0.021)	(0.019)	(0.021)
Elect. Fraud - No Response (=1)	0.478	0.009	0.011	0.012	0.004	0.021	0.004	0.037*	0.033
		(0.015)	(0.019)	(0.019)	(0.020)	(0.017)	(0.021)	(0.020)	(0.022)
Illicit Advert. - No Response (=1)	0.461	0.008	0.017	0.009	-0.003	0.019	0.007	0.030	0.022
		(0.015)	(0.019)	(0.019)	(0.020)	(0.018)	(0.021)	(0.020)	(0.021)

Notes: This table presents the balance checks for a set of post-survey respondent characteristics. The control group mean of each variable is presented in column (1). In each of the remaining columns the difference in means is reported for the shown treatment groups. Clustered standard errors at the municipal-level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A29: Covariate Balance - Post-Treatment Survey

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Panel A. Previous Reports Covariates									
Reports to MOE 2018	0.502	0.031	0.092	-0.041	0.042	-0.088	-0.173	-0.004	0.169
		(0.087)	(0.117)	(0.109)	(0.113)	(0.107)	(0.119)	(0.130)	(0.127)
Reports to MOE 2015	3.443	0.109	0.452	-0.255	0.129	0.053	0.164	-0.055	-0.219
		(0.390)	(0.509)	(0.468)	(0.519)	(0.449)	(0.537)	(0.523)	(0.562)
Panel B. Socioeconomic Covariates									
Log(Population)	10.001	-0.030	0.022	-0.060	-0.053	-0.040	-0.032	-0.049	-0.017
		(0.060)	(0.077)	(0.075)	(0.075)	(0.071)	(0.082)	(0.081)	(0.080)
Facebook Penetration 2019	0.420	0.017	0.031	0.022	-0.001	0.003	-0.002	0.008	0.011
		(0.025)	(0.034)	(0.032)	(0.031)	(0.030)	(0.034)	(0.036)	(0.037)
GDP p.c. 2016 (Ms of Pesos)	14.825	-0.815	-1.211	-2.299	1.038	2.123	1.381	2.854	1.473
		(2.100)	(2.301)	(2.112)	(2.783)	(1.530)	(1.691)	(2.149)	(2.366)
Poor 2005 (%)	43.841	0.349	0.529	0.806	-0.280	-3.040	-2.813	-3.263	-0.450
		(1.750)	(2.339)	(2.264)	(2.172)	(2.156)	(2.497)	(2.458)	(2.432)
Rural Population 2017	53.809	-2.927	-5.377**	-3.370	-0.073	1.222	1.680	0.770	-0.910
		(1.986)	(2.637)	(2.439)	(2.525)	(2.426)	(2.792)	(2.751)	(2.672)
Homicide Rate 2017	28.381	-1.462	-2.704	0.146	-1.823	0.660	-1.164	2.458	3.622
		(2.677)	(3.243)	(3.709)	(3.285)	(3.004)	(3.579)	(3.577)	(3.883)
Panel C. Political Covariates									
Number of Candidates 2019	5.966	-0.228	-0.119	-0.196	-0.366*	-0.202	-0.169	-0.234	-0.066
		(0.163)	(0.207)	(0.206)	(0.206)	(0.198)	(0.227)	(0.221)	(0.209)
Index - All Forensic Stats 2015	-0.027	0.102	0.312***	0.002	-0.005	-0.142	-0.141	-0.142	-0.001
		(0.082)	(0.119)	(0.101)	(0.100)	(0.107)	(0.125)	(0.119)	(0.117)
P-val From Forensics < 0.05 2015	0.386	0.052	0.103*	0.019	0.034	-0.057	-0.058	-0.057	0.001

Continued on next page

Table A29 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
		(0.042)	(0.054)	(0.054)	(0.053)	(0.051)	(0.059)	(0.059)	(0.059)
Turnout 2018 (%)	0.491	0.008	0.016	0.005	0.004	-0.007	-0.007	-0.008	-0.002
		(0.008)	(0.010)	(0.010)	(0.010)	(0.009)	(0.011)	(0.011)	(0.011)
Share Blank Votes 2018 (%)	4.563	0.083	0.390	-0.224	0.083	-0.088	-0.289	0.110	0.399
		(0.359)	(0.508)	(0.468)	(0.467)	(0.458)	(0.525)	(0.560)	(0.581)
Mayor Margin of Victory 2015	13.645	-0.262	0.262	-0.659	-0.387	1.412	0.471	2.340	1.869
		(1.015)	(1.413)	(1.192)	(1.250)	(1.159)	(1.273)	(1.427)	(1.397)
Santos Vote Share 2014	48.119	0.428	1.782	0.709	-1.183	1.466	0.875	2.050	1.174
		(1.870)	(2.316)	(2.322)	(2.370)	(2.056)	(2.467)	(2.340)	(2.482)
Zuluaga Vote Share 2014	49.208	-0.607	-1.973	-0.902	1.031	-1.619	-0.951	-2.277	-1.325
		(1.821)	(2.250)	(2.256)	(2.310)	(1.997)	(2.393)	(2.271)	(2.405)
Liberals Vote Share 2018	15.094	0.017	-0.990	0.710	0.328	0.507	-0.010	1.016	1.026
		(0.927)	(1.174)	(1.242)	(1.172)	(1.113)	(1.269)	(1.328)	(1.335)
Cambio R Vote Share 2018	15.053	-0.451	0.288	-1.101	-0.537	-1.244	-0.968	-1.517	-0.549
		(1.052)	(1.347)	(1.327)	(1.298)	(1.209)	(1.414)	(1.374)	(1.382)
Centro Dem Vote Share 2018	13.856	-1.224	-1.412	-1.780	-0.490	0.351	1.511	-0.792	-2.303*
		(0.918)	(1.140)	(1.142)	(1.175)	(1.002)	(1.250)	(1.108)	(1.255)
P de la U Vote Share 2018	14.818	0.012	0.138	0.075	-0.174	-0.202	-0.291	-0.114	0.177
		(0.948)	(1.185)	(1.259)	(1.189)	(1.084)	(1.232)	(1.325)	(1.356)
Green Party Vote Share 2018	4.526	0.145	0.237	0.475	-0.272	-0.043	-0.248	0.160	0.408
		(0.493)	(0.601)	(0.768)	(0.580)	(0.659)	(0.781)	(0.725)	(0.728)
Polo Vote Share 2018	2.805	-0.010	-0.108	0.326	-0.247	0.118	-0.046	0.280	0.326
		(0.261)	(0.284)	(0.421)	(0.270)	(0.335)	(0.344)	(0.395)	(0.314)
Decentes Vote Share 2018	1.240	0.137	0.169	0.031	0.209	0.090	-0.015	0.193	0.208
		(0.104)	(0.127)	(0.161)	(0.156)	(0.152)	(0.158)	(0.199)	(0.193)
Panel D. Geographic Covariates									

Continued on next page

Table A29 – continued from previous page

	Control Mean	Any Treatment vs Control	Information vs Control	Call-to-action vs Control	Info + Call-to-action vs Control	Any Letter vs No Letter	Letter P. Knowledge vs No Letter	Letter F. Knowledge vs No Letter	Letter F. vs Letter P. Knowledge
Caribean Region (=1)	0.197	0.014	0.055	0.013	-0.025	-0.010	-0.012	-0.008	0.004
		(0.034)	(0.046)	(0.044)	(0.042)	(0.042)	(0.048)	(0.048)	(0.048)
Center-East Region (=1)	0.236	-0.002	0.008	-0.020	0.005	-0.016	-0.025	-0.007	0.018
		(0.036)	(0.047)	(0.046)	(0.047)	(0.043)	(0.050)	(0.050)	(0.050)
Center-South Region (=1)	0.163	-0.047	-0.051	-0.072**	-0.018	-0.020	-0.016	-0.024	-0.009
		(0.030)	(0.037)	(0.035)	(0.039)	(0.033)	(0.039)	(0.038)	(0.037)
Coffee-growing Region (=1)	0.158	0.051	0.024	0.087**	0.042	0.038	0.043	0.033	-0.010
		(0.032)	(0.041)	(0.044)	(0.042)	(0.040)	(0.048)	(0.047)	(0.049)
Llanos Region (=1)	0.064	0.006	0.006	-0.022	0.033	0.002	0.010	-0.005	-0.015
		(0.021)	(0.027)	(0.024)	(0.030)	(0.026)	(0.031)	(0.029)	(0.030)
Pacific Region (=1)	0.182	-0.022	-0.042	0.014	-0.037	0.006	-0.000	0.011	0.012
		(0.032)	(0.040)	(0.043)	(0.040)	(0.037)	(0.043)	(0.043)	(0.044)
Panel E. Other Covariates									
Number Responses Survey	10.715	-0.168	-0.358	1.116	-1.233*	-0.171	-0.098	-0.244	-0.146
		(0.587)	(0.766)	(0.856)	(0.660)	(0.731)	(0.810)	(0.891)	(0.872)
Population Reached by Ads (%)	0.334	0.005	0.011	0.002	0.003	-0.003	0.003	-0.009	-0.012
		(0.015)	(0.020)	(0.020)	(0.020)	(0.019)	(0.023)	(0.022)	(0.024)
Users Reached by Ads (Thousands)	7.418	-0.888	-0.245	-1.342	-1.074	-0.573	-0.483	-0.662	-0.179
		(0.812)	(1.025)	(0.886)	(0.950)	(0.763)	(0.888)	(0.864)	(0.858)
No Users Reached by Ads (=1)	0.000	0.012**	0.007	0.014	0.014	-0.013	-0.013	-0.013	-0.000
		(0.005)	(0.007)	(0.010)	(0.010)	(0.013)	(0.014)	(0.014)	(0.010)
N Responses Post-Treat Survey	8.379	0.189	-0.016	0.250	0.331	-0.190	-0.353	-0.030	0.324
		(0.327)	(0.428)	(0.428)	(0.430)	(0.429)	(0.496)	(0.479)	(0.461)

Notes: This table presents the balance checks on the sample of municipalities with post-treatment survey responses for a selected set of covariates. The control group mean of each variable is presented in column (1). In each of the remaining columns the difference in means is reported for the shown treatment groups. Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A30: Impacts on Survey-Based Irregularity Measures

	(1) Irregularity Index (z-score)	(2) Vote buying (z-score)	(3) Voter intimidation (z-score)	(4) Registration fraud (z-score)	(5) Public servant campaigning (z-score)	(6) Electoral fraud (z-score)	(7) Illicit Advertising (z-score)
Panel A. Pooled Treatment							
[T] Any treatment	-0.158*	-0.096	0.019	-0.042	-0.215**	-0.169**	-0.174**
	(0.088)	(0.088)	(0.083)	(0.086)	(0.086)	(0.084)	(0.084)
	[0.080]	[0.282]	[0.796]	[0.648]	[0.004]	[0.040]	[0.052]
Panel B. Subtreatments by Types of Ad							
[IA] Information Ad	-0.120	-0.021	-0.066	-0.077	-0.204*	-0.085	-0.116
	(0.118)	(0.116)	(0.103)	(0.115)	(0.113)	(0.110)	(0.108)
	[0.340]	[0.876]	[0.526]	[0.466]	[0.100]	[0.434]	[0.270]
[CA] Call-to-Action Ad	-0.151	-0.100	0.017	0.102	-0.126	-0.251**	-0.204*
	(0.110)	(0.117)	(0.105)	(0.108)	(0.110)	(0.105)	(0.109)
	[0.142]	[0.394]	[0.882]	[0.360]	[0.226]	[0.022]	[0.052]
[$I + CA$] Info + Call-to-Action Ad	-0.203*	-0.166	0.107	-0.151	-0.315***	-0.170	-0.201*
	(0.117)	(0.113)	(0.109)	(0.112)	(0.114)	(0.106)	(0.112)
	[0.078]	[0.124]	[0.340]	[0.164]	[0.004]	[0.126]	[0.046]
Test $IA = CA$, p-value	0.81	0.55	0.44	0.15	0.52	0.15	0.46
Test $IA = I + CA$, p-value	0.53	0.26	0.12	0.56	0.38	0.47	0.49
Test $CA = I + CA$, p-value	0.68	0.61	0.43	0.04	0.12	0.47	0.98
Panel C. Subtreatments by Letter - No Letter							
[NL] No Letter - Any Ad	-0.079	-0.088	0.142	0.023	-0.171	-0.179*	-0.110
	(0.117)	(0.117)	(0.104)	(0.111)	(0.113)	(0.107)	(0.112)
	[0.504]	[0.424]	[0.182]	[0.844]	[0.138]	[0.074]	[0.302]
[L] Letter - Any Ad	-0.199**	-0.100	-0.044	-0.076	-0.238**	-0.163*	-0.208**
	(0.094)	(0.095)	(0.090)	(0.094)	(0.093)	(0.090)	(0.090)
	[0.036]	[0.276]	[0.666]	[0.420]	[0.008]	[0.062]	[0.020]
Test $NL = L$, p-value	0.29	0.91	0.05	0.36	0.54	0.88	0.36
Control Mean	0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00
Sample Size	634	642	639	639	640	639	644

Notes: The outcomes in columns (2)-(7) are the average responses to questions about the likelihood of the occurrence of different types of irregularities. In column (1) it is an index of these variables. All variables have been standardized with respect to the control group mean and standard deviation. All specifications include the covariates selected using the method described in Chernozhukov et al. (2015) and Belloni et al. (2014). Robust standard errors are shown in parentheses and randomization inference p-values are shown in square brackets; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

B Dataset of News about Electoral Irregularities

We constructed our dataset of the electoral irregularities covered in the news from three different and complementary sources. We begin by discussing each of these sources and then describe further details about the coding of news.

1.Private News Monitoring Company: We hired *Siglo Data*, a data analysis firm based in Colombia, which is specialized in monitoring and classifying news in mass media and social networks. They actively monitor and classify news coming from TV channels, radio stations, written press, both offline and online. More concretely, they claim to cover news from over 100 newspapers and 60 magazines, 10 national and 60 local radio stations, 15 national and 10 regional TV channels and 370 news websites. They provided the data for the universe of news that had mentions about electoral irregularities related to the 2019 local elections, and that appeared in either the written press, the radio or newspapers on the internet up to November 10, two weeks after the date of the elections.

2. Press releases from Colombia's Electoral Court: We additionally relied on the official press releases produced by the *Consejo Nacional Electoral* (CNE), the National Electoral

Court of Colombia. As part of its regular activities, this entity is in charge of monitoring and controlling all the electoral activity of political groups and their candidates, as well as overseeing electoral organization and ensuring the proper development of electoral processes. In this vein, the CNE published on its website a list of news covering the 2019 local elections, which we included in our analysis.[47]

3. Electoral irregularity monitoring from an NGO: Finally, we use information from the *Fundación Paz y Reconciliación* (PARES), a well-known NGO in Colombia focused on producing independent research on conflict, security, governance, and democracy, among others. To this end, they monitor elections, and in 2019 they produced a report of instances of electoral irregularities gathered from media, citizen reports and their own monitoring activity in the field.[48]

Using the exhaustive list of news coming from these three sources, we hired a research assistant to go through each of these and classify them in terms of three main variables: (i) whether the news reported an electoral irregularity or not, (ii) whether the news came from information obtained from citizen reports submitted to the MOE, (iii) the types of electoral irregularities reported in the news. In our final data set, we exclude news that is not about electoral irregularities according to (i), and that in (ii) so that we do not confound the effect of the campaign on reporting from the effect on actual irregularities.

Using the cleaned data set, which contains over 160 news reports satisfying our criteria, we create an indicator of whether any electoral irregularity was reported by the media in each of one the municipalities in our study group, as well as a variable of the number of distinct irregularities reported per municipality.

C Covariates Included in the Analysis

The covariates included in the analysis can be broadly categorized in three groups:

Past reports: We include the total number of reports made to the MOE in the 2015 local elections and the 2018 congressional elections, as a way to control both for previous experience with reporting channels and the prevalence of electoral irregularities. This data was provided by the MOE, aggregated at the municipal level.

Socioeconomic characteristics: As geographical and demographic variables we use the municipal population in 2018, the proportion of rural population in 2017, and dummies

[47] The website they use to publicize the coverage of elections can be accessed through the following link:https://www.cne.gov.co/prensa/cne-en-medios?start=78.

[48] The full report is publicly available online at this link: https://pares.com.co/2019/10/29/un-balance-nacional-de-estas-elecciones-locales-2019/.

for the six main regions in the country. As measures of economic activity and development we use an estimate of GDP per capita in 2016, the % of the poor population in 2005 and the homicide rate in 2017. All of these variables were taken from the National Department of Statistics (DANE) except for the last two which were taken from the "Muncipal Characteristics" database created and updated by the Centro de Estudios sobre Desarrollo Económico (CEDE), at the Universidad de los Andes in Bogotá, Colombia. Finally, we also included Facebook's penetration rate (defined as the number of active Facebook users divided by total population), which we construct using user data from Facebook's Marketing API.

Political preferences: In order to get a rich set of political characteristics for each municipality we used the turnout in the congressional elections of 2018, the margin of victory in the 2015 mayoral elections, and the share of blank votes, as well as the vote share for each major party in the 2018 congressional elections and the vote share for each candidate in the second round of the 2014 presidential elections. Lastly, we include the number of candidates running in the 2019 mayoral elections, as well as the main forensic measures we used in the analysis but measured for the 2015 mayoral elections. All of these variables were constructed from the official voting records held by the Registraduría Nacional.

When running regressions at the candidate-level, we also included the following covariates for each of the mayoral candidates running in the municipalities in our sample:

Political platform: We included information about the political platform used by the candidate to register her candidacy. Specifically, whether she is running with a coalition of parties or as an independent. We also created a variable of whether the candidate is running as an incumbent or not. In Colombia, there is not immediate reelection for candidates. Thus, we computed being an incumbent as either running with exactly the same party or coalition of parties as the incumbent mayor (strict measure) or with at least one party shared by the party or coalition of parties as the incumbent mayor (lax measure).

Sociodemographic characteristics: We used data on the candidates' gender and age, which was provided by the Registraduría Nacional.

Survey-based characteristics: Using the pre-treatment survey (described in Section 2.5.2), we additionally created proxies for the likelihood that each candidate would engage in electoral irregularities, which are described in Section 3.3, and a measure of the popularity of each candidate, which corresponds to the percentage of respondents in each municipality that says a particular candidate will win the local election.

D Measures of Deviation from Benford's 2nd Digit Law

We use the following tests to determine whether the observed distribution of second digits in the voting booth counts differs from Benford's distribution in each municipality in our sample:

1. Pearson X^2:

$$X^2 = n \times \sum_{i=0}^{9} \frac{(o_i - e_i)^2}{e_i}$$

where o_i is the observed proportion of digit i and e_i is the expected proportion according to Benford's distribution. This statistic is distributed χ^2 with 9 degrees of freedom.

2. Kolmogorov-Smirnov D:

$$D = \sup_{x \in \{0,1,2,\ldots,9\}} |F_n(x) - F(x)|$$

where F_n is the empirical distribution of digits and $F(x)$ is the target distribution.

3. Kuiper V:

$$V = \max_{x \in \{0,1,2,\ldots,9\}} [F_n(x) - F(x)] + \max_{x \in \{0,1,2,\ldots,9\}} [F(x) - F_n(x)]$$

While the Pearson χ^2 test is probably the most commonly used in the literature, it has been shown to be under-powered in small samples (Nigrini 2012), such as the municipalities in our sample, which typically have 100-200 observations to compute these tests. The latter two tests are more appropriate for these types of samples and, in particular, the Kuiper test takes into account the "circular" nature of the distribution of second-digits. To further correct for small sample inference, we compute p-values simulating draws from the distribution under the null hypothesis that the data come from a Benford distribution.

E Robustness Using the Forensic Test Suggested by Beber and Scacco (2012)

In this section we begin by explaining the forensic tests suggested by Beber and Scacco (2012) and we then report the results of intervention using these tests as outcomes. The tests considered rely on psychological biases that humans have in *manually* manipulating numbers. As such, these tests do not detect more sophisticated ways of manipulating vote counts, such as the use of random number generators.

In particular, Beber and Scacco (2012) report three types of failures humans have in reproducing randomly occurring numbers: 1. the last digits of humanly manipulated numbers do not follow a uniform distribution; 2. humans *underestimate* the likelihood of the repetition of consecutive digits – e.g. "22" or "66"; 3. humans *overestimate* the likelihood of adjacent pairs of numbers – e.g. "34" or "43".

To test for the first phenomenon, we use tests akin to the ones used for Benford's 2nd Digit Law using the uniform distribution of the last digits as a reference. For the last two, we use indicators that take the value of one if there is a smaller than expected proportion of repeated digits, or there is a larger than expected proportion of pairs of adjacent digits, according to the distributions derived by Beber and Scacco (2012), focusing on the last two digits of voting counts.

Table A11 reports the results of using these tests as outcomes in our main specifications. Across the different tests, we see that the main conclusions from the previous results still hold, although some results lose precision, possibly since this test captures only manual manipulations of the voting counts, as suggested earlier. The intervention, and in particular the information advertisements, reduced deviations of the last digit from a uniform distribution, they decreased the probability of observing less than expected repeated digits, and they reduced the probability of observing more than expected adjacent pairs of digits. All of these results are consistent with fewer irregularities occurring in treated municipalities, according to the tests suggested by Beber and Scacco (2012).

F Measure of Electoral Irregularities Using Post-Treatment Survey

In addition to the media-based and forensic electoral irregularity measures, we also conducted an online post-treatment survey and used it to construct a survey-based measure of electoral irregularities. In the following section we explain how the survey was conducted and the outcome variables we construct from it, and we then proceed to report the effects of the intervention on those outcomes.

F.1 Post-Treatment Survey Recruitment and Outcomes

The roll-out of the post-treatment survey started immediately after the main intervention ended (starting October 29), and it lasted for 18 days (until November 15). Respondents were recruited using two different strategies. First, we recontacted through email the respondents from the pre-treatment survey who expressed interest in participating in this follow-up survey. Second, we also conducted a Facebook ad campaign identical to the one done in the pre-treatment survey to get additional respondents. Once again, we encouraged participation through a raffle for tablets.

The main goal of this survey was to obtain a measure of respondents' perception of the occurrence of electoral irregularities in their municipalities. We thus asked respondents how likely different types of irregularities had occurred in the previous elections in their municipality on a scale from 1 to 4 (with larger values representing higher likelihood). We did this for the most common irregularities described in Section 1: vote-buying, illicit advertisement, campaigning by public servants, voter intimidation, fraud in voter registration, and electoral fraud.

We then computed the average response across each municipality, and we standardized these variables using the control groups' mean and standard deviation for ease of interpretation. We also created a standardized index across all irregularity types to avoid the issue of multiple hypothesis testing.

We obtained approximately 5600 responses from 660 municipalities in the sample, but out of these, only 2964 of respondents from 634 municipalities answered the questions about the likelihood of occurrence of irregularities. Given that the intervention might have affected respondents' propensity or willingness of answering these questions, we check for balance of the response rate across the questions regarding the likelihood of irregularities, along with demographic characteristics of respondents. Results of this exercise (reported in Table A28), suggest that, in fact, response rates and respondents' demographic characteristics are well balanced. Furthermore, we find that the subsample of municipalities with responses to this survey are also balanced across municipal characteristics, including the number of responses to the survey (see Table A29).

F.2 Effect of the Intervention on Survey-Based Irregularity Measures

We report the effects of the different interventions on the survey-based irregularity measures in Table A30. Consistent with our other proxies for electoral irregularities, we find that the reporting campaign reduced the occurrence of irregularities. In particular, receiving any of the treatments reduced the irregularity index by 0.16 standard deviations ($p < 0.1$).

This effect is strongest for irregularities like campaigning by public servants, electoral fraud and illicit political advertising, but we find negative estimates for other types of irregularities – except for voter intimidation, which has a small and insignificant effect. Furthermore, we do not find statistically significant differences across types of interventions.

G Bootstrap Procedure to Account for Variance in Estimating Candidate-Level Variables

Denote the data by the triplet (y, X, Z), where y and X are at the candidate-municipality level, while Z corresponds to data from our pre-treatment survey, and thus is at the respondent-candidate-municipal level. The candidate level variables $\hat{z}_{ic}$ are estimated from Z but are included as regressors in the estimation so $\hat{z}_{ic} \in X$.

Given this notation, we we build on the wild bootstrap procedure proposed by Cameron et al. (2008), and incorporate an extra resampling stage of Z, in order to incorporate the variance in estimating the candidate level variables $\hat{z}_{ic}$. The procedure can be summarized as follows:

1. From the original sample, estimate $t = \frac{\hat{\gamma}-\gamma_0}{s_{\hat{\gamma}}}$, where $s_{\hat{\gamma}}$ is the standard error of $\hat{\gamma}$ clustered at the municipal level.

2. Estimate the restricted model which imposes the null hypothesis (i.e. $\gamma = 0$). Call the restricted estimates $\hat{\beta}^R$ and the corresponding residuals $\left\{(\hat{u}_1^R, ..., \hat{u}_M^R)\right\}$.

3. Do B iterations of this step. On the $b - th$ iteration:

   (a) For each candidate-municipality combination, create a sample of respondents Z^*_{ci} by resampling with replacement N_{ic} times from the original sample of respondents—where N_{ic} is the original number of respondents for candidate c in municipality i.

   (b) Compute the $\hat{z}^*_{ic}$, the measure of how likely each candidate is to engage in irregularities from Z^*.

   (c) Create a pseudo-sample (y^*, X^*) using the following method. For each cluster $i = 1, ..., M$, generate $\hat{u}_i^{R*} = a_i \hat{u}_i^R$, where a_i is a random variable that takes the value $\frac{1-\sqrt{5}}{2}$ with probability $\frac{1+\sqrt{5}}{2\sqrt{5}}$, or $1 - \frac{1-\sqrt{5}}{2}$ with probability $1 - \frac{1+\sqrt{5}}{2\sqrt{5}}$.[49] Then

[49] As explained by Cameron et al. (2008) these weights are preferred when the distribution of the estimates is potentially asymmetric. We use this alternative since it produces the most conservative p-values (i.e. the ones most likely not to reject the null) when applied to our setting.

define X^* as X but using the sampled $\hat{z}^*_{ic}$ instead of the original $\hat{z}_{ic}$. Finally, define $y^*_i = X^*_i \hat{\beta}^R + \hat{u}^{R*}_i$.

(d) Compute $t_b = \frac{\hat{\gamma}^*_b - \gamma_0}{s_{\hat{\gamma}^*_b}}$, where $\hat{\gamma}^*_b$ and $s_{\hat{\gamma}^*_b}$ are estimated as in step (1) but using the $b-th$ pseudo-sample.

4. Compute the bootstrapped p-value as $p = \frac{\sum_{b=1}^{B} \mathbb{1}\{|t_b|>|t|\}}{B}$

H Cost-Benefit Calculations and Comparisons

Intervention considered in this paper: 1. The Facebook advertisements cost \$10,870 USD in total. This implies a cost of \$15.57 USD per municipality or \$1.71 USD per polling station (since there are 6,349 polling stations in the municipalities in our sample). Given that the effect of the reporting campaign was to reduce the vote share of candidates above average in the percentage of people who say they will engage in electoral irregularities by 2.48 p.p., the cost of reducing a single percentage point is \$0.69 USD (= 1.71/2.48) per polling station. 2. The average votes in each municipality are 13,352, which means that the reporting campaign reduced by 331.13 (= $2.48/100 \times 13,352$) the votes for candidates above average in the percentage of people who say they will engage in electoral irregularities. This then implies that a single dollar invested in the campaign was responsible for a reduction of 21.26 votes (331.13/15.57) per municipality.

Electoral observers: Using the cost estimates by polling station provided by Callen et al. (2016), and the 11 p.p drop in the vote share estimated by Enikolopov et al. (2013), a 1 percentage point change in the vote share costs between \$545.45 (= 6000/11) and \$1,818.18 (= 20000/11) USD per polling station.

Top-down ICT monitoring: 1. Callen and Long (2015) report a "[...]total budget of just over US\$100,000" for 471 treated polling stations, which implies a cost of approximately \$210 USD per polling station. Given the effect of the intervention is to reduce by 6 the votes for candidates "connected" to electoral authorities, this implies that a single dollar reduced 0.029 votes for these candidates (= 6/210). 2. Callen et al. (2016) report a cost of \$40 USD per polling station and a reduction of the vote share of the incumbent candidate by 3 percentage points,[50] so reducing a single percentage point cost approximately \$13.33 USD (= 40/3).

Voter-education interventions: Schechter and Vasudevan (2021) directly report that the intervention they study had a cost-effectiveness of one-dollar investment translating into 109 fewer votes for candidates that are part of vote-buying parties.

[50]This is the largest estimate reported by them.

References

Beber, Bernd, and Alexandra Scacco. 2012. "What the Numbers Say: A Digit-Based Test for Election Fraud." *Political Analysis* 20:211–234.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *The Review of Economic Studies* 81 (2): 608–50.

Callen, Michael, Clark C. Gibson, Danielle F. Jung, and James D. Long. 2016. "Improving Electoral Integrity with Information and Communications Technology." *Journal of Experimental Political Science* 3:4–17.

Callen, Michael, and James D. Long. 2015. "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review* 105 (1): 354–381.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements For Inference With Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414–427.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments." *American Economic Review: Papers and Proceedings* 105 (5): 486–490.

Enikolopov, Ruben, Vasily Korovkin, Maria Petrova, Konstantin Sonin, and Alexei Zakharov. 2013. "Field Experiment Estimate of Electoral Fraud in Russian Parliamentary Elections." *Proceedings of the National Academy of Sciences* 110 (2): 448–452.

Nigrini, Mark J. 2012. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection.* Vol. 586. John Wiley & Sons.

Schechter, Laura, and Srinivasan Vasudevan. 2021. "Persuading Voters to Punish Corrupt Vote Buyers: Experimental Evidence from a Large-Scale Radio Campaign in India." Working paper.