

Monitoring the Vote or Voting to Monitor? Evidence from Two Large Scale Field Experiments in Colombia*

Mateo Montenegro[†] and Natalia Garbiras-Díaz[‡]

January 13, 2020

Click [here](#) for latest version

Abstract

Can crowdsourcing technologies aimed at augmenting civil oversight of elections increase electoral integrity? We report the results of two large-scale field experiments designed to assess the effectiveness of online crowdsourcing technologies in increasing the engagement of civil society in electoral monitoring around elections in Colombia. We leveraged Facebook advertisements to encourage citizen reporting of electoral irregularities through official websites, and also varied whether candidates were informed about the campaign in a subset of municipalities. In addition to the expected *informational effects* – whereby citizen reports increased, and politicians reduced their engagement in electoral irregularities – the results highlight powerful *salience effects*, which operated by making electoral irregularities more top-of-mind to citizens. Specifically, the advertisements generated a large shift in the vote share of candidates perceived to be less corrupt and away from those perceived to be more corrupt. We argue that these salience effects are driven by a shift in voter preferences towards candidates they perceived as ‘cleaner’. We formally test this hypothesis in a follow-up experiment around the 2019 mayoral elections in which we vary the salience of electoral irregularities in the advertisements sent through Facebook. As expected, we find that the advertisements featuring messages emphasizing the salience of electoral misdeeds generate a larger shift in the votes for ‘cleaner’ candidates than the ones only providing information about the reporting website.

*We are grateful for the guidance provided by Daron Acemoglu, Esther Duflo and Ben Olken. This paper has benefited greatly from the conversations with David Atkin, Abhijit Banerjee, Aicha Ben Dhia, Augustin Bergeron, Leopoldo Fergusson, Ray Fisman, Allan Hsiao, Stuti Khemani, Horacio Larreguy, Juliana Londoño, Francine Loza, Pablo Querubín, Frank Schilbach, Cory Smith, Román A. Zarate and all of the participants at the MIT Development and Political Economy lunches. We would also like to thank Laura Pulecio, Juliana Barberena, Sofía Díaz and Diana Velazco at the *Procuraduría General de la Nación*, Juan Esteban Lewin at *La Silla Vacía* and Marlon Pabón, along with the other members of the MOE that helped us out, 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 and the George and Obie Schultz Fund. The two experiments were approved by MIT’s IRB (the *Committee on the Use of Humans as Experimental Subjects*) with reference #1805347582 and #1904805455. The RCT is registered in the AEA RCT Registry with unique identifying number “AEARCTR-0004678”.

[†]Department of Economics, MIT. Contact: mateo.m@mit.edu.

[‡]Department of Political Science, U.C. Berkeley. Contact: nataliagarbirasdiaz@berkeley.edu.

1 Introduction

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.¹ 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 harm the economic and political stability of countries. By increasing the political returns of targeted transfers, clientelism leads to the underprovision of public goods and it generates public policy 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). Finally, voter intimidation might also help to perpetuate violence in weak states (Acemoglu, Robinson, et al., 2013; Robinson and Torvik, 2014).

Bottom-up monitoring technologies – broadly defined as technologies that involve civil society in the oversight of public goods and service provision – constitute a promising tool to fight electoral irregularities. Spurred by the World Bank’s 2004 *World Development Report*, governments and NGOs alike have heeded the call to use these types of technologies in areas as diverse as education, health, public works and elections as a way of deepening social accountability of governments.² Moreover, a ‘second generation’ of these technologies has taken advantage of the increase of the availability and use of the internet in the developing world, and has used online tools to further crowdsource monitoring tasks to civil society (Fox, 2015; Peixoto and Fox, 2016).

In this paper we investigate whether these crowdsourcing technologies can increase electoral integrity when applied to citizen oversight of elections. We do so by studying two field experiments designed to assess the effectiveness of a large-scale Facebook ad campaign aimed at encouraging citizens to report electoral irregularities through online official websites in Colombia.

Our findings highlight that these types of bottom-up monitoring campaigns operate on two different margins. Not only do these campaigns have effects that operate by increasing the available information to citizens about online reporting channels – which we call *informational effects*– but they also generate powerful *salience effects* that operate by making citizens more aware about the issues being monitored. In the context of our intervention, the Facebook ad campaign made electoral irregularities more top-of-mind to citizens, which then reacted by voting for candidates that they perceived to be ‘cleaner’, or less involved in electoral irregularities. These effects thus acted as *complements* to the objectives of the campaign in this setting. However, in theory, these salience effects could have also worked against the objectives of the campaign if they had made citizens too pessimistic about elections to act against electoral irregularities.

¹See Schedler (2002), Collier and Vicente (2012) and Gans-Morse et al. (2014) for a discussion of the different types of electoral irregularities and how politicians combine them strategically.

²See Fox, 2015 for a review of the literature evaluating interventions in these areas.

In a first experiment deployed around the 2018 Presidential elections, we designed and implemented an experiment that allowed us to disentangle the demand and the supply-side responses to our intervention. In a first stage, we randomized two thirds of our sample of 652 municipalities into a treatment condition in which citizens received Facebook ads that contained a message encouraging people to report electoral irregularities through a website hosted by the Office Attorney-Inspector General of Colombia (AG). In the second randomization stage, municipalities were cross-randomized into a ‘political awareness’ treatment. Candidates and their campaign staff were told that municipalities in this group were part of a grassroots campaign to oversee elections. Conceptually, the first treatment arm was designed to test the effect of the advertisement campaign without the (or at least with little) awareness of politicians about it, while the second treatment allows us to study the full equilibrium, when candidates and parties become aware of the intervention and have time to react and change their electoral strategies.

The advertisement campaign reached 1.4 million Facebook users, which represent over a third of the voting population of the municipalities in this treatment group, and each viewer saw the ad on average 3.5 times on their screen. This generated over 12 thousand clicks on the link to the AG’s reporting website –with a 0.9% click through rate and an average cost of \$0.5 USD per click – as well as substantial user engagement with the ads in the form of likes and comments. Despite the substantial engagement with the ad campaign, this treatment generated only a modest increase of 1.5 percentage points in the likelihood that citizens from treated municipalities completed and filed a report. The reason for this gap from clicks to full reports seems to have been that 95% of the citizens viewing the ad did so through their cellphone, while the AG’s reporting website was not fully compatible for cellphone use.

The letters sent to candidates and their parties reduced the occurrence of electoral irregularities by 0.1 standard deviations. This effect was driven by illicit political advertising and vote buying – which were reduced by 100% and 75%, respectively, compared to the control group mean – but had no effect other, less conspicuous, types of irregularities. We rationalize this through a model in which candidates substitute away from more conspicuous types of electoral irregularities towards less conspicuous types in face of the increased civilian monitoring triggered by a decrease in the cost of monitoring. Moreover, the ad campaign itself had no effect on electoral irregularities – i.e. independently from the letters sent to candidates – which validates our strategy to disentangle the demand and supply effects of the intervention through our different treatments.

In contrast to these modest *informational effects*, the ad treatment had a large effect on voting outcomes. The vote share for traditional candidates – defined as those coming from parties which have held substantial power in national and local posts historically – dropped by approximately 2 percentage points in the municipalities receiving the ads. Turnout was unaffected, so the decrease in the vote share for traditional candidates came exclusively from an almost identical increase in the vote share for non-traditional candidates, who had centered their campaigns around fighting corruption and clientelism. The implied *persuasion rate* (DellaVigna and Kaplan, 2007; DellaVigna and Gentzkow, 2010) of these estimates is approximately 6%, which means that one in every

17 citizens who viewed the ad changed their vote in favor of non-traditional candidates.

As mentioned before, we attribute this shift in electoral results to an increase in the *saliency* of electoral irregularities, which led citizens to change their vote towards candidates they perceived to be less engaged in electoral irregularities. However, an alternative interpretation is that the ad campaign decreased the occurrence of actual electoral irregularities, which then ‘freed’ voters from one set of candidates to the other. Consistent with our favored interpretation, we find that municipalities receiving the ad did not in fact experience a decrease in electoral irregularities, which is the premise of the alternative explanation. Also consistent with our interpretation, the ad treatment generated a 150% increase in the probability of participating in protests against the national government, which was led by the traditional candidate who won the election. This suggests that the campaign generated a persistent shift in the demand for ‘clean’ candidates (or at least what the citizens perceived as such) that expressed itself not only on their voting decisions but also on other subsequent forms of non-electoral opposition to traditional politicians.

To further probe into the mechanisms we conducted an original post-treatment survey that asked citizens about their attitudes towards different institutions in Colombia. In line with our interpretation, respondents from municipalities receiving the ads report less trust in elections – which might be caused by an increase in the saliency of electoral irregularities – while respondents from municipalities included in the letter sent to politicians report *higher* levels of trust in elections, which might have been generated by the decrease in conspicuous electoral irregularities in these municipalities.³

To more formally test our interpretation about the saliency effects of the intervention, we designed and implemented a second experiment conducted around the Colombian 2019 mayoral elections in which we unbundled the ad campaign’s purely ‘informational’ content from its message to act against electoral irregularities by reporting them. More concretely, we randomly exposed a set of municipalities to different ad versions which contained either (a) a message informing citizens about the reporting website, (b) a saliency message, drawing attention about the urgency to act against electoral irregularities and inviting citizens to report electoral them, or (c) both. We hypothesize that the saliency message would be the main one responsible for increasing the citizens awareness about electoral irregularities, and that it would thus create a larger shift in the vote share for candidates that were perceived to be ‘cleaner’.

The advertisement campaign in this second experiment had a similar scale to the first one, reaching approximately a third of the people registered to vote in the municipalities in the sample, and it was viewed three times by each viewer on average.

In order to identify which candidates were perceived to be ‘cleaner’ among the large set of candidates running for mayoral elections, we conducted a large online survey in which we asked

³An alternative interpretation is that this decrease in trust about elections comes from the technical issues in the reporting website – an issue that has been stressed in another context by Marx et al. (2017) However, we disfavor this interpretation since there is not a significant change on the trust for the AG, which would have been responsible for the technical issues.

citizens about their views on each candidate in their municipality three weeks ahead of the intervention. Confirming our interpretation about the first experiment, we find that municipalities receiving the advertisement featuring the salience message experienced an increase in the vote share of candidates that were ex-ante identified as ‘cleaner’ by respondents of an online survey of approximately 5%, while the ones that only received the information message about the reporting website did not experience any such change in the voting behavior of citizens. The shift in the vote share of ‘clean’ candidates implies a persuasion rate of 15%, which indicates that 1 in every 7 citizens viewing the advertisement changed their vote.

For this second experiment we partnered with an ONG, called the *Misión de Observación Electoral* (MOE), which hosts a cellphone and user friendly reporting website and is also the most popular reporting channel in Colombia. In contrast to the first experiment, we find a large effect of the information message about the MOE’s reporting website on citizens reports, which suggests that the technical issues in the AG’s website were important in explaining the small effects found in the first experiment. More precisely, we find that the information message increased the reports made through the MOE’s website by 33% or, equivalently, by 0.6 standard deviations. Moreover, we find that it was not only the total number of reports that increased, but also the subset of reports containing hard-evidence about the occurrence of irregularities, which are the ones that are useful from a policy perspective.

In this second experiment we also cross-randomized whether candidates running for mayor in certain municipalities would be informed through letters about the advertisement campaign. In line with the findings from the first experiment, we find that sending these letters reduced the occurrence of electoral irregularities by approximately 0.15 standard deviations, and that this effect was driven by illicit political advertising and fraud in voter registration – the last of which is often done in conjunction to vote buying.

This paper makes contributions and builds on at least five strands of literature. First, we contribute to the literature that studies ways to fight electoral irregularities. Two broad strategies have been studied to counter them. One first strand in the literature has evaluated campaigns aimed at mobilizing civil society against electoral irregularities, either through education campaigns against vote-buying (Vicente, 2014; Hicken et al., 2018; Blattman et al., 2019) and electoral violence (Collier and Vicente, 2013), or through campaigns that explicitly try to persuade citizens to vote against vote-buying candidates (Green and Vasudevan, 2016). Alternatively, second strand in the literature has examined how effective top-down monitoring approaches are. In particular, a long tradition of papers have examined the effects of domestic and international electoral observers (Hyde, 2007; Hyde, 2010; Enikolopov, Korovkin, et al., 2013; Ichino and Schündeln, 2012; Leeffer and Vicente, 2019), and a more recent literature has examined the use of technological innovations to monitor voting aggregates (Callen and Long, 2015; Callen, Gibson, et al., 2016). We contribute to this literature by studying the effectiveness of bottom-up monitoring technologies, which constitutes a third approach that combines elements from both the mobilization strategies and the

monitoring strategies, and that has been properly studied to our knowledge.⁴ Moreover, we follow Blattman et al. (2019) in analyzing efforts against electoral irregularities as a general equilibrium issue which might give way to substitution in strategies by affected parties. The difference, however, is that we focus on the substitution between types of electoral irregularities, while Blattman et al. (2019) study the spatial spillovers of an citizen education campaign against vote buying.⁵

Second, it contributes to the literature on bottom-up monitoring of public good provision and services by studying the effects of a type of e-governance platforms to monitor elections. Earlier papers concentrated on offline interventions promoting citizens' oversight of public goods in the areas of education, health, public works and sanitation, but few had directly addressed the monitoring of elections using ICT-enabled technologies.⁶ Three exceptions are worth mentioning. First, Driscoll and Hidalgo (2014) show that an information campaign aimed at educating citizens about how to file formal complaints about electoral irregularities around elections in Georgia increased electoral irregularity reports but depressed turnout, which they interpret as a consequence of citizens belief that they were being monitored by the regime or by researchers and that retaliation might ensue. Second, Blair et al. (2019) test whether SMS messages encouraging the reporting of electoral misdeeds or a film featuring characters doing so can spur citizen reporting in Nigeria, and find a substantial increase in reports from both interventions. Third, Ryvkin et al. (2017) study the effectiveness of bribe reporting platform in the lab, modeled on *I paid a bribe*, an Indian online website, and find that these sorts of initiatives might be improved by disclosing specific information about bribes. This paper makes a twofold contribution to this literature. First, it corroborates the positive effects of bottom-up monitoring campaigns in the context of electoral integrity. Second, it brings attention to the salience effects that are inherent to the information campaigns typically used to mobilize citizens in these interventions. The latter, to the best of our knowledge, have not been studied yet, and, as our paper shows, are drivers of behavioral change that may or may not complement the effects that are expected from these monitoring tools.

Third, it adds to the literature on the unintended and side effects of communications campaigns. Early contributions to this literature come from health campaigns that cause unintended

⁴A few papers study the effects on bottom-up monitoring interventions, but they do so only tangentially and their focus does align with our paper's. Aker et al. (2017) study whether different interventions, including newspapers, information about elections and access to an electoral reporting hotline increase political accountability, measured by turnout, voting patterns and text messages sent to the elected president in Mozambique. Their focus on direct measures of accountability, as well as their design diverges from ours. Gonzalez (2019) uses a regression discontinuity design to show 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.

⁵Other papers that study spatial spillovers in the context of interventions that randomly assigned electoral observers are Ichino and Schündeln (2012) and Asunka et al. (2019).

⁶Fox (2015) reviews this early literature which found mixed results of the effects of bottom-up monitoring. Some prominent early papers related to the topic of monitoring public servant malfeasance and leakage are Olken (2007), which examines the impacts of grassroots monitoring on the construction of public works in Indonesia, Reinikka and Svensson (2004) and Reinikka and Svensson (2011) which study a newspaper campaign to disseminate information about education fund capture in Uganda.

effects such as boomerang effects, desensitization with the issues advertised, culpability or ‘social norming’, among others.⁷ The results from our first experiment show that bottom-up monitoring campaigns can increase the salience of the issues that they try to combat, in a way that generates responses from citizens other than just the ones that the campaign tries to promote. In this respect, it echoes the finding in Chong et al. (2015) that informing voters about past corruption of politicians depresses turnout and creates democratic disengagement. By highlighting the importance of studying issue salience it also addresses a long literature in political science that has argued that the salience that specific issues have – understood broadly as how ‘top-of-mind’ they are – determines citizens voting behavior (see Dennison (2019) for a review of this literature). In particular, our finding that voters respond to the salience of electoral irregularities by voting for candidates campaigning on anti-corruption appeals echoes Klačnjak et al. (2014) who argue that the emergence of anti-corruption parties increases the salience of corruption in a way that depresses the vote share of incumbent parties.

Fourth, it contributes to the growing literature studying the effects of social media campaigns on elections, which has typically failed to show large and significant results of these campaigns. Methodologically, we build on Broockman and Green (2014) who randomly expose clusters composed of county and age brackets to political ads around legislative elections in the US and find no effects on ad recall, nor on candidate recognition or favorability. Similarly, Kobayashi and Ichifuji (2015) randomly expose Twitter users to tweets from Osaka’s mayor and find no effects on votes for his party. Bond et al. (2012) study a massive Facebook “get out the vote” campaign and find no effects on turnout except when users were exposed to pictures of friends who had reported voting on the platform. While most of this literature finds null results,⁸ we find that our campaign had significant and large results on vote shares, suggesting that information about salient issues can mobilize voters in a stronger way than campaigns explicitly designed to persuade voters in a certain way.⁹

Finally, it adds to the literature that has studied how the expansion of ICT-technologies has facilitated collective action and protests. Theoretical work had argued that information, in general, and social media, in particular, might facilitate protest (Barberà and Jackson, 2019; Little, 2016). Fergusson and Molina (2019) use the release date of Facebook in specific languages to instrument the its use in different countries and show that Facebook penetration causes civilian protests. Acemoglu, Hassan, et al. (2018) show that discontent in Twitter is a predictor of protests during the Arab Spring in Egypt. Enikolopov, Makarin, et al. (2018) show that penetration of a social media platform in Russia led to more protest activity.¹⁰ Our paper contributes to this literature by

⁷See Cho and Salmon (2007) for a review of this literature and Boyle et al. (2017) for a similar review in the case of human right promotion campaigns.

⁸More recent papers finding null or borderline significant results from using ads to influence voting outcomes are Hager (2019) and Haenschen and Jennings (2019).

⁹One exception, coming from a slightly different literature, are Enríquez et al. (2019) who show that Facebook ads informing citizens about past audits in Mexico decrease vote share for corrupt incumbents.

¹⁰Relatedly, García-Jimeno et al. (2018) show that Temperance Crusade protests and events in the 19th century

showing that even short online advertisement campaigns can trigger protests in the medium-term.

The remainder of this paper is organized as follows. Section 2 outlines a model that illustrates the *saliency* and *informational effects* of our intervention. Section 3 provides an overview of the context of the intervention, including a discussion of the most common types of electoral irregularities in Colombia, as well as an overview of the reporting mechanisms available and the elections around which our interventions were deployed. Section 4 describes both the experimental design and the data used in both experiments, while Section 5 presents and discusses the main results. Finally, Section 6 discusses the relevance of the findings from a policy perspective and concludes.

2 Theoretical Framework

In this section we present a simple model which illustrates the potential informational and saliency effects that our intervention had. In the model, citizens have access to a reporting technology and politicians decide how much to invest in illegal and legal ways of getting votes. The campaign in our intervention is modeled to have two effects: (1) it reduces the cost of reporting, (2) it increases the saliency of electoral irregularities, which then benefits the candidates that are perceived to be more honest. The comparative statics of the model will allow us to outline a series of predictions that we seek to validate with the experiments described in the following sections.

2.1 Model Setup

We consider an election contested by two parties indexed by $i \in \{1, 2\}$ and a unit mass of voters, indexed by $j \in [0, 1]$. Each party chooses a triplet $S_i = \{L_i, O_i, U_i\}$ of how much (legal) campaigning (L_i), observable electoral irregularities (O_i) and unobservable electoral irregularities (U_i) it spends on to maximize its probability to win the election. The unit cost of each of these types of expenditures are c_L, c_O and c_U , respectively. For simplicity, we assume these costs are the same for both parties.

Observable electoral irregularities generate a marginal probability $p(R)$ that parties get caught, which is a function of the total number of reports, R , generated by citizens. We assume that this function is increasing and concave, $p(0) = 0$, and that it satisfies a pair of Inada conditions: $\lim_{R \rightarrow 0} p'(R) \rightarrow \infty$ and $\lim_{R \rightarrow \infty} p'(R) \rightarrow 0$.

If the party gets caught, it is fined by an amount $k > 0$, which is proportional to the amount of observable electoral irregularities performed. Without loss of generality, we assume unobservable electoral irregularities do not generate a probability of getting caught. An interpretation of this assumption is that this probability is included in the cost c_U but it does not depend on the amount of reports made.

US were less likely to spread to neighboring towns when railroad strikes and accidents did not happen, highlighting the important role of information transmission through rails and the telegraph on collective action.

The parties' payoff functions are thus given by:

$$\Pi_i(S_i, S_{-i}, R) = \lambda Pr(\text{Win}_i | S_i, S_{-i}) - c_L L_i - c_O O_i - c_U U_i - p(R) O_i k \quad (1)$$

where λ denotes the rents from being in power. Without loss of generality we normalize $\lambda = 1$.

Citizens derive utility from the expected sanctions received by both parties, $\sum_{i=1}^2 p(R) O_i k$. This can represent citizens' preferences for justice, but it can also be interpreted as a reduced form way of capturing the prospective stream of utility citizens get from the fact that punishing parties would not allow them to run in the future.

Citizens have two decisions. First, they decide how much effort they put on reporting, r_j , which has a unit cost of c . The total amount of reporting is thus given by:

$$R = \int_0^1 r_j dj$$

Second, they decide which party to they vote for, $v_j \in \{1, 2\}$.

Voting for party i gives them a utility given by $F(S_i) + \sigma_j^i + \delta^i$, where $F(S_i)$ represents the 'popularity' generated by this party, σ_j^i is a independent random popularity shock and δ^i is a common popularity shock for party i .

We assume $\sigma_j \equiv \sigma_j^2 - \sigma_j^1$ is uniformly distributed on $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$, and $\delta \equiv \delta^1 - \delta^2$ is uniformly distributed on $\left[-\frac{1}{2\psi} + \mu, \frac{1}{2\psi} + \mu\right]$. Parameters $\phi > 0$ and $\psi > 0$ determine the variance of the distribution of shocks, while μ represents the mean of the relative popularity shock for candidate 1.

Moreover, μ in turn depends on the perceived attributes of parties, which are exogenously determined.¹¹ We assume two broad groups of attributes for each party. First, there is an attribute we call 'honesty', which represents how 'clean' parties are perceived to be by citizens, which is denoted by H_i . Second, parties have 'other attributes', which are denoted by A_i .

The relative *saliency* of honesty compared to the other attributes will determine μ . In particular we assume μ is an increasing function of the relative attributes of party 1 compared to party 2, that takes the following form:

$$\mu = g(\omega(H_1 - H_2) + (A_1 - A_2))$$

where $\omega > 0$ captures the *saliency* of honesty compared to other attributes of candidates,¹² and we assume $g'(\cdot) > 0$. Thus, candidate 1 will on average be more popular if she is perceived to be more honest or to have better other attributes than candidate 2, and the relative importance of these attributes depends on their saliency.

The voters' payoff function is then:

¹¹The fact that parties cannot alter their attributes is reasonable in our setting given that the intervention happened so close to the elections that it didn't give candidates time to adjust their image in response.

¹²This way to model attribute saliency follows the classic model of Bordalo et al. (2013).

$$U_j(r_j, l_j, O_1, O_2) = \eta \sum_{i=1}^2 p(R) O_i - cr_j + \sum_{i=1}^2 \mathbb{1} \{v_j = i\} [F(S_i) + \sigma_j^i + \delta^i] \quad (2)$$

where η is a parameter governing how much citizens care about punishing parties engaged in observable irregularities.

The timing of the model is the following:

1. Parties choose their campaigning strategies, S_i for $i = 1, 2$, and simultaneously citizens choose their reporting, r_j for $j \in [0, 1]$.
2. Popularity shocks, δ and σ^j are realized, and citizens vote.

Finally, to get a closed form solution we will assume that $F(S_i)$ has a CES form, namely:

$$F(S_i) = (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}}$$

with parameters such that $\alpha < \gamma < 1$, which ensures the concavity of this function with respect to its arguments.

2.2 Equilibrium

We will restrict the following analysis to Nash equilibria (NE) in pure strategies for conciseness.

Proposition 1. There exists a unique pure strategy NE of the game, $[\{L_i^*, O_i^*, U_i^*\}_{j=1,2}, \{r_j^*\}_{j \in [0,1]}$ in which parties $j = 1, 2$ play identical best responses $S^* = S_i^*, i = 1, 2$ given by:

$$L^*(r^*) = \left(\frac{\psi\alpha}{c_L} \right)^{\frac{1}{1-\alpha}} \left(1 + c_L^{\frac{\gamma}{1-\gamma}} \left[\frac{1}{(c_O + p(r^*)k)^{\frac{\gamma}{1-\gamma}}} + \frac{1}{c_U^{\frac{\gamma}{1-\gamma}}} \right] \right)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}} \quad (BR_i - 1)$$

$$O^*(r^*) = \left(\frac{\psi\alpha}{c_O + p(r^*)k} \right)^{\frac{1}{1-\alpha}} \left(1 + (c_O + p(r^*)k)^{\frac{\gamma}{1-\gamma}} \left[\frac{1}{c_L^{\frac{\gamma}{1-\gamma}}} + \frac{1}{c_U^{\frac{\gamma}{1-\gamma}}} \right] \right)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}} \quad (BR_i - 2)$$

$$U^*(r^*) = \left(\frac{\psi\alpha}{c_U} \right)^{\frac{1}{1-\alpha}} \left(1 + c_U^{\frac{\gamma}{1-\gamma}} \left[\frac{1}{(c_O + p(r^*)k)^{\frac{\gamma}{1-\gamma}}} + \frac{1}{c_L^{\frac{\gamma}{1-\gamma}}} \right] \right)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}} \quad (BR_i - 3)$$

And citizens play identical best responses, $r^*(O^*) = r_j^*(O^*), \forall j \in [0, 1]$, which take the following form:

$$r^* = (p')^{-1} \left(\frac{c}{2\eta O^* k} \right) \quad (BR_j)$$

Proof: Proofs are contained in the appendix.

2.3 Comparative Statics of the Effects of the Intervention

We expect our intervention to have two distinct effects on the model. First, the Facebook ads might have the effect of reducing the costs of reporting, which are represented by parameter c in the model. These reductions in costs might come through a reduction in information costs involved in knowing how and where to report, but might also involve reducing the risks of reporting, since online reporting is more anonymous than in person reporting. Second, the intervention might increase the salience of electoral irregularities, which might benefit candidates which are perceived to be more honest or ‘cleaner’. We model this as a shock to parameter ω , which determines the mean of the popularity shock δ . Notice we do not model these changes in preferences to be associated with candidates’ actual engagement in electoral irregularities, but rather model this as a subjective and psychological shock. We do so because citizens’ perceptions about how honest each candidate is – which determine their vote choice – might be uncorrelated with how clean a candidate actually is.

The following result characterizes the comparative statics related to our intervention:

Proposition 2. (Predictions about the intervention) In the NE defined in Proposition 1, the following comparative statics with respect to parameters c and ω hold:

1. r^* is increasing in $-c$
2. O^* is decreasing in $-c$, while L^* and U^* are increasing in $-c$ if $0 < \alpha < \gamma < 1$.
3. The vote share for party 1 (party 2) is increasing (decreasing) in ω if and only if $H_1 > H_2$, and the converse is true if $H_2 > H_1$.

Proof: Proofs are contained in the appendix.

Part 1 of this proposition simply states the fact that decreasing the costs of reporting will increase the number of reports.

Part 2 tells us that a decrease of the costs of reporting will decrease the incidence of observable electoral irregularities but it might increase the efforts that parties put on legal campaigning and unobservable electoral irregularities if the parameters in the popularity function $F(\cdot)$ are such that $0 < \alpha < \gamma < 1$. Intuitively, this happens because a decrease in the reporting costs increases citizen reports, which in turn increases the probability that parties get caught when performing observable electoral irregularities. In response, if they substitute towards legal campaigning and unobservable electoral irregularities, which are not subject to punishment, if there is enough substitution between the different types of campaigning strategies, which is guaranteed by the condition on parameters α and γ . It is important to note that we only expect these effects to occur when candidates are informed about the increase in reporting. We will argue in the following sections that this is the case only in the treatment group in which we inform politicians about the Facebook ad campaign.

Part 3 of the proposition tells us that an increase in the salience of electoral irregularities might generate an increase in the vote share for candidate 1 if he is perceived to be more honest than

candidate 2, and the reverse is true if candidate 2 is perceived to be more honest. We expect this effect to occur only in municipalities which receive the Facebook ads since the ad campaign rises people’s awareness of electoral irregularities.

3 Context

3.1 Electoral Irregularities in Colombia

Electoral irregularities take many forms and permeate every election in Colombia’s democracy. Despite this, only a few studies shed light on the extent of the problem. Fergusson, Molina, and Riaño (2017) use a list experiment to elicit Colombians’ engagement in clientelist practices (broadly defined as receiving particularistic benefits in exchange for their vote) in a way that overcomes the issue of social desirability bias associated with these type of questions. They find that approximately 18% do so at some point in their lives¹³, and this number is larger for rural and poor respondents . 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.

In order to describe the main types of electoral irregularities used in Colombian elections it becomes necessary to define what we mean by ‘electoral irregularities’ since this term between legal and cultural contexts. Throughout this paper we will use the terms ‘electoral irregularities’ and ‘electoral corruption’ interchangeably to mean any conduct affecting elections that is penalized by Colombian law. There are over fifteen such types of conduct typified in the Penal Code (Law 599 of 2000). Figure 1 shows the most important types of irregularities as approximated by the number of reports made to the government’s unified reporting unit, URIEL, in the 2014 congressional and presidential elections as well as the 2015 mayoral elections. Taken together, the seven types of irregularities shown in this figure represent over 90% of the total reports about electoral irregularities. A rough definition of each of these irregularities is the following:

Vote buying: Also called ‘voter corruption’ by Colombian law, it refers to any attempt to get citizens to vote in a particular way in exchange for money or any type of gift.

Campaigning by public servants: It occurs when public servants attempt to interfere in elections by either trying to favor or harm a particular candidate or party, or when they join a political organization.

Illicit political advertising: Political advertisement is forbidden on election day and is only allowed on the three months prior to election day. It is also forbidden to place ads on public infrastructure such as light posts or monuments.

Fraud in voter registration: This occurs when citizens register to vote in a polling station located in a municipality or district different from their place of residence in order

¹³Similar studies, such as Gonzalez-Ocantos et al. (2012), show similar numbers for other countries in the region.

to obtain an illicit profit or to alter electoral results. This is usually done as a way to facilitate vote buying as explained below.

Voter intimidation: It occurs when someone threatens citizens verbally or physically to vote in a particular way (or not to turnout).

Voter deception: It occurs when someone deceives a citizen to vote in a particular way. Examples include deception about the mechanics of the voting process (e.g. “Blank votes are added to the strongest candidate”) or about candidates belonging to a certain political party.

Electoral fraud: This occurs when electoral results are altered after elections have taken place or by means different to violence, vote buying, deception, such as ballot stuffing.

As shown in Figure 1, vote buying, campaigning by public servants and illicit political advertising represent over 60% of the reports about electoral irregularities. This might be an indicator of how widespread these types of irregularities are, but it might also reflect the fact that these conducts are also more visible, and thus more likely to be reported, than other types of irregularities, such as electoral fraud and fraud in voter registration, which are less conspicuous. Even though the relative distribution of types of electoral irregularities remains similar across types of elections, it is important to mention that the total number of reports in mayoral and congressional elections is substantially larger than in presidential elections: while URIEL gathered approximately 2600 and 2800 reports about electoral irregularities in the 2014 congressional and 2015 mayoral elections, respectively, it only gathered less than 1600 reports in the presidential elections. Interviews with the MOE’s staff reveal that this difference in the number of reports across elections also corresponds to an actual difference of electoral irregularities across the country.

The organizational details about how electoral irregularities are carried out vary according to the type of irregularity. As has been 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 that intermediate between political organizations and voters. These brokers play the important role of providing political organizations with the local information necessary to target and recruit potential voters in clientelistic relationships, as well as in ensuring that these voters actually vote in the intended way. A very common form of ensuring client’s compliance that has been studied both by academics and journalists is by registering voters in polling stations outside of their place of residence, so that brokers can control the votes of their clients (Rueda, 2017; Ardila, 2018).¹⁴ This process is explained in the following way by a broker who worked around Cartagena, in the north of the country, to journalist Laura Ardila:

“Look, doctor, you have to understand that if a leader comes from a particular place it is

¹⁴Evidence from other contexts shows that brokers ensure vote buying by targeting reciprocal individuals (Finan and Schechter, 2012) or, alternatively, by buying turnout from voters likely to sympathize with the candidate supported by the broker (Nichter, 2008). Both of these strategies might also be relevant in Colombia.

useless that his voters vote in that same place. They have to be moved to other areas because that's how one controls them.

I thus tell them 'Do your people vote in El Bosque neighborhood? Well in that case you have to get them to vote in Manga, and you have to tell them in which polling station they are going to vote. If I gave 20 million pesos in exchange for 20 votes, then they have to appear there'." (Ardila, 2018, pp. 47-48)

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 such as mayors, who often collude with running 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, Robinson, et al., 2013), but also by non-armed actors, such as employers who threaten their employees or school principals who threaten parents to lose their jobs or their children's spots in schools if they did not vote in a particular way.

The timing of electoral irregularities also depends on the type of irregularity. Electoral fraud necessarily occurs after election day, and voter registration fraud also occurs during the voter registration period, which ends 3 months before election day. Most types of irregularities, however, occur throughout the pre-election period and up to election day. Such is the case of vote buying, voter intimidation or campaigning by public servants. For instance, as illustrated by the previous recounting of how moving voters to different polling stations helped to buy voters, clientelism and vote buying are irregularities that begin with several months of anticipation to the elections. This is a subject that we will return to when discussing the results of the first experiment.

3.2 Electoral Watchdogs and Reporting in Colombia

Several governmental agencies and NGOs run online electoral reporting channels in Colombia. One of the first and most successful was created by the *Misión de Observación Electoral* (MOE), an NGO whose institutional mission is to promote civil society's engagement with democracy, monitor elections, promote knowledge about and compliance with political rights and to advance research on these same topics. The MOE's reporting website, called *Pilas con el voto* (translated roughly as "keep an eye on your vote"), has been in place since the 2011 elections. Since the MOE does not have the power to directly investigate and take legal action about these reports, it acts as an intermediary between civil society and the government by preparing official reports based on the information provided by citizens and redirects them to the government's unified reporting unit, the URIEL. This unit then processes these reports and sends them to the particular agencies in charge of investigating their claims about electoral irregularities and sanctioning them. The MOE's reporting website has been so popular that it has been responsible of more than 80% of the reports collected by URIEL in recent years.¹⁵

¹⁵Reports redirected by the MOE represented 83% of reports held by URIEL in 2015 and 90.6% in 2011.

The Office Attorney-Inspector General of Colombia¹⁶ (henceforth, AG), also hosts a more recent and less popular reporting website.¹⁷ The AG is an independent institution that oversees the correct conduct of public servants through both preventive faculties and the faculty to sanction. They collect reports about every type of electoral irregularities, but only investigate the ones pertaining to their competence (mainly disciplinary offenses, such as public servants intervening in politics) and redirect the remainder to the competent agencies.

Both the MOE's and the AG's reporting websites share a number of basic features but also have some important differences. Both websites allow users to submit their reports anonymously – making it optional to specify people's names, addresses or email addresses to receive notifications about the status of their reports – and only require them to specify the date and municipality of the irregularities reported. Additionally, they require the users to describe the facts in a free-form field, which is only afterwards classified by their staff as a report about one (or several) of the electoral irregularities typified by Colombian law.

Four main differences arise between the two websites that are important to mention. First, the MOE's staff classify reports by the quality of the report given, into three categories: high, medium and low quality reports. This classification depends on the amount of evidence and facts (places, names and proof such as videos) provided about the electoral irregularities reported. Second, the AG's reporting website requires that users identify the actors of the electoral irregularities reported, by providing their name and their 'affiliation' (such as a public institution or political party) . Similarly, the AG's website requires users to read and agree to a series of legal agreements (three in total) that the MOE's website does not. Finally, the AG's reporting website is not optimized for cellphone use – i.e. it can be accessed through smart-phones but it is hard to navigate. These last three features make the AG's website harder to navigate than the MOE's, and were one of the main reasons why we used the MOE's website for our second experiment.

Both reporting websites have been promoted through campaigns on social media by both the MOE and the AG. In recent elections, for instance, the MOE has spent approximately \$600 USD monthly in the seven months prior to the elections on Facebook, Twitter and Google ads to promote their reporting website. The AG, on the other hand, does not buy ads to promote reports, but it rather uses 'organic' posts on its popular social media accounts for this purpose. In addition to social media, both organizations have also used advertisements on other media sources such as national TV channels, radio stations and even in movie theaters.

The question of how effective reporting is in this context is complicated to answer for several reasons. To begin with, a large fraction of reports do not contain enough evidence for the electoral

¹⁶The spanish name for this institution is the *Procuraduría General de la Nación*. There is a second institution called the *Fiscalía General de la Nación* which is commonly translated as the Attorney General. Both institutions share attributes that are concentrated in only one institution in countries like the US, but are separate institutions. In order to make it easier for readers outside of the Colombian context we have decided to use the AG acronym for the former institution, but the existence of the latter institution should be kept in mind.

¹⁷For instance, the AG collected only 437 reports for the congressional elections of 2018, out of which only 96 came from their reporting website, while the MOE collected over 4000 reports for the same election.

watchdogs to start a judicial case. Similarly, some reports are directed to agencies whose competence does not include the irregularities reported. Finally, many reports are duplicates of other reports and these cases of duplicity are not reported as such in the electoral watchdog’s data sets. Notwithstanding these difficulties, in Section A.1 of the Appendix, we present an analysis of the reports gathered by the AG in recent years and show that at least 2.5% of reports ultimately lead to a judicial decision.

3.3 Obstacles to Reporting

How likely are people to report electoral irregularities when they have witnessed them? Moreover, what obstacles prevent them from reporting? In an online survey we conducted in June 2018, we included a set of questions that allowed us to shed some light on these issues. For the full details about this survey the reader can refer to Section 4.3.1.

Out of 392 respondents included in our survey¹⁸ approximately 15% of them admitted to having witnessed electoral irregularities in the past, but only 8% of them (i.e. approximately 1.2% of the 392) had reported them.

Why didn’t the remaining 92% of respondents report the electoral irregularities they had witnessed? Figure 2 plots the responses to this question. Respondents were given a set of four possible answers, and they were asked to choose all of the options that applied. The first most popular answer, with 43% of mentions, was that respondents did not know where to report the electoral irregularities they had witnessed. A second close response, with 38% of mentions, was that respondents had not reported them because they were afraid of doing so - which is natural in a weakly institutionalized context such as the one at hand, in which disclosure of this information might result in unpunished reprisals from the accused. “Other reasons” were the third most mentioned reason, with 25% of mentions. Although this is only speculative, one such additional reason that might be included in this category is that citizens believe that reports are ineffective since authorities do not sanction offenders, as a study by the MOE has found (Misión de Observación Electoral, 2018). Finally, the least mentioned reason, with under 10% of mentions, was that survey respondents did not have the time to report the irregularities they had witnessed.

Although the conclusions one can draw from this survey are limited, the results are quite sharp in illustrating that even among a population with access to the internet, knowledge about how to report electoral crimes is poor. This provides an initial piece of evidence that an information campaign, such as the one included in our two interventions might be useful in boosting reports and citizen monitoring of elections.

¹⁸This only includes respondents from municipalities in our control group, so that their responses were not affected by the treatment conditions.

3.4 The 2018 Presidential Election

In the year 2018, both legislative and presidential elections took place in Colombia. In mid-March, the congressional elections took place to elect the members of the House of Representatives (the lower house) and the Senate (the upper house). A couple of months later, in May 27, the first round of presidential elections took place.

Five main candidates participated in this round of elections.¹⁹ Table A.7 in the appendix summarizes the position of each candidate in the political spectrum and the coalition of parties supporting them. We also categorize in this table whether candidates were ‘traditional’ or not. We define ‘traditional’ candidates as those supported by parties that have had a substantial power in national and local politics (e.g. as measured by their presence in congress) both presently and in the past decade or longer.

Out of the five candidates that participated in these elections, two of them are non-traditional, according to our definition, and three of them were traditional. In Section A.2 of the Appendix we give more background about the candidates and we further discuss our categorization of candidates.

For the purpose of interpreting the results later on, it is also important to mention that non-traditional candidates, Petro and Fajardo, had made the fight against corruption and clientelism an important part of their campaigns, and in their discourse they closely linked traditional and old-fashioned politics to corruption.²⁰ Either by persuasion from these candidates’ campaigns or because of the bad name traditional parties have, voters also perceived the non-traditional candidates as less corrupt. In the online survey we conducted around our first experiment (see Section 4.3.1 for details) we asked approximately 400 respondents who they thought was the best candidate to fight corruption. The two non-traditional candidates were the two first to get the most mentions, with a combined 75% of responses, with the non-traditional candidates getting the remaining 25% (Figure A.2).

In Colombia’s two-round election system, there is a run-off between the two candidates with most votes in the first round, unless there is a candidate with more than 50% of votes, in which case that candidate wins without a second round. In the 2018 elections this was not the case: Duque and Petro went on to the second round of elections with party alliances forming along the traditional/non-traditional party divide.²¹ Duque was the winner of this second contest getting 54% of votes over Petro’s 42%.

¹⁹A sixth candidate, Jorge Antonio Trujillo, received less than 0.5% of votes and will thus be omitted from this overview.

²⁰Petro, for instance, not only consistently claimed that corruption was the main tool that kept traditional politics in power (see, for example, El Tiempo (2018)), but even accused President Santos and candidate Vargas to prepare a plan to rig the software used to aggregate votes in favor of Vargas. Similarly, Fajardo described his platform as one “in opposition to the traditional clientelistic model” and established as one of its ‘poster child’ policies the fight against corruption (Sergio Fajardo’s Campaign Team, 2018).

²¹Although Fajardo did not himself adhere to Petro, one of his supporting parties, the *Polo Democrático Alternativo*, and a faction of his other main parties, the *Alianza Verde*, did. Similarly, Vargas and De la Calle did not adhere to Duque, but some of their supporting parties, the Conservative and the Liberal parties did.

3.5 The 2019 Mayoral Elections

Local elections were held in October 27, 2019. In these elections, voters chose not only the mayors, but also the council-members in each municipality, and governors and department-level legislators. The candidates with a simple majority of votes won the election in a single-round system.

We decided to focus on the voting behavior for mayors since this post was assigned at the municipality level (which the level at which we performed randomization), and the number of candidates is tractable compared to the council-members, which often had several dozens of candidates.

On average, municipalities in our sample had 5 candidates for mayor, with a minimum of one and a maximum of 13. Many candidates ran as independents or as part of large coalitions between parties, which made it difficult to identify ‘traditional’ candidates as we did in the first experiment. As explained in later sections, we instead rely on citizens’ self-reported perceptions about ‘clean’ candidates to identify the candidates they might vote for in response to the advertisement campaign.

4 Experimental Design

4.1 Study Sample

The sample for the first experiment consisted of 652 municipalities coming from every Colombian department (see Figure 3). This sample was defined as the subset of Colombian municipalities which had at least 1000 active Facebook users and no more than 50’000. This lower bound of users is used because Facebook’s Ad API does not report user populations in areas with less than 1000 users. The upper bound on the number of users was chosen to keep the costs of the ads within our budget.²²

For the second experiment, we defined a slightly larger sample of 681 municipalities chosen according to two criteria: (1) they had a population of people over 18 years old of more than 5000 and less than 97’000, and (2) they had to have a significant ad delivery in the first experiment, meaning that the ad reached more than 5% of the Facebook ad users. These criteria were set given the same considerations mentioned for the first experiment.²³ The set of municipalities chosen this way overlapped considerably, but not fully, with the ones from the first experiment and the resulting sample was quite similar to it in terms of their observable characteristics.

Table 1 presents the summary statistics for a selected set of variables for the municipalities included in the samples for the first and second experiments. As seen in this table, the character-

²²We tried to reach a constant proportion of the population in each municipality (approximately 30% of the Facebook users).

²³For the second experiment we used actual population instead of the number of active users provided by Facebook to define the bounds on population size given that, as proved by our experience from the first experiment, some of the Facebook user estimates seemed to diverge substantially from the final number of people reached by the ad.

istics of the municipalities in each of the experiments are very similar. The average municipality in both experiments has approximately than 25 thousand inhabitants, but the variation in the sample is large in both cases. The municipalities have a relatively large access to Facebook, with over 40% of the population reported as active users by Facebook on average. Despite this large access to Facebook, a large amount of their population is poor and rural: GDP per capita is on average 14 millions pesos (approximately \$4,500 US dollars), over 40% of the population is considered poor and over 50% live in rural areas according to data from the *National Department of Statistics* (DANE).

The average municipality reported electoral irregularities moderately during the 2018 Congressional elections which occurred a couple of months before our first experiment – it sent approximately 0.5 reports to the MOE and 0.14 to the AG – but this conceals substantial variation: some municipalities submitted more than 8 reports to either of these agencies, and some did not submit any reports.

4.2 Experimental Design

4.2.1 First Experiment - 2018 Presidential Elections

Figure 4 provides a map of the different treatment groups involved in the first experiment and the timeline of the different interventions is depicted in Figure 5. Randomization was carried out in two stages in a factorial design that allowed us to maximize power by increasing the sample size per treatment group (Duflo et al., 2008).

In the first stage of the randomization we allocated the 652 municipalities in our sample into two treatment groups.²⁴

TC. Control Group: Municipalities in this group did not receive any ads.

TF. Facebook Ad: Municipalities in this group received a Facebook ad that both informed them about the AG’s reporting website and a hot-line designated to receiving reports, and encouraged them to take action against electoral corruption (e.g. “*If you have witnessed an irregularity or offense in these elections, file your report through the Attorney General’s webpage. Click here: [...] Report! Let’s raise our voices against electoral corruption.*”).

Figure 6 depicts the ad sent, as it was displayed on the Facebook feed of a cellphone user. A few features of the design of the ad are worth mentioning. First, the image associated to the ad – showing a ballot box with the colors of the Colombian flag along with hands crossing hands and lifting flags – was designed so it contained no colors associated to any particular party. We

²⁴We additionally randomly varied the message received in the Facebook ad for half of the municipalities in the sample – so that it included a message highlighting the efficiency of the AG in fighting electoral misdeeds – and within each of these groups we varied the share of Facebook users to be either 50% or 100%. For conciseness, we only present the results for the “pooled” effect of the intervention, but we display the results for the sub-treatments in the Online Appendix.

also made an effort in designing an image that transmitted a positive image about democracy and the prospects of reporting. This is important since, as it will be discussed in further detail in later sections, the ad’s message could also be read in a more pessimistic light, as reminder that elections were not transparent. A second feature of the ad worth mentioning is the inclusion of messages informing readers that the reporting process is easy and anonymous. Especially the former information is important since, as we mentioned earlier, reporting is perceived as a potentially dangerous activity in the Colombian context and informing people that they will not need to provide personal information when reporting might mitigate these concerns. In Section A.3 in the Appendix we further discuss the technical details of how the ads were specifically programmed on Facebook’s Advertisement Manager.

The ad campaign intervention lasted for six days. It started on May 23, four days before the elections, and it ended on the night of May 28, one day after the elections. This extra day potentially allowed citizens who had witnessed electoral irregularities but had not reported them on election day to report them the next day through the AG’s link. The relatively short time span before the elections in which the ad campaign was deployed was chosen to minimize the risk that politicians would find out about the campaign and have time to react to it (which was the objective of the ‘Letter to politicians’ intervention that is explained later on).²⁵

In a second stage of randomization, municipalities in *both* the control group and the ones in any of the treatment groups receiving ads are then cross randomized into the following treatment groups:

TL. Letter to Politicians: All of the candidates running for President and their campaigns managers receive a letter and an email from the AG informing them that the municipalities in this group (which were included in an attached list) might be included in a grassroots campaign to monitor elections.

TN. No Letter: The municipalities in this group were not included in the list sent out to candidates and their campaign managers.

The complete text included in the letters sent to the candidates is displayed in Figure A.1 in the Appendix. Three features of this letter are important to mention. First, since municipalities in the control group were also included in this letter, the letter emphasizes that these municipalities

²⁵As illustrated in Table A.8 in the appendix, in the past two elections many of the main types of irregularities occur before this time span. In particular, over 84% of electoral irregularities reported to the MOE about illicit registration of voters occurs before this period since voter registration is closed three months before the elections. By the time the ad campaign starts most of the illegal campaigning of public servants in favor of candidates is also well underway: 77% of these types of irregularities had occurred before this period in the 2015 local elections, and 60% in the case of congressional elections. Voter intimidation is also an irregularity that begins early on, with 66% and 49% of it occurring six days or more before the elections. An important fraction of vote buying in the 2015 local elections also seemed to start before this period (42%), but in the 2018 congressional elections this fraction was considerably smaller (24%). Only electoral fraud seems to occur predominantly after this period, which is not surprising given that election results are mostly altered before they occur.

might, but not necessarily would be included in the AG’s campaign.²⁶ This phrasing avoids deception or giving misleading information to the candidates. A second important feature of the letter is that it did not reveal to politicians through what way we would conduct the campaign (i.e. it only says it is an online campaign, but it doesn’t say it would be conducted through social media) nor which reporting channel would be promoted through the campaign.²⁷ This was done to prevent the candidates from engaging in *signal jamming*, which they could have done either by interfering with the Facebook ad campaign or by filing false reports in the AG’s reporting website. Finally, since clientelistic networks and political brokers usually operate at a local level (as explained in Section 3), the letter explicitly asked the recipients to pass the information to the campaigns “regional offices” as a way of making this treatment more effective.

The letters were sent in May 16, eleven days before the first round of elections. Due to logistical reasons we were not able to send these letters before, which would have been ideal to maximize the time candidates would have had to react to the news and, perhaps, adjust their campaign strategies. This is one feature of the experiment that was improved in the design of the second experiment, to which we now turn.

4.2.2 Second Experiment - 2019 Mayoral Elections

Figure 7 displays the experimental arms involved in the second experiment and Figure 8 shows the timeline of the different interventions and data collection milestones. As in the first experiment, we used a factorial design, with two stages of randomization. In the first stage we randomized municipalities into four treatment conditions:

TP. Placebo Control Group: Municipalities in this group receive a Facebook ad containing a ‘placebo’ message, reminding people about the elections – “*Don’t forget that local elections will take place on Sunday, October 27.*”

TI. Information message: Municipalities in this group receive a Facebook ad informing them of the existence of 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.*”

TS. Salience message: Municipalities in this group receive a Facebook ad encouraging them to report electoral corruption and to act against it – “*In these elections let’s stop electoral irregularities. Report them! Don’t forget that local elections will take place on*

²⁶Two reasons lead us to include a subset of the control group in the letter sent to politicians: (1) it gave us more power to test the significance of this treatment arm by reaching a 1:1 proportion with respect to the group that was not included in the letter (Glennerster and Takavarasha, 2013); (2) this allowed us to test if politicians reacted to this information in absence of the campaign actually occurring. However, we found no such effects and decided not to report results leveraging these heterogeneous effects in the main analysis.

²⁷The AG’s reporting website is relatively unknown and hard to find so it is unlikely the candidates would have linked this campaign with the website.

Sunday, October 27. ”

TB. Information + Salience message: Municipalities in this group receive a Facebook ad containing both messages in T1 and T2.

The rationale for each of these experimental groups is the following. First, we include a placebo message in the control group to net out the effect of politically-oriented advertisement on citizens’ behavior – which represents a slight change from what we did in the first experiment, where we had a ‘pure’ control group.

Treatments TI,TS and TB are designed to understand whether we can manipulate two different elements which were combined in the first experiment: (1) the cost of reporting, which would be reduced by informing citizens’ about the MOE’s reporting website; (2) the salience about electoral irregularities and the urgency to action against it. As mentioned before, the first experiment did not manage to decrease the cost of reporting due to technical reasons, and thus the effects on electoral outcomes might have been due to the salience effect. Moreover, one hypothesis is that citizens’ reacted to the increased salience of electoral irregularities by substituting between two actions against this issue: instead of reporting (which they were not able to do), they voted for the non-traditional candidates, which they perceived to be ‘cleaner’.

Some details about the overall design of the ads were changed from the ones in the first experiment. First, the ads were sent only three days before the intervention in an attempt to further reduce the scope for politicians to react to the campaign in the treatment groups in which they were not sent letters explicitly informing them about it.²⁸ Secondly, we opted to use videos with slide shows instead of images as part of the main ads in the second experiment. Figure 9 displays the slides used on the different ads, which follow the text mentioned above.

In the second stage of randomization we further assigned municipalities receiving any of the treatments TI,TS and TB to two groups:²⁹

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

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

As in the first experiment, these treatment groups allow us to test whether politicians react to the campaign by changing their electoral strategies – both in respect to their engagement in electoral irregularities and in their (legal) campaigning strategies, as predicted by the model presented in 2. For reference, Figure A.1 shows the letter sent to candidates which follows closely the one from

²⁸The ads also ran the day of the election and the day after, as had been done in the first experiment.

²⁹For the second experiment we did not allow part of the control group to be included in the Letter to politicians treatment condition. The reason was to avoid having to tell candidates that the campaign *might* be occurring in their municipality, which might have weakened the effect of the intervention, compared to the opted phrasing which states that the campaign would certainly occur.

the first experiment.³⁰

In order to maximize the effect of these interventions, we sent the letter to politicians approximately two months before the elections and then we sent out a remainder three weeks before. Both physical letters and emails were sent to maximize the chances of getting the candidates' attention.

4.3 Outcome Variables and Data

4.3.1 First Experiment - 2018 Presidential Elections

For the first experiment we use four main groups of outcome variables:

Reporting to the AG: Our main outcome variable to assess whether this campaign is successful in getting citizens to report will be the number of reports per municipality collected by the AG through both their website and the call center that was advertised through our campaign. We distinguish short and medium term effects of the campaign by considering separately the reports made around the first round of elections from those made 21 days later on the second round of elections.

Electoral irregularities: We are also interested in understanding whether our intervention reduced actual electoral irregularities (independently from whether it affected reporting). Since direct measures of electoral irregularities are unavailable, we instead resort to a proxy of electoral irregularities. We use the reports made to the MOE in the first round of elections that are deemed to be of a 'high quality' (see Section 3 for an explanation of how reports are classified) as proxies for electoral irregularities, and dis-aggregate the reports by type of electoral irregularities to test for strategic responses by politicians.

Voting outcomes: We use the official voting records provided by the *Registraduría Nacional* aggregated at the municipality level to understand the effects of the intervention on elections. We also compare this to the responses collected in the post-treatment survey, explained in the following section.

Protest Participation: Results from the first experiment reveal that the ad campaign led vote share for non-traditional candidates to increase at the expense of traditional candidates. Since the main traditional candidate got elected, protests against the national government following the elections serves as a measure additional of collective action and discontent against traditional politics. The data used is part of the Social Struggle database the collected by the *Centro de Educación y Educación Popular* (CINEP) from local and

³⁰We additionally randomly varied the text sent in the letters to politicians. Half of the municipalities received information in letters about the exact reporting website promoted by the AG, and half of them did not include this information. We pool the results from these treatment variations in the main paper for conciseness, but refer the reader to the Online Appendix for the results and details about these treatment arms.

national media. We use data about protests made against the national government from May 27th, after the intervention, until December 31st (the last date available).

We also collected a rich set of municipal level covariates to test for balance checks and to include in the main specifications as robustness checks. We discuss the covariates included in Section A.4 of the Appendix.

Post-Treatment Survey

After the main intervention in the first experiment was concluded, we administered an online survey to collect data on complementary outcomes which were not available from external sources. Recruitment for the survey was conducted through Facebook, by sending users in the municipalities in the sample an ad inviting them to participate in an survey to collect their “opinion about the past presidential elections”. This type of survey recruitment strategy on Facebook has been studied in both developed and developing contexts in the previous literature (Kosinski et al., 2015; Samuels and Zucco, 2013; Sances, 2019; Zhang et al., 2018). It has been shown to be particularly effective at reaching populations that are costly to reach through conventional survey methods, such as the one at hand (Samuels and Zucco, 2013), and to approximate the representativeness of common recruitment methods such as phone surveys.³¹ These recruitment ads were sent from June 20th, three days after the second round of elections, to June 24th. Take-up was incentivized by raffling a Samsung tablet (valued at, approximately \$120 USD) among the survey respondents.

Our final sample of survey respondents includes 1029 responses coming from 328 municipalities (i.e. an average of approximately 3 respondents per municipality).³² As shown in Table A.1, the sample of municipalities with survey respondents have a larger population, a higher Facebook penetration, are less rural and have a lower percentage of poor than the full sample of municipalities in the first experiment. They have also had a higher number of reports to the MOE, but not to the AG in the past. These differences suggest that extrapolating the results from the survey

³¹Zhang et al. (2018) show that Facebook recruitment approximates the degree of representativeness of traditional phone-surveys if population quotas and post-stratification are used. In our setting we only gathered data about the gender of respondents so that they felt the survey respected their anonymity, so we could not perform post-stratification.

³²Over 1470 responses were collected, but we dropped the following cases for our main analysis: (1) respondents who claimed to reside in municipalities outside of the experimental sample (≈ 350 responses) ; (2) respondents who claimed to reside in municipalities in other treatment arms (≈ 90 responses). The size of these two groups allow us to get an approximation of how accurate Facebook targeting functions are. Overall, Facebook seems to be doing a decent job at targeting the desired municipalities, with approximately 70% of the survey ads delivered to the correct municipalities. This is relatively low compared to the results in Sances (2019), who report a ‘correct’ delivery rate that varies between 81.5% and 99% in several surveys recruited through Facebook ads and targeted at medium sized US cities. However, our estimate is probably a lower bound on how accurate Facebook’s targeting is, since some of the respondents might have misreported their residence location in order to guarantee their anonymity given the delicate nature of some of the questions in the survey. Additionally, respondents in case (1) mostly claimed to be from large cities (over %80 of the cases) which might correspond to people who commute to work to these large cities, but actually live in smaller ones.

sample to the full sample should be done with caution. However, in a later section we show that the sub-sample of municipalities with survey respondents is well balanced across a wide range of observable characteristics.

Three sets of outcomes were collected in this survey:

Reporting and electoral irregularities: We asked respondents whether they had witnessed electoral irregularities, whether they had reported them and, in case they had, to what agency. We also asked them about how effective and how easy they thought the reporting process would be. In case they had witnessed electoral irregularities but they did not reported them, we asked them why they had not done so.

Trust in institutions: We elicited how much respondents trusted the AG, elections, the president, NGOs and the judiciary on a scale from 1 to 7.

Voting and political preferences: We asked respondents whether they had voted and, in case they did, who they voted for. We also asked them who the best candidate to tackle corruption, and what the most pressing problem in the country was.

4.3.2 Second Experiment - 2019 Mayoral Elections

As in the first experiment, the main outcomes of interest are the people’s reporting (in this case, to the MOE), the extent of electoral irregularities and voting outcomes. However, the first two of these outcomes are measured differently in the second experiment and require some explanation.

Reports made to the MOE: We use the aggregate number of reports made to the MOE as a way of testing the citizens’ ‘demand’ for the reporting website. Additionally, we disaggregate the reports by the quality of the reports, which are classified by the MOE as high, medium or low quality, depending on the evidence and the information contained in the reports about the electoral irregularities, as explained in Section 3. This allows us to test whether the campaign successfully manages to induce useful, evidence-backed reports that ultimately can put checks on corrupt behavior, or if it only affects the margin of low quality reports.

In the new experiment, we also change our strategy for measuring electoral irregularities. We conduct both an online pre-treatment survey and a post-treatment survey directly asking citizens whether they experienced different types of electoral irregularities, and which candidate or parties was behind those irregularities. We now go on to explain how each of these surveys were conducted and what variables we collected.

Pre-Treatment Survey

The pre-treatment survey was conducted in the 3 weeks prior to the elections (see Figure 8). Respondents were recruited through a Facebook ad, as we did in the first experiment and participation was also incentivized by including participants in a raffle for several Samsung tablets

(valued at, approximately, \$120 USD). The survey took approximately 10-15 minutes to answer, and the recruitment ad made no reference to its content, nor about the upcoming elections.

The main goal of this survey was to identify the candidates that were perceived to be more or less ‘honest’ by citizens before the intervention. Given the difficulty of characterizing the large number of candidates running in the mayoral elections, this strategy provides a data-driven way of identifying the perceptions of citizens about these candidates. We collected citizens’ responses to two main main questions for this purpose:

Best candidates to fight corruption and vote buying: We asked respondents to name one candidate who they thought was the best at preventing fiscal corruption and vote buying, separately.

Prospects of electoral irregularities by each candidate: We asked respondents how likely each candidate in their municipality was going to engage in the seven main types of electoral irregularities mentioned in Section 3.

Our final sample is made of 6581 complete responses coming from 641 municipalities.

Post-Treatment Survey

The post-treatment survey was conducted immediately after the intervention and lasted for eleven days. We used two recruitment methods. First, we recontacted through email the respondents from the pre-treatment survey who express 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.

Two main outcomes were collected in this survey:

Witnessing electoral irregularities: We asked respondents whether they had witnessed the seven main types of electoral irregularities mentioned in Section 3 in the 2019 local elections.

Voting outcomes: We asked respondents whether they had voted and who they had voted for in the elections.

As in the survey conducted for the first experiment, we collected citizens’ reported trust in several institutions and we asked them about the most important issues in their municipality.

To ensure accurate measures of the occurrence of electoral irregularities we used two complementary strategies. First, we reminded respondents about the importance of collecting their truthful responses about these matters. Second, we asked them how confident they were about their responses.

We gathered 2720 complete responses coming from 625 municipalities, but we restrict our analysis to respondents which say that are at least “somewhat confident” about their responses. This leaves us with 2579 respondents from 621 municipalities. Table A.2, shows that this sample of municipalities with survey respondents are not statistically different from the full set of 681 municipalities in the full sample along a number of observable characteristics.

4.4 Randomization in Practice: Stratification and Balance Checks

In order to increase the likelihood of a randomization balanced on potential confounds, we randomized municipalities within strata in each of our experiments. For the first experiment, we defined these strata using the intersection of the bins generated by partitioning the sample using the 33.4% and 66.7% percentiles of the number of Facebook active users, the 33.4% and 66.7% percentiles of the turnout in the 2018 congressional elections, and an indicator of whether they had made any report to the AG during the 2018 congressional elections. For the second experiment, randomization was carried out within strata defined by the intersection of bins partitioning in the sample in three ways: (1) by the 50% and 85% percentiles of 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; (3) by whether the municipalities filed reports to the MOE around congress elections 2018 above or below the median.

Table A.3 reports the balance checks for the first experiment in 2018, using 29 covariates – including measures for the number of reports collected by the AG and the MOE in previous elections, the socioeconomic covariates and the political covariates mentioned in Section 4.3. We present comparisons of the control group to the Facebook Ad treatment as well as the group for which the politicians received the letter to the group without it. The first three columns of this table report balance checks for the full sample of 652 municipalities, while the last columns present the balance checks for the set of 328 municipalities for which we collected responses for the post-treatment survey.

The results presented in this table suggest that municipalities are well balanced across treatment arms in the first experiment in our main sample, as well as in the municipalities with responses in the post-treatment survey. Only two differences in means out of 58 comparisons are statistically significant at the 10% level for the full sample of municipalities, and four such differences occur in the sample of municipalities with responses in the post-treatment survey. While these imbalances might have only arisen by chance, this justifies including covariates in our main specifications as a robustness check.

Table A.4 displays the balance checks for the second experiment in 2019, using the same 29 covariates used in the first experiment as well as the number of candidates participating in the mayoral elections. We test for mean differences across the multiple treatment arms with respect to the control group, and also the differences between the groups of municipalities receiving the letters sent to candidates to those that did not. Columns (1)-(5) report balance checks for the entire sample of 681 municipalities, while columns (6)-(10) do so for the set of 621 municipalities for which we collected responses for the post-treatment survey. Additionally, Table A.5 reports the same balance checks for the collected demographic characteristics of the respondents to the post-treatment survey.³³

³³We did not collect these variables for respondents in the post-treatment survey in the first experiment, so we can only report the balance checks for the respondents in the second experiment.

Once again, the results suggest that randomization was carried out successfully: only five out of 150 comparisons are statistically significant in the full sample, while three of them are in the sub-sample of municipalities with respondents in the post-treatment survey. The demographic characteristics of the respondents to the post-treatment survey are also balanced, with only one difference being marginally significant.

4.5 Empirical Specification

Our main specification for the regressions in the first experiment takes the following form:

$$y_m = \alpha FacebookAd_m + \beta_1 Letter_m + X'_m \gamma_1 + \phi_m^1 + \epsilon_m^1 \quad (3)$$

in which y_m is the outcome variable for municipality m , $FacebookAd_m$ is an indicator that takes the value of one if municipality m is in the Facebook ad treatment group, $Letter_m$ is an indicator for whether municipality m was included in the letter sent to presidential candidates, X_m is a set of municipal covariates, and ϕ_m^1 are a set of fixed effects for the strata used in randomization and ϵ_m^1 is the error term.

For the second experiment we instead use the following specification:

$$y_m = \alpha_{TI} Info_m + \alpha_{TS} Saliencem_m + \alpha_{TB} Info\&Saliencem_m + \beta_2 Letter_m + X'_m \gamma_2 + \phi_m^2 + \epsilon_m^2 \quad (4)$$

where $Info_m$, $Saliencem_m$, and $Info\&Saliencem_m$ are indicators for whether municipality m was sent the Information, Saliencem or both messages, respectively; $Letter_m$ is an indicator for whether the letter to mayoral candidates was sent in m , and X_m , ϕ_m^2 and ϵ_m^2 are defined in analogous way as in (3).

For all specifications we report Huber-White standard errors (White, 1980) but also randomization inference p-values to allow for inference that does not depend on distributional assumptions or asymptotic theory (Athey and Imbens, 2017; Young, 2017). In most specifications we use the double-post-lasso covariate selection method proposed by Chernozhukov et al. (2015) and Belloni et al. (2014) in order to choose the covariates included in vector X_m without running into overfitting issues. Unless otherwise specified, the set of covariates in Table A.3 is the one considered in this method.

When using survey data, we report results at the individual response level, but we also include individual-level covariates X_{im} , and we cluster standard errors at the municipality level – which is the level of randomization.

5 Main Results

5.1 Ad Campaign Scale and Results on Facebook Metrics

We begin by describing the scale of the Facebook ad campaign and its success in creating engagement by its audience in both experiments.

Table 2 provides a summary of several metrics of the scale of the ad campaigns in absolute terms, as well as the average results per municipality, per capita, and per people registered to vote. In the first experiment, the campaign reached over 1.4 million people and it appeared almost 4.5 million times on peoples’ screens. This implied that the average viewer saw the ads roughly $3^{1/2}$ times. In each municipality it reached on average almost 3300 people, which represents approximately 14% of population or 36% of those registered to vote.

The ad was quite successful in getting viewers to click on the AG’s reporting website link: over 12,000 people clicked on the link, which means that approximately 0.9% of viewers clicked on the link³⁴, and that we got over 28 clicks per municipality on average. This also implies that each click cost about \$0.5 USD. The ad also got substantial engagement from the viewers: in each municipality, on average, 7.5 people municipality ‘reacted’ to the ad (by liking it, loving it, etc...), 2.4 people shared the ad with their friends and 0.3 people commented on the ad.

In the second experiment, the ad campaign had a larger scale, and it created more engagement than the first experiment. It reached approximately 4.4 million Facebook users, which saw the ad on average approximately 3 times. An average of 6400 people saw the ad per municipality, which represents 23% of the population in these municipalities and 31% of the registered voters.³⁵ Engagement was also substantial, with over 23,000 people clicking on the link to the MOE’s reporting website, representing an average of 75 people per municipality, at a cost of approximately \$0.23 per link click. The average municipality had 14 people reacting to the different ads, 6.7 of them sharing the ads and 0.64 of them commenting on it.

5.2 Effects on Reporting

We now move on to study whether the social media campaign was effective in getting citizens to report electoral irregularities.

5.2.1 First Experiment - 2018 Presidential Elections

We begin by analyzing the effects on the first experiment. Columns 1-2 of Table 3 display the effects of the different treatment arms on the number of reports collected through the AG’s website and the call center we enabled in the first round of elections. Since there was no municipality from which the AG received more than one report, we can interpret the effects as percentage point changes. Overall, the ads were successful in increasing citizen’s reports but the effects were modest in magnitude. Receiving any of the ads through Facebook increased the likelihood that a municipality sent a report by 1.4 percentage points (p-value < 0.05).

³⁴In comparison, Broockman and Green (2014) finds a click rate of 0.02% per impression using ads for US legislative candidates, whereas results in Table ?? imply a click rate of approximately 0.03%.

³⁵Notice that in the second experiment, the ad campaigned reached a higher proportion of the population but a *lower* proportion of citizens registered to vote. This explained by the fact that there was a larger increase in voter registration over this time period.

As mentioned in Section 3, the second round of presidential elections occurring three weeks after our intervention allow us to test for medium-term effects of the ad campaign. Columns 3-4 of Table 3 show the treatment effects on the number of reports collected in the second round of elections. While estimates of the effects are all positive, they are not statistically significant from zero. Columns 5-6 report the treatment effects on the combined reports from both the first and the second round of elections. Naturally, the effects are higher than in each of the reports from each round of elections considered separately. On average, the ad treatments increased the likelihood of reports by 1.8 percentage points ($p\text{-value} < 0.01$).

Although significant, the magnitude of the effects of the ad campaign on reporting present a puzzle. As mentioned in Section 5.1, over 12'000 viewers of the ads clicked on the link directing them to the AG's reporting website in the first experiment, but only eight reports were actually received by the AG from treatment municipalities and none from control ones. While it is certainly possible that not all of the people clicking on the link actually wanted to file a report – other possible reasons being curiosity or clicking on the link by mistake – it is hard to think that this by itself explain this click-to-report gap.

The most likely explanation of this gap is, instead, that there was a feature of the AG's reporting website that prevented citizens to file their reports after landing on the website. In particular, as we later found out, approximately 95% of people viewing the ad did so through their cellphones, but the website was not fully cellphone compatible: while cellphone users were indeed able to reach the website it was difficult to navigate, making it often necessary to scroll in every direction to read the text in the website or to find where to click to reach the next set of questions. To get an idea of what this problem looked like, Figure A.3 illustrates how the AG's landing page, with an initial agreement box only partially displayed, was visualized on a cellphone.

This cellphone compatibility issue suggests that the results presented in this section underestimate significantly the effect of our reporting campaign under appropriate conditions. As we will see next, the effects on reporting in the second experiment – which used the MOE's reporting website which was fully cellphone-compatible – are substantially larger, suggesting that small costs of reporting can have important effects on citizen's decision to complete a report.

5.2.2 Second Experiment - 2019 Mayoral Elections

Table 4 displays the estimated effects of the different treatment arms involved in the second experiment on reporting.

Columns (1)-(2) display the results on an indicator for whether municipalities issued any report through the MOE. Municipalities receiving the information message saw an increase in the probability of filing a report of approximately 24 percentage points ($p\text{-value} < 0.01$), which represents a 100% increase compared to the control mean. Municipalities only receiving the salience message also increased the probability of a report by 14 percentage points ($p\text{-value} < 0.05$), which is significantly different ($p\text{-value} < 0.1$), and lower, than the effect of receiving the information message. The fact that there is an increase in the probability of reports from the Salience message group

indicates that despite the fact that reporting remains more costly than in the Information Message group, citizens are still able to find a way to report and do so in response to the campaign. Finally, the letter sent to politicians also decreases the probability of reports by 12 percentage points, which anticipates that the ad campaign might have had an effect on lowering the extent of irregularities (which is in fact what we show in a later section).

In columns (3)-(4) we explore whether the treatment ads also had an effect on the extensive margin by using the inverse hyperbolic sine transformation of the number of reports sent to the MOE.³⁶ We see that the number of reports increased by approximately 33% in municipalities which received the information message (p-value<0.01), and by 21% in the group only receiving the salience message (p-value<0.01). Finally, the letters sent to politicians decreased the number of reports by 15%.

Columns (5)-(8) explore the effects of the interventions only on the reports of a ‘higher’ quality³⁷ to examine whether the intervention also had an effect on evidence-backed reports and not only increased the discontent of citizens expressed by lower quality reports. As seen in these columns, the different Facebook treatment ads did indeed increase the higher quality reports on both the intensive and extensive margins. Moreover, the effects on the salience message group are now quite similar and not statistically different from the ones of the groups receiving the information message, which suggests that either the information or the salience message are equally effective in getting citizens to report “higher” quality reports despite the differences in the costs of reporting in both groups.

5.3 Effects on Electoral Irregularities

We now examine whether information about the reporting campaign can prevent candidates from engaging in electoral irregularities – or at least some types of them.

5.3.1 First Experiment - 2018 Presidential Elections

For the first experiment we use the reports collected by MOE about different types of electoral irregularities that are deemed of ‘high quality’, according to the evidence presented in the reports, as a proxy for electoral irregularities. There are two reasons to think this is a valid proxy for the actual occurrence of electoral irregularities. First, these reports are independent from the ones collected by the AG, which alleviates the concern that we might conflate an increase in reports due to the ad campaign with an increase in electoral irregularities. This last possibility would be a threat to our measurement strategy if citizens viewing the ad campaign decided to file reports

³⁶The effects using the inverse hyperbolic sine transformation can be interpreted in the same way as a logarithmic transformation (i.e. as elasticities), but this transformation is well defined for observations with zero reports.

³⁷We define these to be those reports of either a high or medium quality as judged by the MOE. See Section 3 for a discussion about how quality of reports is assessed by the MOE.

not only through the AG’s reporting website but also through MOE, but the results reported in this section will suggest that this is not the case. Second, the MOE classifies reports according to their quality, so using the highest quality reports, which contain the best evidence available about actual irregularities.

Table 5 displays the estimated effects of the different treatments on indicators of whether high quality reports about each type of electoral irregularities (described in Section 3) were issued from the municipalities in the sample.³⁸ Since we are now concerned with the occurrence of actual electoral irregularities, and not with the number of reports (several of which might be about the occurrence of a single electoral irregularity) we use an indicator which takes the value of one if there is any report coming from a given municipality about a given type of electoral irregularity.

Overall, the intervention seems to agree well with the prediction 2 from the model presented in Section 2, but the estimated effects are modest in magnitude. The letter sent to politicians reduced the overall occurrence of irregularities (column 1) by 0.1 standard deviations (p-value<0.1), as measured by an index including the main types of irregularities mentioned in Section 3.³⁹ As seen in the following columns, this effect was driven by the reduction in the probability of illicit political advertising by approximately 1 percentage point (p-value < 0.1), and the probability of vote buying by 1.5 percentage (p-value < 0.1). The letter sent to politicians did not significantly reduce the incidence of other types of irregularities. This is consistent with the predictions from model, which predicts that information about the ad campaign should reduce only observable electoral irregularities – of which illicit political advertising and vote buying are prime examples.

The ad treatment did not affect the probability of receiving reports about electoral irregularities of any type. This is consistent with the assumption that citizens viewing the ads did not file reports through MOE instead of the AG.

5.3.2 Second Experiment - 2019 Mayoral Elections

For the second experiment we use respondents’ self-reported occurrence of electoral irregularities as a measure of actual electoral misdeeds. We asked respondents to the post-treatment survey how likely each of the main types of irregularities presented in section 3 happened in the past elections in their municipality.⁴⁰ As our main outcome variable, we code indicators, d_{ki} for whether each type of irregularity k was deemed ‘likely’ or ‘very likely’ to have occurred by respondent i . We

³⁸Notice that two types of electoral irregularities – namely voter deception and fraud in voter registration – did not present any high quality reports by any municipality in our sample, and thus they are not included in this table.

³⁹This index was defined as

$$I_m = \sum_k z_{mk}$$

which is just a sum of the z-scores z_{mk} of the types of irregularities k in municipality m . The control mean and standard deviations were used to construct the z-scores.

⁴⁰We left out voter deception from the options in the survey due to the difficulty in explaining the cases that constitute voter deception to other forms of campaigning.

also computed an index across the different types of irregularities to estimate the effect of the intervention on electoral misdeeds, as a whole.⁴¹

Table 6 reports the results of the different treatment arms on these measures of electoral irregularities. Results in this table lead to very similar conclusions to the ones obtained in the first experiment. The letters sent to mayoral candidates reduced the incidence of electoral irregularities by 0.14 standard deviations (p-value<0.05), which is slightly higher than the effect found for the first experiment. This seems to be driven by a reduction in illicit political advertising and fraud in voter registration, both of which are become close to zero in places receiving the letter sent to politicians. As in the first experiment, these are types of irregularities that are conspicuous, as the model presented in Section 2 would predict.

As a whole, none of the Facebook advertisement messages seems to have altered the likelihood of electoral irregularities, as in the first experiment. Finally, as reported in last row of Table 6, the effect of the letters sent to politicians, net of the average effect of receiving any of the ad messages is lower than in the control group, which in indicates that the overall effect of sending these letters was to reduce electoral irregularities (p-value<0.1).

The results presented in this section confirm the findings from the first experiment concerning the effects of informing politicians about the advertisement campaign, but we find slightly larger effects. There are two potential reasons why these effects were larger in the second experiment. On the one hand, the municipalities included in the sample for the first experiment were relatively small in terms of their number of voters, which would make the information about our campaign not as relevant to presidential candidates, as for mayoral candidates. For instance, León and Azuero (2018) report that candidates and parties running in presidential elections only focus on the 200 largest municipalities known as the ‘Pareto municipalities’, which amass more than 70% of popular vote. Most of the sample for the first experiment is not part of this group of municipalities and, as such, the reporting campaign might have not been problematic for candidates which probably did not have them under their radar. On the other hand, the letters sent to presidential candidates in the first experiment were only sent 11 days prior to the elections, which might have been too short to generate a sizable reaction from candidates in their campaigning strategies. In contrast, in the second experiment we managed to send these letters almost two months before the elections, which gave a larger margin to politicians to alter their campaigning strategies.

⁴¹This index was defined analogously to the one used in the first experiment. More precisely, the index had the following form:

$$I_i = \sum_k z_{ki}$$

which is a sum of the z-scores z_{ik} of indicators d_{ki} for each of the types of irregularities k according to respondent i . The control mean and standard deviations were used to construct the z-scores.

5.4 Effects on Legal Campaigning

In this section we examine whether candidates reacted to information about the advertisement campaign by shifting their efforts towards legal types of campaigning. As highlighted in Proposition 2, given enough substitutability between the different campaigning strategies used by politicians, the advertisement campaign might not only generate a decrease in politicians’ engagement in observable electoral irregularities, but also a substitution towards legal forms of campaigning.

In the post-treatment survey collected around the second experiment in 2019 we asked respondents to tell us whether they had witnessed different types of campaigning – namely, public speeches, fliers, social media ads and radio ads – for each candidate running for Mayor in their municipality in week before the elections. We thus test for effects on legal campaigning by running regressions of the following type:

$$y_{imc} = \alpha_{TI}Info_m + \alpha_{TS}Saliencem + \alpha_{TB}Info\&Saliencem + \beta_2Letter_m + X'_{im}\gamma_2 + \phi_m^2 + \epsilon_m^2$$

where y_{imc} is an indicator for whether respondent i in municipality m witnessed one of the mentioned forms of legal campaigning done by candidate c , and the rest is defined as in equation (4), except that we use individual level covariates X_{im} , as in the other specifications using survey data.

Table A.11 displays the estimated effects of the different interventions performed in the second experiment in 2019 on the measures of legal campaigning collected in the post-treatment survey. Results in this table show that we cannot reject the null hypothesis that the interventions had no effect on any of the forms of campaigning collected in the post-treatment survey. In fact, we find a negative (but insignificant) effect of informing politicians about the advertisement campaign ahead of the elections on an index of the different forms of campaigning (column 1), which suggests that, if anything, legal campaigning might be a *complement* instead of a *substitute* to observable irregularities.

5.5 Effects on Voting Behavior

In this section we describe the estimated effects of the intervention on voting outcomes, and we delay the interpretation of the results for the next section.

5.5.1 First Experiment - 2018 Presidential Elections

Table 7 reports the effect of the first intervention carried out in 2018 on several voting outcomes in the first round of the presidential elections. Neither the ad campaign nor the letters sent to politicians had any effect on turnout or on blank votes, which are generally used as ‘protest votes’ in this context. The ad campaign, however did have a significant and meaningful effect on voting: municipalities receiving any of the ad treatments saw on average an increase of 1.8% in the

votes for non-traditional candidates and an equivalent reduction in the votes share for traditional candidates (p-value < 0.05). We cannot reject the null hypothesis that the letter sent to politicians did not have an effect on any of the voting outcomes, which suggests that the reduction in illicit campaigning reported in the last section was not strong enough to generate a shift in the vote shares for candidates.

These findings are robust to the use of the voting outcomes gathered in the survey data, as shown in Table A.9. The magnitude of the effect on vote for non-traditional (and traditional) candidates is larger using this data: the pooled treatment effects indicate a 5.9 percentage point increase (p-value < 0.1) in the votes for non traditional candidates and a 7 percentage point decrease in the vote share for traditional ones (p-value < 0.05). The difference in the estimates from the administrative data and the survey data is almost completely explained by the fact that survey respondents are part of the Facebook population targeted by the ads: using the fact that 36% of registered voters (see Table 2) were reached by the ad, we compute an estimate of the effect of the treatment on the treated of $1.8/0.36 = 5$ percentage points, which is almost identical to the one found using data from the survey respondents.⁴²

For reference, we discuss the effects of the intervention on the vote shares for all of the candidates’ vote shares separately in Section A.5 of the appendix.

Table 8 explores whether the intervention had longer-term effects on voting behavior by looking at the second round of elections that took place three weeks after the intervention and the first round of elections. The effects on the vote share for the traditional and non-traditional candidates go in the same direction as before, although they are smaller and lose some of their statistical significance when controls are included: municipalities receiving any of the ads exhibited a decrease in the vote share for Duque of 1.2% (p-value ≈ 0.12) and a increase in the vote share for Petro of about 1.1% (p-value ≈ 0.17).

5.5.2 Second Experiment - 2019 Mayoral Elections

We now turn to examine the effects of the second experiment interventions on the 2019 mayoral elections. Due to the large number of candidates involved in these elections, we now use a data-driven way of identifying which candidates are perceived to be ‘cleaner’ by citizens using the responses collected in the pre-treatment survey. In particular, we define candidates to perceived to be ‘clean’ by citizens as those who are mentioned above the median as as being the best ones to either fight corruption or vote buying. More precisely, for each candidate c in municipality m , we define this candidate to be perceived as ‘clean’ if $\sum_i b_{cmi} \geq median_m$, where b_{cmi} is as indicator for whether candidate c is named as the best candidate to fight corruption or vote buying by respondent i , and $median_m$ is the median number of times that the candidates are perceived to be clean by respondents in municipality m .

⁴²A second mayor difference is that in the survey data there seems to be a significant negative effect of the ads on respondents voting blank. The average effect of the pooled ads was to decrease the blank votes by 1.2 percentage points (p-value < 0.1).

Table 9 displays the results of the different interventions involved in the second experiment on different electoral variables. Results in columns (1)-(2) show that none of the interventions had a statistically significant effect on turnout, as was the case in the first experiment. As seen in columns (3)-(6) ads containing the salience message increased the vote share for candidates who perceived to be clean and decreased the vote share for candidates perceived to be less clean by almost exactly the same amount. In particular, the advertisement containing only the salience message increased the vote share of candidates perceived to be ‘clean’ by almost 4% (p-value <0.05) – although this result loses significance when including the indicator for the letter treatment, which is possibly due to a loss of power in this latter specification – while the ad featuring both the salience and the information messages had a larger effect of about 6% (p-value <0.01). These effects are, however not statistically distinguishable. On the other hand, the effect of the ads only containing the information message follows the same pattern but is not statistically different from zero.⁴³ Finally, the letter sent to mayoral candidates had no effect on these variables, suggesting that – as in the first experiment – the decrease in electoral irregularities in this group were not sufficiently large to generate a significant change in the voting patterns.

Columns (7)-(8) show that the different ad variations had the effect of increasing blank votes by approximately 0.2%-0.5% (p-value <0.1), depending on the treatment group, but these effects are not statistically different from one another, and they become insignificant when including the indicator for the letter sent to politicians. As in the other variables, the letter sent to politicians did not have any significant effect over the share of blank votes.

5.6 Interpretation(s) of the Effects on Voting Behavior

As mentioned in the introduction, the interventions might have affected voting in favor of more ‘clean’ candidates by one of two channels: (1) by affecting the extent of electoral irregularities it might have altered electoral results directly – for instance, it might have reduced vote buying for certain candidates which would reduce their vote shares; (2) it might have altered the salience of electoral irregularities, which might have indirectly affected citizen’s willingness to vote for particular candidates, as highlighted in the model presented in Section 2. In this section we will argue that the effects on voting behavior described in the last section were due to the latter channel. Four pieces of evidence suggest that this is the case.

First, as shown in Section 5.3, in neither of the experiments did the different ad treatments have an effect on the prevalence of electoral irregularities. This suggests that the change in vote shares caused by the ad campaigns is unlikely to have been caused by a change in candidates electoral irregularity strategies.

Second, as we will show in the following section, in the first experiment the ad campaign generated an increase in the number of protests against that national government lead by Duque, the

⁴³Moreover, the effect of the the only containing the information message is statistically different from the one with both the information and the salience messages (p-value <0.05), as seen in the lower rows of this table.

traditional candidate that won the 2018 presidential elections, in the first six months of his mandate. It seems unlikely that a change in electoral irregularities might have generated this increase in protests. Instead, this seems to square up better with explanation (2), which suggests that the ad might have shifted preferences against traditional candidates, and that this was reflected not only in the voting behavior of citizens in the presidential elections, but also in the protests that occurred later on when Duque got to power.

Third, it also seems unlikely that candidates and their parties had time to react so quickly to the ad campaigns – which, as a reminder, started four days before the elections – as to generate such a large movement in votes as the one reported in the last section. As discussed in Section 4.2 a large proportion of electoral irregularities have occurred before election day, and even before six days prior to election day, in previous elections. To some extent this reflects the logistical and organizational issues that electoral irregularities imply for parties. As mentioned in Section 3, for instance, vote buying often begins three months before elections when voters are reassigned to polling stations so that brokers can monitor whether they are voting in the way are paid to.

Finally, the results from the second experiment suggest that only ads containing the salience message which encouraged citizens to report and to fight against electoral irregularities generated a increase in the vote share for ‘clean’ candidates and against those perceived to be less ‘clean’. This suggests that the specific content of the ads generated differential responses by voters, which is easier to conciliate with explanation (2).

Even though these pieces of evidence suggest that the ad interventions changed voters’ behavior by directly affecting their preferences about ‘cleaner’ candidates and not by altering the extent to which parties and candidates engaged in electoral irregularities, it is important to point out that the exact mechanism through which it changed their preferences is unknown. We explore this issue in the following sections.

5.7 Discussion about the Magnitude of the Effects on Voting Outcomes

How large were the effects on voting behavior reported in the previous sections? One convenient way to benchmark the results found is using what DellaVigna and Gentzkow (2010) call the *persuasion rate*. This statistic measures how powerful a message is at swaying votes in a particular way, and it has been computed in many settings, which might serve as a comparisons. If we denote a treatment group as T and the control group by C , the persuasion rate is defined by the following formula:

$$f = \frac{v_T - v_C}{e_T - e_C} \frac{t_T}{100 - v_C} \times 100$$

where e_i represents the share of group $i = T, C$ receiving the message, v_i is the vote share and t_i is turnout. Intuitively, the first term captures the change in vote shares on the population receiving

a message and the second provides a correction by the fact that larger turnout and/or larger vote shares in the control group will imply larger persuasion rates.

Using the estimates from Section 5 ⁴⁴ we estimate a persuasion rate of $f \approx 6$ away from traditional candidates in the first experiment, and a persuasion rate of $f \approx 15$ away from the candidates perceived to be less ‘clean’ when considering the ad featuring both the salience and information message in the second experiment. ⁴⁵ This implies that approximately 1 in every 17 people that saw the ad in the first experiment changed their vote from the traditional candidates perceived to be less clean, and that 1 in every 7 that saw the ad featuring both the salience and the information message changed their vote from the candidates perceived to be less ‘clean’ to the cleaner ones. ⁴⁶

In comparison, DellaVigna and Kaplan (2007) estimate $f = 11.6$ for the effect of Fox News entry on Republican vote share; Enikolopov, Petrova, et al. (2011) find $f = 7.7$ for the effect of an independent Russian TV station on anti-Putin vote share; Enríquez et al. (2019) estimate large values of f between 39 and 84 for the effect of an information campaign disclosing low government expenditure irregularities on incumbent candidate vote share; finally, and closer to the topic at hand, Green and Vasudevan (2016) estimate f between 7 and 24 for a radio campaign encouraging citizens to vote against candidates engaged in vote-buying in India. Given that our Facebook ad intervention was not directly designed to encourage vote for any particular candidate as some of these interventions, this estimate reveals that our campaign had a particularly large persuasive effect in both experiments.

However, it is important to notice that the estimation of persuasion rates relies on the assumption that only citizens that viewed the ads were persuaded by them. If instead, there are spillover or social multipliers due to citizens changing other peers’ voting behavior this persuasion rates would be lower. Thus, the persuasion rates estimated are probably upper bounds on the effects of the interventions over the whole population, as has been documented by Enríquez et al. (2019).

5.8 Effects on Protests Against the Government

As shown previously, the ad campaign in the first experiment had the effect of shifting citizen’s votes from traditional candidates towards non-traditional ones which were perceived to be ‘cleaner’. If this shift of voting outcomes was driven by a change in the underlying preferences for ‘clean’ candidates, it is possible that citizens expressed these preferences in other political arenas different from elections. In particular, participation in the series of protests that occurred against Duque’s

⁴⁴From Table 7 we have $v_T - v_C \approx -1.87$, $t_T \approx 49.2$ and $100 - v_C \approx 40$. From Table ??, we have $e_T - e_C \approx 36$, under the reasonable assumption that citizens in the control group did not view the ad.

⁴⁵From Table 9 we have $v_T - v_C \approx -5.4$, $t_T \approx 67$ and $100 - v_C \approx 79.5$. Finally, from Table 2, we have $e_T - e_C \approx 31$.

⁴⁶Alternatively, the persuasion rate in favor of ‘clean’ candidates is $f = 4$ in the first experiment and $f = 57$ in the second experiment. As explained below the large persuasion rates might be explained by spillovers or social multipliers generated by the experiment.

government in the aftermath of his victory might be used to test this change in preferences against traditional candidates.

Table 10 reports the effects of the interventions on indicators of whether there were protests against the national government (columns 1-2) or other levels of government (columns 3-4) in the period after the second round of elections and until the end of 2018.⁴⁷ The ad campaign had a large and statistically significant effect on the probability of protests against the national government: it increased the probability of these type of protests occurring by over 3 percentage points (p-value < 0.05), more than double the control group mean. Moreover, the letter sent to politicians sent to politicians did not have a significant effect on the probability of protests against the national government.

However, as seen in columns 3-4, the different ads did not have any effect on the probability of protests occurring against other, sub-national, forms of government. This suggests that the way that the campaign affected protests was through a change of preferences against traditional politicians and not merely through an increase in discontent or a general increase in the ability of municipalities to organize collective action.

5.9 Survey Outcomes and Mechanisms

In this section we conclude by exploring some potential mechanisms, using the outcomes measured in the post-treatment surveys conducted after the interventions in both of the experiments.

5.9.1 First Experiment - 2018 Presidential Elections

We begin by considering whether the interventions in the first experiment affected trust in institutions such as elections, the president, the judiciary and the AG. The ad treatments could have affected trust in these institutions in both a positive and a negative way. On the one hand, they could have decreased trust by increasing the salience of electoral irregularities. On the other hand, they could have increased it by providing information about the existence of a reporting site to fight electoral irregularities, or by giving citizens the sense that they had a tool at hand to counter electoral irregularities. As such, it is difficult to predict ex-ante whether the intervention would increase or decrease trust in these institutions.⁴⁸

Table 11 reports the effects of the different interventions on measures of trust of the institutions previously mentioned on a scale from 1 to 7, with higher scores representing higher trust in each institution. We standardize these variables so they are easier to interpret. Overall, the average effect of the ad treatments led to a decrease in trust all of the institutions considered, but this effect is only statistically significant for the measure of trust in elections.

⁴⁷This is the longest period for which the CINEP has records of protests at the moment.

⁴⁸Additionally, interpretation of these possible effects might be complicated due to the fact that the post-treatment surveys were conducted after the elections, so the results from the election might also be a mediating variable in explaining any potential shift in institutional trust.

The Facebook ad treatment caused a decrease of 0.16 standard deviations in trust in elections (p-value < 0.05). The decrease in trust about elections caused by the ad campaign provides supporting evidence to our interpretation that the ads increased the salience of electoral irregularities in a way that caused pessimism about the transparency of elections. A contending interpretation of this result is, however, that the technical issues surrounding the AG’s reporting website might have generated distrust about elections. However, this interpretation seems to be contradicted by the fact that trust in the AG was not significantly affected by the intervention (see Table 11, columns 7-8).

The letter sent to candidates had the effect of increasing the trust in elections by 0.12 standard deviations (p-value < 0.1). This might be explained by an increase in trust of elections due to the reduction in conspicuous electoral irregularities in municipalities included in the letter to candidates, as reported in previous sections.

A second set of outcomes of interest answer the question of whether the ad treatments changed the perception of citizens about how easy or effective reporting is. In particular, the Effectiveness ad treatment was designed to increase citizens’ perception about how effective reporting is, so this provides a direct test of whether this objective was successfully achieved. One caveat to keep in mind, however, is that the fact that the AG’s reporting website was not cellphone compatible – as explained earlier – might have created the opposite effect of decreasing citizen’s perceptions of reporting ease and effectiveness. Columns 1-4 of Table A.12 report the estimated effects of the interventions on measures of perceived effectiveness and easiness of reporting on a scale from 1 to 7, with higher values representing perceptions that reports are more effective or easier. As before, we standardized these variables for ease of comparison. Results show that none of the interventions had a statistically significant effect on the effectiveness or easiness of reporting. If anything, the estimates suggest that the ad treatments had a *negative* effect on these perceptions, consistent with the observation that the issues with the AG’s reporting website might have had a perverse effect on citizens’ views on reporting.

5.9.2 Second Experiment - 2019 Mayoral Elections

As with the first experiment, we begin by analyzing the effects of the interventions of the second experiment on trust in institutions. We measured trust in institutions by asking respondents how much they trusted different types of institutions, and we coded indicator variables for whether they “trusted” or “trusted a lot” each institution. Table 12 reports the effects of the different treatment arms on these outcomes. Results in this table show that none of the interventions had any significant effect on the trust on any of the institutions considered. As mentioned in the discussion for the first experiment, this could have happened because ex-ante the interventions could have had positive or negative effects on trust, even in the presence of salience effects that lead citizens to vote for ‘cleaner’ candidates.

In Table 13 we further explore whether the different interventions affect citizens’ perceptions of how important corruption is as a problem in the whole country or in their municipalities. Results

in columns (1)-(2) reveal that the importance of corruption was increased particularly in municipalities that received the Facebook ad message featuring both the salience and the information message by 0.13 standard deviations (p-value < 0.1), using an index of corruption importance as an outcome, but this effect is not significant for any of the other ad treatments. Results in columns (3)-(6) reveal that this effect is driven by an increase in the perception of how important corruption is a problem in the country, but not in the particular municipalities of respondents receiving the salience and the information messages combined. This finding is consistent with the fact that this particular treatment group was the one which experienced the largest shift in the vote share for ‘clean’ candidates, so that the salience effects seem to have been larger for this group. One hypothesis of why this might have been the case is that in this group the salience message was further intensified by the belief that the were official websites put in place at the national level to counter the issue of corruption.

6 Conclusions and Final Remarks

This paper provides evidence that campaigns to promote bottom-up monitoring of elections not only generate informational effects, but also give rise to large salience effects, which operate by making citizens more aware about the possibility that elections might be rigged or, at least, tarnished by electoral irregularities. In our specific context, these salience effects acted as complements to the general objective of reducing corruption in elections since citizens reacted to our intervention by voting for candidates they perceived to be less engaged in electoral malpractice. However, these salience effects could have also generated a backlash by making citizens pessimistic about the transparency and worth of elections, as has been documented in other settings – i.e. Chong et al. (2015) . Understanding which factors determine whether the salience of corruption spurs or depresses citizens’ engagement in democracy is an important avenue of future research.

The implications for policy are numerous. Many governments and NGOs have headed the call made in the *World Development Report* of 2004 (World Bank, 2004) to adopt and promote the use of bottom-up mechanisms, with the goal of increasing the oversight of the provision of public goods and services. The results in this paper suggest that an assessment of the success or failure of these initiatives must incorporate both the informational effects of these campaigns, as well as the behavioral changes that are triggered by the alterations in the salience of the issues being monitored through these interventions. In our context, the cost-effectiveness estimates of the intervention should not only incorporate the reduction of electoral irregularities caused by the intervention, but also the changes in citizens attitudes towards democratic institutions and their support for candidates they perceive to be ‘cleaner’. A back-of-the-envelope calculation suggests that our intervention shifted the vote share towards candidates that were perceived to be ‘cleaner’ at a cost of approximately \$0.03-0.17 USD per vote.⁴⁹

⁴⁹On the upper bound, in the first experiment, a 1.8% shift in vote share is equivalent to 81 votes, and the average cost of advertisements in treated municipalities was \$13.7 USD per municipality, which implies a cost of

Throughout the paper we have tried to stressed the fact that the advertisement campaign increased the vote share of candidates *perceived* to be cleaner, but not necessarily the *actual* candidates that are cleaner. Understanding whether these two effects are aligned or not, and whether campaigns as the ones presented in this paper can be combined fruitfully with information campaigns about the candidates' actual engagement in corruption, is an important question for future research. Moreover, the results from our second experiment suggest that the salience effects of these types of campaigns can be in fact modulated and tailored for specific contexts. For instance, if citizens in fact have poor knowledge about which candidates are more or less 'clean', it might be desirable to avoid salience effects which might lead them to change their vote in favor of candidates that are not in fact as transparent as citizens believe them to be. On the other hand, in contexts in which citizens are well informed about which candidates have a better history of being 'clean', it might be desirable to enable the salience effects.

The results in this paper also suggest that using social media to promote bottom-up engagement in electoral monitoring is a cost-effective way of curbing electoral irregularities. Callen, Gibson, et al. (2016) estimate that the European Union spends about \$6000-20000 USD per polling station deploying electoral observers in missions in regions with weak state capacity, with no concluding evidence about their effectiveness in reducing electoral irregularities. In comparison, our intervention is substantially cheaper and we have provided evidence from two different experiments that it indeed reduces electoral irregularities by a meaningful amount. More concretely, we estimate that our intervention cost between \$13-19 USD per municipality – or \$1.5-2 USD per polling station – and reduced the extent of electoral irregularities by 0.1-0.15 standard deviations across the two experiments. Moreover, this intervention is easy to scale-up, requiring only that the targeted areas possess substantial internet coverage and use of social media – both of which are still low in the developing world, but quickly increasing (World Bank, 2016).

\$0.17 per vote. On the lower bound, in the second experiment, a 5% shift in vote share is equivalent to 693 votes and the average cost of advertisements per municipality was \$19 USD, which implies a cost of \$0.03 per vote.

References

- Acemoglu, Daron, Tarek A. Hassan, and Ahmed Tahoun (2018). “The Power of the Street: Evidence from Egypt’s Arab Spring”. *Review of Financial Studies* 31(1), pp. 1–42.
- Acemoglu, Daron, James Robinson, and Rafael J. Santos-Villagran (2013). “The Monopoly of Violence: Evidence from Colombia”. *Journal of the European Economic Association* 11, pp. 5–44.
- 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), pp. 185–200.
- Ardila, Laura (2018). “Profesión: Puya ojos”. In: *El dulce poder: Así funciona la política en Colombia*. Penguin Random House, pp. 43–49.
- Arenas, Natalia (2018). “El primer eslabón: La compra de los ediles”. In: *El dulce poder: Así funciona la política en Colombia*. Penguin Random House, pp. 51–58.
- Asunka, Joseph, Sarah Brierley, Miriam Golden, Eric Kramon, and George Oforu (2019). “Electoral Fraud or Violence: The Effect of Observers on Party Manipulation Strategies”. *British Journal of Political Science* 49(1), pp. 129–151.
- Athey, Susan and Guido Imbens (2017). “The Econometrics of Randomized Experiments”. In: *Handbook of Economic Field Experiments*. Ed. by Abhijit Banerjee and Esther Duflo. Vol. 1. North-Holland, pp. 73–140.
- 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*. Ed. by Frederic Charles and Schaffer Andreas. Lynne Rienner Publishers.
- Barberà, Salvador and Matthew O. Jackson (2019). “A Model of Protests, Revolution, and Information”. SSRN Working paper. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2732864.
- 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), pp. 608–50.
- Blair, Graeme, Rebecca Littman, and Elizabeth Levy Paluck (2019). “Motivating the Adoption of New Community-Minded Behaviors: An Empirical Test in Nigeria”. *Science Advances* 5, pp. 1–8.
- 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.

- Bond, Robert M., Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler (2012). “A 61-million-person experiment in social influence and political mobilization”. *Nature* 489, pp. 295–298.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2013). “Salience and Consumer Choice”. *Journal of Political Economy* 121(5), pp. 803–843.
- Boyle, Elizabeth Heger, Cosette D. Creamer, Amy Hill Cosimini, Yagmur Karakaya, Suzy McElrath, Florencia Montal, and J. Siguru Wahutu (2017). “Making Human Rights Campaigns Effective While Limiting Unintended Consequences: Lessons from Recent Research”. USAID Research and Innovation Grants Working Papers Series.
- Broockman, David and Donald P. Green (2014). “Do Online Advertisements Increase Political Candidates’ Name Recognition or Favorability? Evidence from Randomized Field Experiments”. *Political Behavior* 36(2), pp. 263–289.
- 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, pp. 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), pp. 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), pp. 486–490.
- Cho, Hyunyi and Charles T. Salmon (2007). “Unintended Effects of Health Communication Campaigns”. *Journal of Communication* 57(1), pp. 293–317.
- Chong, Alberto, Ana De La O, Dean Karlan, and Leonard Wantchekon (2015). “Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification”. *Journal of Politics* 77(1), pp. 55–71.
- Collier, Paul and Pedro C. Vicente (2012). “Violence, bribery, and fraud: the political economy of elections in Sub-Saharan Africa”. *Public Choice* 153, pp. 117–147.
- Collier, Paul and Pedro C. Vicente (2013). “Votes and violence: Evidence from a field experiment in Nigeria”. *The Economic Journal* 124, pp. 327–355.
- DellaVigna, Stefano and Matthew Gentzkow (2010). “Persuasion: Empirical Evidence”. *Annual Review of Economics* 2(1), pp. 643–669.
- DellaVigna, Stefano and Ethan Kaplan (2007). “The Fox News Effect: Media Bias and Voting”. *The Quarterly Journal of Economics* 3(122), pp. 1187–1234.

- Dennison, James (2019). “A Review of Public Issue Salience: Concepts, Determinants and Effects on Voting”. *Political Studies Review*, pp. 1–11.
- Driscoll, Jesse and Daniel Hidalgo (2014). “Intended and Unintended Consequences of Democracy Promotion Assistance to Georgia After the Rose Revolution”. *Research and Politics*, pp. 1–13.
- Duflo, Esther, Rachel Glennerster, and Michael Kremer (2008). “Using Randomization in Development Economics Research: A Toolkit”. In: *Handbook of Development Economics*. Ed. by T. Schultz and John Strauss. Vol. 4. Elsevier, pp. 3895–3962.
- El Tiempo (May 2018). *Corrupción, Fuerzas Militares y gobernabilidad, los otros temas*. "<https://www.eltiempo.com/elecciones-colombia-2018/presidenciales/corrupcion-fuerzas-militares-y-gobernabilidad-temas-de-el-debate-222044>". Accessed: 2018-04-03.
- 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), pp. 448–452.
- Enikolopov, Ruben, Alexey Makarin, and Maria Petrova (2018). “Social Media and Protest Participation: Evidence from Russia”. SSRN Working paper, available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2696236.
- Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya (2011). “Media and Political Persuasion: Evidence from Russia”. *American Economic Review* 7(101), pp. 3253–3285.
- Enríquez, José Ramón, Horacio Larreguy, John Marshall, and Alberto Simpser (2019). “Online Political Information: Facebook Ad Saturation and Electoral Accountability in Mexico”. Working paper.
- Facebook (2019). *Ad Help Documentation*. <https://www.facebook.com/business/help/355670007911605>. Accessed: 2019-07-19.
- Fergusson, Leopoldo and Carlos Molina (2019). “Facebook Causes Protests”. Working Paper.
- 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* (20), pp. 1–13.
- Finan, Frederico and Laura Schechter (2012). “Vote-Buying and Reciprocity”. *Econometrica* 80(2), pp. 863–881.
- Fox, Johnathan A. (2015). “Social Accountability: What Does the Evidence Really Say?” *World Development* 72, pp. 346–361.
- Gans-Morse, Jordan, Sebastián Mazzuca, and Simeon Nichter (2014). “Varieties of Clientelism: Machine Politics during Elections”. *American Journal of Political Science* 2(58), pp. 415–432.
- García-Jimeno, Camilo, Angel Iglesias, and Pinar Yildirim (2018). “Women, Rails, and Telegraphs: An Empirical Study of Information Diffusion and Collective Action”. Working paper.

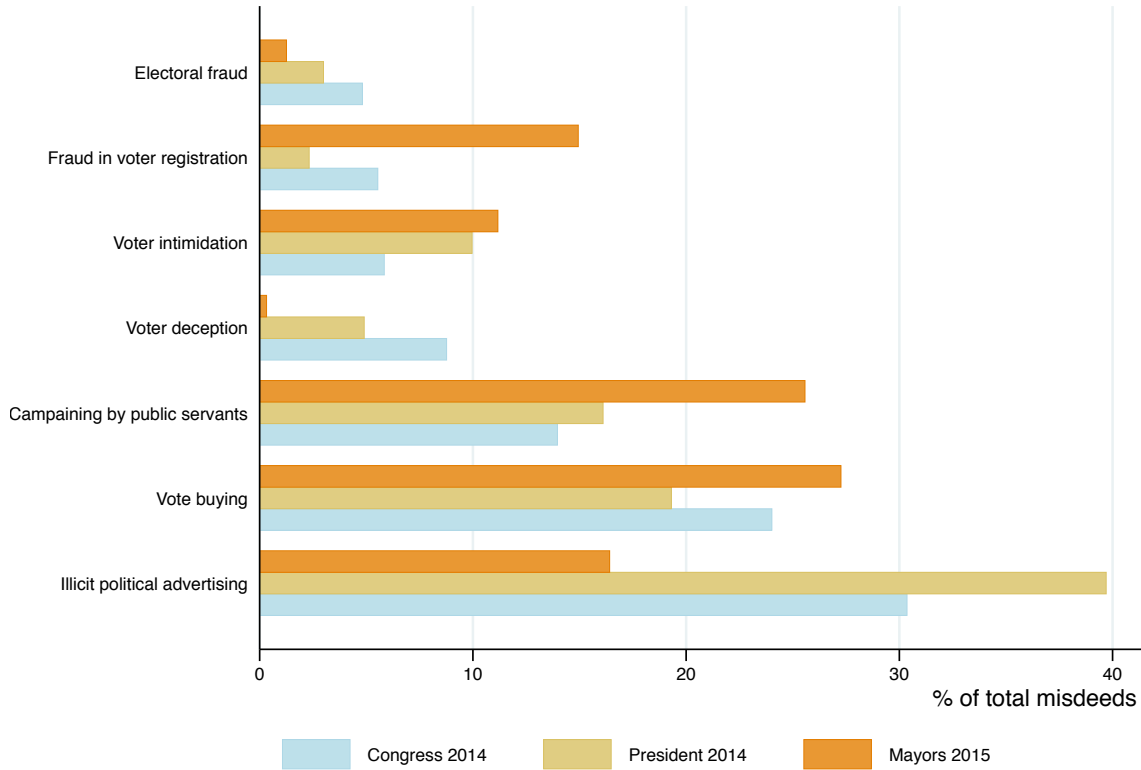
- Garcia, Miguel and Sebastian Pantoja (2015). “Incidencia del clientelismo segun riesgo electoral y de violencia: Un análisis de las elecciones presidenciales de 2014 en municipios de consolidación territorial”. In: *Mapas y factores de riesgo electoral. Elecciones de autoridades locales*. Mision de Observación Electoral, pp. 291–313.
- Glennerster, Rachel and Kudzai Takavarasha (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton and Oxford: Princeton University Press.
- 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), pp. 202–217.
- Gonzalez, Robert (2019). “Cell Phone Access and Election Fraud: Evidence from a Spatial Regression Discontinuity Design in Afghanistan”. Working Paper.
- Green, Donald P. and Srinivasan Vasudevan (2016). “Diminishing the Effectiveness of Vote Buying: Experimental Evidence from a Persuasive Radio Campaign in India”. Working paper.
- Haenschen, Katherine and Jay Jennings (2019). “Mobilizing Millennial Voters with Targeted Internet Advertisements: A Field Experiment”. *Political Communication* 36 (3), pp. 376–393.
- Hager, Anselm (2019). “Do Online Ads Influence Vote Choice?” *Political Communication* 36, pp. 376–393.
- Hicken, Allen (2011). “Clientelism”. *Annual Review of Political Science* 14, pp. 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, pp. 1–14.
- Hyde, Susan D. (2007). “The Observer Effect in International Politics: Evidence from a Natural Experiment”. *World Politics* 60(1), pp. 37–63.
- Hyde, Susan D. (2010). “Experimenting in Democracy Promotion: International Observers and the 2004 Presidential Elections in Indonesia”. *Perspectives on Politics* 8(2), pp. 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), pp. 292–307.
- Khemani, Stuti (2015). “Buying Votes Versus Supplying Public Services: Political Incentives to Under-Invest in Pro-Poor Policies”. *Journal of Development Economics* 117, pp. 84–93.
- Klašnja, Marko, Joshua A. Tucker, and Kevin Deegan-Krause (2014). “Pocketbook vs. Sociotropic Corruption Voting”. *British Journal of Political Science*, pp. 1–28.
- Kobayashi, Tetsuro and Yu Ichifuji (2015). “Tweets That Matter: Evidence From a Randomized Field Experiment in Japan”. *Political Communication* 32, pp. 574–593.

- Kosinski, Michal, Sandra C. Matz, Samuel D. Gosling, Vesselin Popov, and David Stillwell (2015). “Facebook as a Research Tool for the Social Sciences: Opportunities, Challenges, Ethical Considerations, and Practical Guidelines”. *American Psychologist* 70(6), pp. 543–556.
- 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, pp. 42–58.
- León, Juanita and Manolo Azuero (2018). “El Día D”. In: *El dulce poder: Así funciona la política en Colombia*. Penguin Random House, pp. 87–93.
- Little, Andrew T. (2016). “Communication Technology and Protest”. *Journal of Politics* 78(1), pp. 152–166.
- Marx, Benjamin, Vincent Pons, and Tavneet Suri (2017). “The Perils of Voter Mobilization”. NBER Working paper N.23946.
- Misión de Observación Electoral (2018). *Irregularidades electorales en Colombia: Informe final Pilas con el voto, elecciones presidencia y congreso 2018*. Mision de Observación Electoral.
- Nichter, Simeon (2008). “Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot”. *American Political Science Review* 102(1), pp. 19–31.
- Olken, Benjamin (2007). “Monitoring Corruption: Evidence from a Field Experiment in Indonesia”. *Journal of Political Economy* 115, pp. 200–249.
- Olken, Benjamin and Rohini Pande (2012). “Corruption in Developing Countries”. *Annual Review of Economics* 4, pp. 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*, pp. 1–26.
- Reinikka, Ritva and Jakob Svensson (2004). “The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture”. *Quarterly Journal of Economics* 119, pp. 678–704.
- Reinikka, Ritva and Jakob Svensson (2011). “The Power of Information in Public Services: Evidence from Education in Uganda”. *Journal of Public Economics* 95, pp. 956–966.
- Robinson, James A. and Ragnar Torvik (2014). “The Real Swing Voter’s Curse”. *American Economic Review: Papers and Proceedings* 99(2), pp. 310–315.
- Rueda, Miguel R. (2017). “Small Aggregates, Big Manipulation: Vote Buying Enforcement and Collective Monitoring”. *American Journal of Political Science* 61(1), pp. 163–177.
- Ryvkin, Dmitry, Danila Serra, and James Tremewan (2017). “I Paid a Bribe: An Experiment on Information Sharing and Extortory Corruption”. *European Economic Review Volume* 94, pp. 1–22.

- Samuels, David and Cesar Zucco (2013). “Using Facebook as a Subject Recruitment Tool for Survey-Experimental Research”. SSRN Working paper.
- Sances, Michael W. (2019). “Missing the Target? Using Surveys to Validate Social Media Ad Targeting”. *Political Science Research and Methods*, pp. 1–8.
- Schedler, Andreas (2002). “Elections Without Democracy: The Menu of Manipulation”. *Journal of Democracy* 13(2), pp. 36–50.
- Sergio Fajardo’s Campaign Team (2018). *Resumen ejecutivo plan de gobierno: Campaña presidencial de Sergio Fajardo*. <http://sergiofajardo.co/wp-content/uploads/2018/05/plan-completo.compressed-1.pdf>. Accessed: 2019-04-01.
- 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), pp. 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), pp. 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), pp. 292–305.
- White, Halbert (1980). “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”. *Econometrica* 48(4), pp. 817–838.
- World Bank (2004). *World Development Report: Making Services Work for the Poor*. Washington, DC: World Bank.
- World Bank (2016). *World Development Report: Digital Dividends*. Washington, DC: World Bank.
- World Bank (2017). *World Development Report: Governance and the Law*. Washington, DC: World Bank.
- Young, Alwyn (2017). “Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results”. London School of Economics.
- Zhang, Baobao, Matto Mildemberger, Peter D. Howe, Jennifer Marlon, Seth A. Rosenthal, and Anthony Leiserowitz (2018). “Quota Sampling Using Facebook Advertisements”. *Political Science Research and Methods*, pp. 1–7.

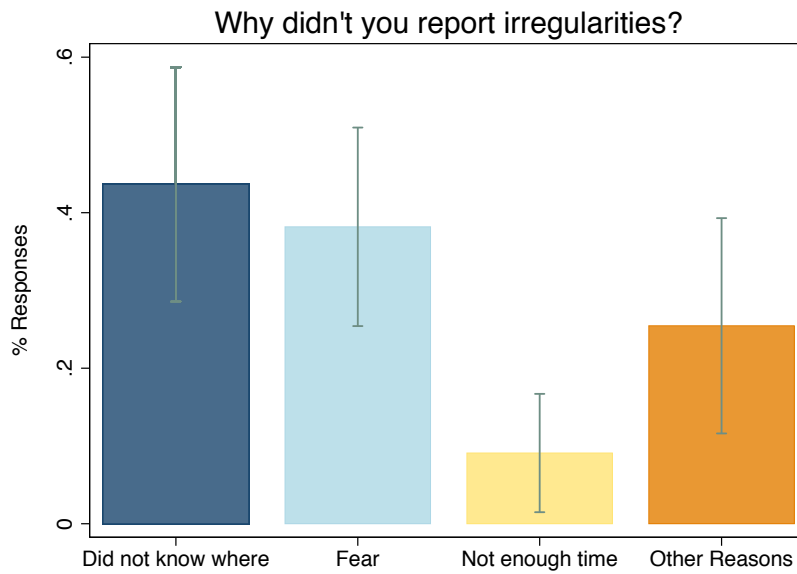
7 Figures

Figure 1: Electoral irregularities Reported in 2014-2015 Elections



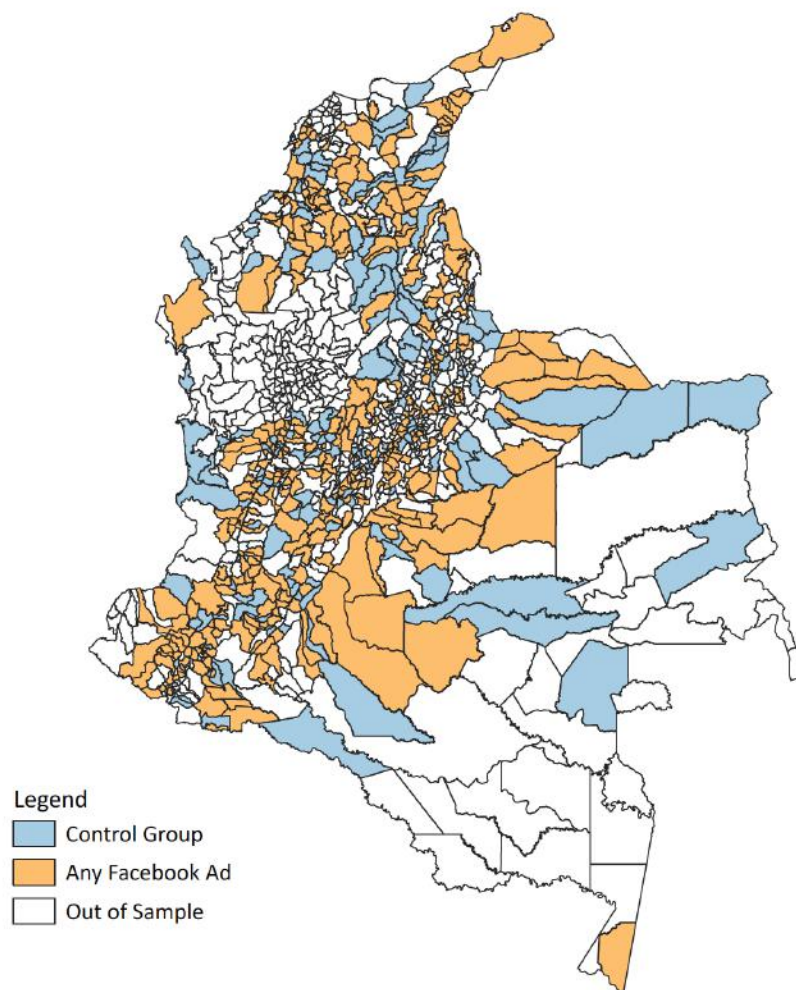
Notes: This figure displays the proportion of electoral irregularities of different types as a percentage of total irregularities reported to URIEL for three different elections: the 2014 congressional and presidential elections and the 2015 mayoral elections. The definitions for each type of electoral irregularity are presented in Section 3.

Figure 2: Survey Responses: Reasons not to Report



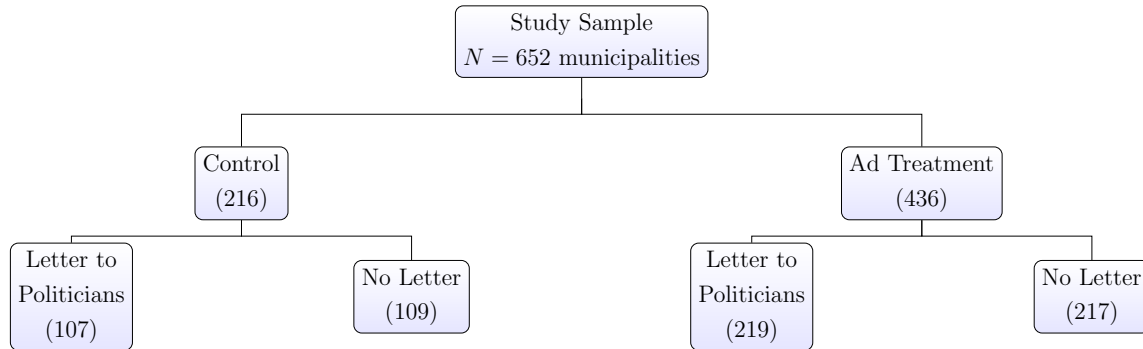
Notes: This figure plots the responses to the question “Why didn’t you report the electoral irregularities you witnessed?” among survey participants in the first experiment who said they had witnessed an electoral irregularity but did not report it. Details about the survey performed in the first experiment are found in Section 4.3.1. The four possible answers shown in the figure were not exclusive and respondents could pick more than one option. The sample is restricted to only respondents coming from control municipalities so that responses are not affected by the different treatment conditions. The total number of control group respondents who report witnessing electoral irregularities but did not report them is 55. 95% confidence intervals for the average response to each mentioned reason are shown in gray.

Figure 3: Municipalities in Sample by Treatment Arm - First Experiment



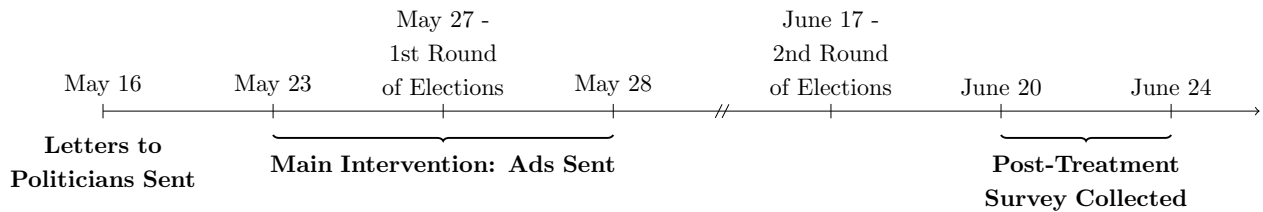
Notes: This figure shows a map of Colombia with the administrative boundaries of municipalities and the assignment to treatment arms in the first experiment. Municipalities in blue are part of the control group; those in orange are part of any of the treatment groups that receive any type of ads; finally, those in white are not in the experimental sample.

Figure 4: Randomization Design - 2018 Experiment



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

Figure 5: Timeline - 2018 Experiment



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

Figure 6: Basic Ad - 2018 Experiment



Translation:

If you have witnessed an irregularity or offense in these elections, file your report through the Attorney General's webpage. Click here: [...] You can also call [...] to file your report.

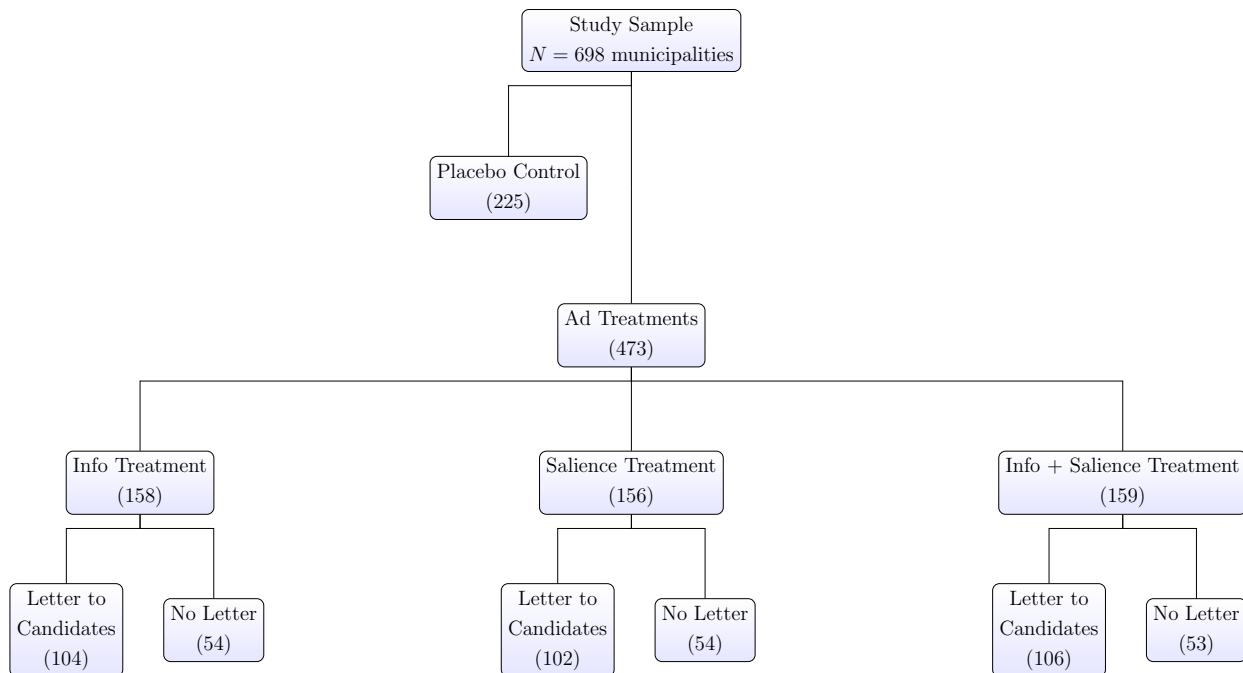
Report! Let's raise our voices against electoral corruption.

*File your report **here***

It's easy and anonymous

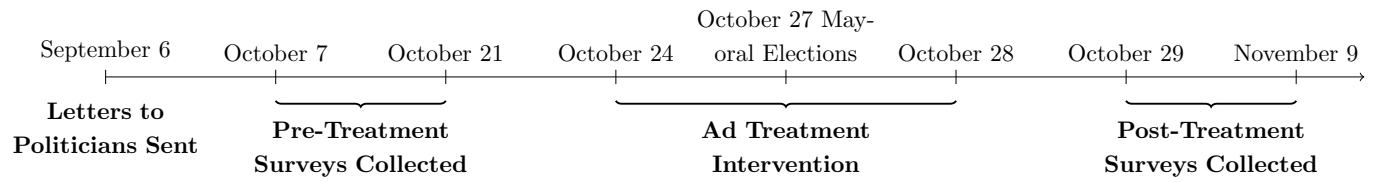
Notes: The image on the left displays the Basic Ad used in the first experiment, as it would be displayed on a cellphone screen. On the right is a translation to English of the text contained in the ad.

Figure 7: Randomization Design - 2019 Experiment



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

Figure 8: Timeline - 2019 Experiment



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

Figure 9: Ad Slideshow - 2019 Experiment

(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 in the second experiment 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 Salience message group was shown slides A, C and D. Finally, the group with both the Salience and the Information message was shown all of the slides, A-D.

Tables

Table 1: Summary Statistics

	2018 Sample			2019 Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
	Mean	Min	Max	Mean	Minimum	Maximum
Population 2018 (Thousands)	24.69	3.23	195.50	27.54	7.20	139.36
Facebook Penetration 2018	0.41	0.02	2.21	0.41	0.02	2.21
Population density (per km2)	106.18	0.52	3594.27	108.01	0.60	3594.27
Reports to AG 2018	0.14	0.00	6.00	0.14	0.00	6.00
Reports to MOE 2018	0.46	0.00	7.00	0.52	0.00	8.00
Reports to MOE 2015	3.15	0.00	29.00	3.47	0.00	33.00
Per Capita GDP 2016 (Millions of Pesos)	14.20	2.49	349.12	14.06	2.55	349.12
% Rural Population 2017	51.03	1.65	95.61	52.25	1.65	95.61
% Poor 2005	43.64	6.84	100.00	44.95	6.84	100.00
Sample size	652			681		

Notes: This table displays summary statistics for the sample of municipalities in the first experiment in 2018 (columns 1-3) and the second experiment in 2019 (columns 4-6) on a selected group of variables.

Table 2: Scale of Ad Campaigns

	2018 Experiment				2019 Experiment			
	Total	Per municipality	Per capita	Per registered to vote	Total	Per municipality	Per capita	Per registered to vote
People Reached	1,423,832	3265.67	0.14	0.36	4,358.870	6400.69	0.23	0.31
Impressions	4,443,565	10,191.66	0.45	1.12	12,886,427	18,922.8	0.69	0.92
People Clicking on Link*	12,396	28.43	1.40 (per 1000)	3.15 (per 1000)	23,418	75.30	2.7 (per 1000)	3.67 (per 1000)
People Reacting to Ad	3276	7.51	0.35 (per 1000)	0.83 (per 1000)	9623	14.13	0.51 (per 1000)	0.69 (per 1000)
Post Shares	1053	2.4	0.13 (per 1000)	0.27 (per 1000)	4531	6.65	0.24 (per 1000)	0.33 (per 1000)
Comments on Ad	130	0.3	0.01 (per 1000)	0.03 (per 1000)	437	0.64	0.02 (per 1000)	0.03 (per 1000)

Notes: This table reports several metrics of the scale of the Facebook advertisement campaigns in both the first and second experiments, as well as metrics of the engagement of Facebook users with the the ads. Notice that the metrics from the first experiment do not include the control group, since it did not receive any ads, but the metrics from the second experiment do indeed include the municipalities in the placebo control group. The variables reported in this table are defined as follows. People reached are the number of distinct individuals who saw the ads at least once. Impressions are the number of times the ads appeared on any screen. People clicking on the link are the number of distinct individuals who clicked on the link landing on the AG's or the MOE's reporting website, for the first and the second experiments, respectively . People reacting to the 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. *: For the metrics about the number of people clicking on the link in the second experiment we only considered the municipalities in actually receiving the link to the MOE's website (i.e. the ones including the Information message) and not all of the municipalities, as we do for the other metrics.

Table 3: Effects on Reports to the AG - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)
	1st Round reports		2nd Round reports		1st+2nd Round reports	
[<i>TF</i>] Facebook Ad Treatment	0.014** (0.006) [0.124]	0.014** (0.006) [0.111]	0.005 (0.003) [0.354]	0.005 (0.003) [0.506]	0.018*** (0.007) [0.033]	0.018*** (0.006) [0.031]
[<i>TL</i>] Letter to Politicians	0.006 (0.008) [0.424]	0.006 (0.007) [0.412]	0.000 (0.005) [0.746]	0.000 (0.004) [0.950]	0.007 (0.009) [0.444]	0.006 (0.009) [0.458]
Control Mean	0.00	0.00	0.00	0.00	0.00	0.00
Sample Size	652	652	652	652	652	652
Selected Controls		X		X		X

Notes: The outcome in the columns 1-2 is the number of reports received by the AG in the first round of Presidential elections. In columns 3-4 it is the number of reports received by the AG in the second round of Presidential elections. In columns 5-6 it is the combined number of reports received by the AG in both the first and second rounds of Presidential elections. Specifications in even-numbered columns 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: Effects on Reports to the MOE - 2019 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Reports (=1)		asinh(Number of Reports)		Higher Quality Reports (=1)		asinh(Higher Quality Reports)	
[TI] Info Message	0.160*** (0.046) [0.004]	0.243*** (0.057) [0.000]	0.232*** (0.064) [0.002]	0.335*** (0.077) [0.000]	0.104*** (0.040) [0.018]	0.153*** (0.049) [0.010]	0.103** (0.042) [0.032]	0.150*** (0.054) [0.038]
[TS] Saliency Message	0.061 (0.046) [0.254]	0.141** (0.055) [0.000]	0.108* (0.062) [0.120]	0.207*** (0.077) [0.000]	0.075* (0.040) [0.086]	0.123** (0.049) [0.004]	0.092** (0.047) [0.088]	0.138** (0.060) [0.000]
[TB] Info and Saliency Messages	0.152*** (0.045) [0.000]	0.233*** (0.055) [0.000]	0.224*** (0.063) [0.004]	0.325*** (0.076) [0.000]	0.134*** (0.040) [0.002]	0.183*** (0.050) [0.000]	0.168*** (0.049) [0.002]	0.214*** (0.058) [0.002]
[TL] Letter to Politicians		-0.123*** (0.046) [0.002]		-0.153** (0.065) [0.002]		-0.074* (0.042) [0.050]		-0.071 (0.049) [0.150]
Control Mean	0.24	0.24	0.30	0.30	0.13	0.13	0.13	0.13
Sample Size	681	681	681	681	681	681	681	681
Test $TI = TS$, p-value	0.06	0.06	0.10	0.09	0.54	0.51	0.84	0.82
Test $TB = TS$, p-value	0.08	0.07	0.12	0.11	0.21	0.21	0.20	0.19
Test $TB = TI$, p-value	0.88	0.86	0.91	0.90	0.52	0.54	0.24	0.25
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value		0.02		0.01		0.01		0.01
Selected Controls	X	X	X	X	X	X	X	X

Notes: The outcome in columns (1)-(2) is an indicator for whether any reports about the seven main types of irregularities presented in Section 3 was issued from each municipality. In columns (3)-(4) it is the inverse hyperbolic sine transformation of the number of such reports. In columns (5)-(6) the outcome is an indicator for whether any high or medium quality reports were issued from each municipality, while in columns (7)-(8) it is the inverse hyperbolic sine transformation of the number of such sub-set of reports (see Section 3 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 5: Effects on Electoral Irregularities - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)
	Irregularity Index	Illicit political advertising	Vote buying	Public servant campaigning	Voter intimidation	Electoral fraud
[TF] Facebook Ad Treatment	0.022 (0.061) [0.716]	-0.000 (0.006) [0.904]	0.007 (0.009) [0.562]	-0.005 (0.005) [0.194]	0.002 (0.002) [0.616]	-0.002 (0.005) [1.000]
[TL] Letter to Politicians	-0.101* (0.059) [0.078]	-0.009* (0.005) [0.198]	-0.015* (0.009) [0.122]	0.003 (0.003) [0.546]	0.003 (0.003) [0.442]	-0.000 (0.004) [0.978]
Control Mean	-0.00	0.01	0.02	0.00	0.00	0.01
Selected Controls	X	X	X	X	X	X

Notes: This table displays the effects of the different interventions performed in the first experiment in 2018 on reports about electoral irregularities that are collected by the MOE and are judged of a “high quality” (see Section 3 for a discussion about how quality is assessed). The outcome in column 1 is an index of electoral irregularities reported to the MOE which is defined in the main text. In the following columns the outcome is an indicator for whether each of the displayed types of electoral irregularities were reported and deemed of a high quality. Note that two types of electoral irregularities – “Voter deception” and “Fraud in voter registration” – did not have any high quality reports in our sample, and thus they are not reported in this table. 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 6: Effects on Electoral Irregularities - 2019 Experiment

	(1) Irregularity Index	(2) Illicit political advertising	(3) Vote buying	(4) Public servant campaigning	(5) Voter intimidation	(6) Fraud in voter registration	(7) Electoral fraud
[<i>TI</i>] Info Message	0.111 (0.077) [0.476]	0.034 (0.031) [0.428]	0.068** (0.032) [0.162]	0.052 (0.033) [0.100]	0.048 (0.037) [0.768]	0.037 (0.036) [0.206]	0.027 (0.039) [0.466]
[<i>TS</i>] Salience Message	0.015 (0.080) [0.878]	0.010 (0.033) [0.488]	0.019 (0.031) [0.658]	-0.002 (0.033) [0.736]	0.018 (0.036) [0.356]	0.021 (0.035) [0.954]	-0.060 (0.039) [0.008]
[<i>TB</i>] Info and Salience Messages	-0.024 (0.076) [0.340]	0.005 (0.033) [0.136]	-0.011 (0.032) [0.396]	-0.042 (0.033) [0.292]	0.038 (0.037) [0.036]	-0.014 (0.036) [0.972]	-0.038 (0.038) [0.344]
[<i>TL</i>] Letter to Politicians	-0.143** (0.061) [0.002]	-0.060** (0.025) [0.018]	-0.037 (0.024) [0.112]	-0.027 (0.026) [0.218]	-0.031 (0.029) [0.162]	-0.063** (0.028) [0.014]	-0.013 (0.030) [0.538]
Control Mean	0.00	0.64	0.74	0.68	0.36	0.61	0.42
Sample Size	2095	2437	2407	2369	2332	2333	2364
Test $TI = TS$, p-value	0.19	0.39	0.08	0.08	0.38	0.61	0.02
Test $TB = TS$, p-value	0.60	0.88	0.30	0.18	0.57	0.29	0.53
Test $TB = TI$, p-value	0.05	0.31	0.01	0.00	0.77	0.14	0.07
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value	0.06	0.07	0.63	0.30	0.90	0.07	0.19
Selected Controls	X	X	X	X	X	X	

Notes: This table displays the effects of the different interventions performed in the second experiment in 2019 on self-reported measures of the occurrence of electoral misdeeds collected in the post-treatment survey. The outcome in column 1 is an index of electoral irregularities defined in the main text. In the following columns the outcome is an indicator for whether each of the displayed types of electoral irregularities were deemed ‘likely’ or ‘very likely’ to have occurred by respondents. 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 municipality 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 7: Effects on Presidential Elections - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Turnout (%)		Non-traditional votes (%)		Traditional votes (%)		Blank votes (%)	
[TF] Facebook Ad Treatment	-0.157 (0.743) [0.810]	0.056 (0.400) [0.888]	3.592*** (1.296) [0.010]	1.839** (0.788) [0.016]	-3.638*** (1.308) [0.006]	-1.869** (0.792) [0.026]	0.070 (0.062) [0.276]	0.068 (0.054) [0.194]
[TL] Letter to Politicians	0.043 (0.697) [0.964]	0.339 (0.379) [0.362]	-1.180 (1.274) [0.328]	-0.592 (0.751) [0.430]	1.380 (1.286) [0.276]	0.732 (0.753) [0.306]	-0.153** (0.060) [0.010]	-0.075 (0.051) [0.126]
Control Mean	49.37	49.37	37.14	37.14	60.05	60.05	1.79	1.79
Sample Size	652	652	652	652	652	652	652	652
Selected Controls		X		X		X		X

Notes: The outcome in columns 1-2 is the turnout rate in the first round of presidential elections. In columns 3-4 and 5-6 it is the vote share for non-traditional and traditional candidates, respectively. In columns 7-8 it is the share of blank votes. For a discussion on how traditional candidates were defined see the main text. Specifications in even-numbered columns 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 8: Medium Term Effects on Second Round of Presidential Elections - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Turnout (%)		Duque votes (%)		Petro votes (%)		Blank votes (%)	
[TF] Facebook Ad Treatment	0.031 (0.677) [0.976]	0.236 (0.420) [0.578]	-3.463** (1.525) [0.024]	-1.194 (0.791) [0.138]	3.377** (1.562) [0.030]	1.137 (0.811) [0.144]	-0.001 (0.114) [0.996]	0.026 (0.082) [0.776]
[TL] Letter to Politicians	0.349 (0.632) [0.548]	0.624 (0.406) [0.116]	1.390 (1.488) [0.338]	0.996 (0.723) [0.168]	-1.290 (1.520) [0.382]	-1.056 (0.740) [0.182]	-0.089 (0.109) [0.370]	0.046 (0.076) [0.570]
Control Mean	51.22	51.22	60.76	60.76	34.70	34.70	2.92	2.92
Sample Size	652	652	652	652	652	652	652	652
Selected Controls		X		X		X		X

Notes: The outcome in columns 1-2 is the turnout rate in the second round of presidential elections. In columns 3-4 and 5-6 it is the vote share for the traditional candidate (Iván Duque), and the non-traditional candidate (Gustavo Petro). In columns 7-8 it is the share of blank votes. For a discussion on how traditional candidates were defined see the main text. Specifications in even-numbered columns 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 9: Effects on Mayoral Elections - 2nd Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Turnout (%)		Vote best candidates to fight corruption (%)		Vote worst candidates to fight corruption (%)		Blank Vote (%)	
[TI] Info Message	0.135 (0.692) [1.000]	0.126 (0.782) [0.574]	2.218 (1.999) [0.388]	1.405 (2.376) [0.146]	-1.624 (1.984) [0.502]	-0.902 (2.354) [0.296]	0.200* (0.108) [0.132]	0.407 (0.259) [0.036]
[TS] Saliency Message	0.411 (0.662) [0.618]	0.402 (0.798) [0.578]	3.966** (2.007) [0.084]	3.172 (2.374) [0.316]	-3.176 (1.983) [0.152]	-2.470 (2.346) [0.594]	0.204* (0.109) [0.106]	0.402 (0.254) [0.614]
[TB] Info and Saliency Messages	-0.201 (0.814) [0.836]	-0.210 (0.977) [0.494]	6.076*** (2.051) [0.012]	5.257** (2.436) [0.086]	-6.133*** (2.005) [0.018]	-5.405** (2.410) [0.042]	0.560* (0.326) [0.060]	0.761 (0.533) [0.000]
[TL] Letter to Politicians		0.014 (0.663) [0.974]		1.224 (1.853) [0.482]		-1.088 (1.845) [0.518]		-0.304 (0.331) [0.192]
Control Mean	67.72	67.72	76.87	76.87	20.52	20.52	1.57	1.57
Sample Size	681	681	641	641	641	641	681	681
Test $TI = TS$, p-value	0.67	0.67	0.40	0.39	0.45	0.45	0.98	0.97
Test $TB = TS$, p-value	0.43	0.43	0.32	0.32	0.16	0.16	0.26	0.26
Test $TB = TI$, p-value	0.68	0.68	0.07	0.07	0.03	0.03	0.26	0.26
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value		0.85		0.01		0.02		0.01
Selected Controls	X	X	X	X	X	X	X	X

Notes: The outcome in columns 1-2 is the turnout rate in the mayoral elections of 2019. In columns 3-4 and 5-6 it is the vote share for clean and non-clean candidates, respectively, as defined in the main text. In columns 7-8 it is the share of blank 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 10: Effects on Protests - 2018 Experiment

	(1)	(2)	(3)	(4)
	Protests against national government (=1)		Protests against other levels of government (=1)	
[TF] Facebook Ad Treatment	0.029* (0.015) [0.096]	0.032** (0.016) [0.062]	-0.010 (0.024) [0.662]	-0.015 (0.021) [0.484]
[TL] Letter to Politicians	0.007 (0.017) [0.708]	0.012 (0.017) [0.488]	0.026 (0.022) [0.224]	0.017 (0.020) [0.458]
Control Mean		0.02	0.02	0.09
Sample Size		652	652	652
Selected Controls			X	X

Notes: The outcome in columns 1-2 is an indicator that takes the value of 1 if the municipality presented any protest against the national government in the remaining months of 2018 after Iván Duque's victory, as recorded by CINEP. In columns 3-4 it is an indicator that takes the value of 1 if the municipality presented any protest against other levels of government. Specifications in even-numbered columns 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 11: Effects on Trust as Reported in Survey - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Trust	Trust	Trust	Trust	Trust	Trust	Trust	Trust	Trust	Trust
	index	(z-score)	elections	(z-score)	President	(z-score)	Judiciary	(z-score)	AG	(z-score)
[<i>TF</i>] Facebook Ad Treatment	-0.121*	-0.108	-0.164**	-0.158**	-0.092	-0.098	-0.075	-0.046	-0.120	-0.103
	(0.072)	(0.067)	(0.072)	(0.064)	(0.072)	(0.067)	(0.070)	(0.074)	(0.127)	(0.120)
	[0.096]	[0.126]	[0.018]	[0.028]	[0.214]	[0.150]	[0.276]	[0.530]	[0.390]	[0.402]
[<i>TL</i>] Letter to Politicians	0.107	0.076	0.121*	0.106*	0.040	0.029	0.061	0.050	0.234*	0.164
	(0.070)	(0.064)	(0.068)	(0.063)	(0.069)	(0.064)	(0.067)	(0.069)	(0.126)	(0.117)
	[0.134]	[0.264]	[0.082]	[0.134]	[0.520]	[0.658]	[0.386]	[0.438]	[0.086]	[0.164]
Control Mean	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.09	3.09
Sample Size	1029	1029	1029	1029	1029	1029	1029	1029	1029	1029
Selected Controls		X		X		X		X		X

Notes: The outcome in columns 1-2 is an index of the self-reported trust in the institutions shown in the remaining columns. The outcome in columns 3-4 is a measure of how much the respondent trusts elections on a scale of 1 to 7, where higher numbers represent more trust, which has been standardized with the mean and standard deviation of the control group. In columns 5-6 it is a measure of how much the respondent trusts the President on the same scale as before. In columns 7-8 it is a measure of how much the respondent trusts the judiciary on the same scale as before. In columns 9-10 it is a measure of how much the respondent trusts the AG on the same scale as before. Specifications in even-numbered columns include the covariates selected using the method described in Chernozhukov et al., 2015 and Belloni et al., 2014. Clustered standard errors at the municipality 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 12: Effects on Trust as Reported in Survey - 2019 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Trust		Trust		Trust		Trust		Trust	
	index	(z-score)	elections	(=1)	Mayor	(=1)	Judiciary	(=1)	the MOE	(=1)
[<i>TI</i>] Info Message	-0.027	-0.015	-0.022	-0.027	-0.001	0.001	-0.008	0.008	-0.030	-0.019
	(0.063)	(0.070)	(0.031)	(0.035)	(0.029)	(0.033)	(0.027)	(0.029)	(0.029)	(0.031)
	[0.680]	[0.430]	[0.494]	[0.850]	[0.972]	[0.552]	[0.818]	[0.650]	[0.392]	[0.684]
[<i>TS</i>] Salience Message	-0.004	0.008	0.015	0.010	-0.024	-0.022	0.005	0.021	-0.017	-0.005
	(0.061)	(0.074)	(0.030)	(0.037)	(0.029)	(0.033)	(0.026)	(0.029)	(0.028)	(0.033)
	[0.958]	[0.702]	[0.628]	[0.690]	[0.400]	[0.158]	[0.870]	[0.400]	[0.584]	[0.660]
[<i>TB</i>] Info and Salience Messages	0.006	0.018	0.002	-0.003	-0.018	-0.015	0.010	0.026	-0.006	0.005
	(0.058)	(0.067)	(0.027)	(0.033)	(0.028)	(0.032)	(0.025)	(0.028)	(0.029)	(0.035)
	[0.920]	[0.826]	[0.942]	[0.666]	[0.572]	[0.650]	[0.704]	[0.552]	[0.850]	[0.518]
[<i>TL</i>] Letter to Politicians		-0.018		0.008		-0.004		-0.024		-0.017
		(0.056)		(0.028)		(0.026)		(0.023)		(0.026)
		[0.748]		[0.754]		[0.888]		[0.290]		[0.480]
Control Mean	-0.00	-0.00	0.50	0.50	0.49	0.49	0.32	0.32	0.47	0.47
Sample Size	2460	2460	2548	2548	2514	2514	2530	2530	2527	2527
Test $TI = TS$, p-value	0.74	0.74	0.29	0.29	0.47	0.47	0.67	0.65	0.66	0.64
Test $TB = TS$, p-value	0.87	0.88	0.68	0.68	0.83	0.83	0.86	0.87	0.72	0.72
Test $TB = TI$, p-value	0.62	0.61	0.45	0.45	0.59	0.59	0.54	0.53	0.43	0.43
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value		0.78		0.96		0.52		0.80		0.34
Selected Controls	X	X	X	X	X	X	X	X	X	X

Notes: The outcome in columns 1-2 is an index of the self-reported trust in the institutions shown in the remaining columns. The outcome in columns 3-4 is an indicator of whether the respondent reports “trusting” or “trusting a lot” elections. In columns 5-6 it is an index of whether the respondent reports trusting the Mayor, defined as before. In columns 7-8 it is an index of whether the respondent reports trusting the Judiciary, defined as before. In columns 9-10 it is an index of whether the respondent reports trusting the MOE, defined as before. Specifications in even-numbered columns include the covariates selected using the method described in Chernozhukov et al., 2015 and Belloni et al., 2014. Clustered standard errors at the municipality 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 13: Effects on Perceived Importance of Corruption - 2019 Experiment

	(1) Importance of Corruption Index	(2) Importance of Corruption Index	(3) Corruption Most Important Problem in Country (=1)	(4) Corruption Most Important Problem in Country (=1)	(5) Corruption Most Important Problem in Municipality (=1)	(6) Corruption Most Important Problem in Municipality (=1)
[TI] Info Message	0.004 (0.069) [0.906]	0.092 (0.065) [0.362]	0.035 (0.032) [0.364]	0.061* (0.033) [0.054]	-0.027 (0.030) [0.430]	0.003 (0.028) [0.460]
[TS] Salience Message	-0.077 (0.069) [0.216]	0.011 (0.068) [0.782]	-0.016 (0.035) [0.656]	0.018 (0.033) [0.618]	-0.042 (0.030) [0.142]	-0.016 (0.030) [0.672]
[TB] Info and Salience Messages	0.066 (0.068) [0.604]	0.129* (0.067) [0.016]	0.064* (0.034) [0.106]	0.081** (0.034) [0.010]	-0.011 (0.029) [0.558]	0.016 (0.030) [0.226]
[TL] Letter to Politicians		-0.073 (0.052) [0.112]		-0.024 (0.026) [0.380]		-0.026 (0.023) [0.224]
Control Mean	0.01	0.00	0.36	0.35	0.28	0.27
Sample Size	2347	2339	2347	2363	2324	2339
Test $TI = TS$, p-value	0.29	0.20	0.17	0.17	0.65	0.51
Test $TB = TS$, p-value	0.06	0.06	0.04	0.05	0.35	0.27
Test $TB = TI$, p-value	0.42	0.55	0.44	0.52	0.64	0.66
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value		0.93		0.24		0.27
Selected Controls	X	X	X	X	X	X

Notes: The outcome in columns 1-2 is an index of how important corruption is perceived to be by respondents of the post-treatment survey in the second experiment in 2019, which is composed by the z-scores of the variables in the remaining columns. The outcome in columns 3-4 is an indicator of whether the respondent says corruption is the most important problem in the country. The outcome in columns 5-6 is an indicator of whether the respondent says corruption is the most important problem in their municipality. Specifications in even-numbered columns include the covariates selected using the method described in Chernozhukov et al., 2015 and Belloni et al., 2014. Clustered standard errors at the municipality 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.

Theoretical Appendix

Proof of Proposition 1

We begin by analyzing the voting behavior of citizens. Given equation (2), citizens will vote for party 1 with probability $Pr_j^1(S)$ given by:

$$\begin{aligned} Pr_j^1(S) &= Pr [F(S_1) + \sigma_j^1 + \delta^1 > F(S_2) + \sigma_j^2 + \delta^2] \\ &= Pr [F(S_1) - F(S_2) + \delta > \sigma_j] \\ &= \frac{1}{2} + \phi(F(S_1) - F(S_2) + \delta) \end{aligned}$$

Since there is a unit mass of voters and σ_j is independent across voters, this will be the actual total votes for party 1. However, since parties do not observe the realization of δ before taking their actions, the ex-ante probability of winning the election is:

$$\begin{aligned} Pr(Win_1 | S_1, S_2) &= Pr \left[\phi(F(S_1) - F(S_2) + \delta) + \frac{1}{2} > \frac{1}{2} \right] \\ &= Pr [\delta > F(S_2) - F(S_1)] \\ &= \frac{1}{2} + \psi(F(S_1) - F(S_2) + \mu) \end{aligned} \quad (5)$$

The parties' maximization problem for a given level of reporting R and other firms' strategies S_{-i} is given by:

$$\max_{S_i \in \mathbb{R}_+^3} \Pi_i(S_i, S_{-i}, R)$$

Notice $\Pi_i(S_i, S_{-i}, R)$ is concave in S_i and the form we chose for $F(S_i)$ satisfies the Inada conditions for S_i – for instance, for L_i :

$$\begin{aligned} \lim_{L_i \rightarrow 0} \frac{\partial F}{\partial L_i} &= \lim_{L_i \rightarrow 0} \alpha (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}-1} L_i^{\gamma-1} \rightarrow \infty \\ \lim_{L_i \rightarrow \infty} \frac{\partial F}{\partial L_i} &= \lim_{L_i \rightarrow \infty} \alpha (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}-1} L_i^{\gamma-1} \rightarrow 0 \end{aligned}$$

Thus, we can take the first order conditions for this problem and a interior solution is guaranteed:

$$\begin{aligned} \frac{\partial \Pi_i(S_i, S_{-i}, R)}{\partial L_i} &= \eta \alpha (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}-1} L_i^{\gamma-1} - c_L = 0 \\ \frac{\partial \Pi_i(S_i, S_{-i}, R)}{\partial O_i} &= \eta \alpha (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}-1} O_i^{\gamma-1} - c_0 - p(R)k = 0 \\ \frac{\partial \Pi_i(S_i, S_{-i}, R)}{\partial U_i} &= \eta \alpha (L_i^\gamma + O_i^\gamma + U_i^\gamma)^{\frac{\alpha}{\gamma}-1} U_i^{\gamma-1} - c_U = 0 \end{aligned}$$

Solving out this system of equations we get the best response functions for the parties described by equations $(BR_i - 1)$ - $(BR_i - 3)$ in Proposition 1.

We now analyze the citizens' problem. Since $p(R)$ is concave in r_j , and it satisfies the Inada conditions, the citizens' problem is concave and it has an interior maximum. Also notice, that the citizens' problem is symmetrical for all citizens, and thus $r_j = r$ is the same for all $j \in [0, 1]$.

We can thus take the first order conditions to find an interior solution:

$$2\eta p'(r^*)O^* - c = 0 \quad (6)$$

$$\Leftrightarrow r^* = (p')^{-1} \left(\frac{c}{2\eta O^* k} \right) \quad (7)$$

We now show existence and uniqueness of this equilibrium. First notice that $O^*(r)$ is (strictly) decreasing in r :

$$\frac{\partial O^*(r)}{\partial r} = p'(r)k \left[\underbrace{-\frac{\tilde{c}_O^{-\frac{1}{1-\alpha}-1}}{1-\alpha} (1 + \tilde{c}_O z)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}}}_{<0 \text{ if } \alpha < 1} + \underbrace{\frac{\alpha-\gamma}{(1-\gamma)(1-\alpha)} \tilde{c}_O^{-\frac{1}{1-\alpha} + \frac{\gamma}{1-\gamma}-1} (1 + \tilde{c}_O z)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)-1}}}_{<0 \text{ if } \alpha < \gamma} \right] < 0$$

Also notice $O^*(0)$ is a positive constant.

On the other hand, $r^*(O^*)$ is strictly increasing – since $p'(\cdot)$ is increasing –, and we have that $\lim_{O^* \rightarrow 0} r^*(O^*) \rightarrow 0$ and $\lim_{O^* \rightarrow \infty} r^*(O^*) \rightarrow \infty$, which follow directly from the Inada conditions on $p'(\cdot)$.

Thus, $O^*(r^*)$ and $r^*(O^*)$ intersect exactly once, showing that the equilibrium exists and is unique.

Proof of Proposition 2

Part 1 of Proposition 2 follows directly from derivating (7) and noticing that $((p')^{-1})'(\cdot) < 0$ since $p''(\cdot) < 0$:

$$\frac{\partial r^*}{\partial c} = ((p')^{-1})' \left(\frac{c}{2\eta O^* k} \right) \frac{1}{2\eta O^* k} < 0$$

Since $\frac{\partial r^*}{\partial c} < 0$ to prove part 2 we just need to show that $\frac{\partial O^*(r^*)}{\partial r^*} < 0$ – which we did in proving Proposition 1– and that $\frac{\partial L^*(r^*)}{\partial r^*} > 0$ and $\frac{\partial U^*(r^*)}{\partial r^*} > 0$. The two latter inequalities follow from differentiating best responses $(BR_i - 1)$ and $(BR_i - 3)$:

$$\frac{\partial L^*(r^*)}{\partial r^*} = \left(\frac{\psi\alpha}{c_L}\right)^{\frac{1}{1-\alpha}} \frac{(\gamma-\alpha)\gamma}{\gamma(1-\gamma)(1-\alpha)} c_L^{\frac{\gamma}{1-\gamma}} \left(1 + c_L^{\frac{\gamma}{1-\gamma}} \left[\frac{1}{(c_O + p(r^*)k)^{\frac{\gamma}{1-\gamma}}} + \frac{1}{c_U^{\frac{\gamma}{1-\gamma}}} \right]\right)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}-1} > 0$$

$$\frac{\partial U^*(r^*)}{\partial r^*} = \left(\frac{\psi\alpha}{c_U}\right)^{\frac{1}{1-\alpha}} \frac{(\gamma-\alpha)\gamma}{\gamma(1-\gamma)(1-\alpha)} c_U^{\frac{\gamma}{1-\gamma}} \left(1 + c_U^{\frac{\gamma}{1-\gamma}} \left[\frac{1}{c_L^{\frac{\gamma}{1-\gamma}}} + \frac{1}{(c_O + p(r^*)k)^{\frac{\gamma}{1-\gamma}}} \right]\right)^{\frac{\alpha-\gamma}{\gamma(1-\alpha)}-1} > 0$$

Finally part 3 follows directly from (5):

$$\frac{\partial Pr(Win_1 | S_1^*, S_2^*)}{\partial \omega} = \psi g'(\cdot)(H_1 - H_2) > 0 \quad \text{if } H_1 > H_2$$

A Appendix: Additional Materials

A.1 Effectiveness of Reporting

In this section we present a brief analysis of the AG’s records of reports received in order to illustrate whether reporting is an effective method of punishing electoral misdeeds. Table A.6 shows a breakdown of the reports received by the AG in the period 2010-2018 by their status in August 2019. Of the 6060 reports received in this period, only 2.5% have reached the final stage in which a decision was made over the case; a quarter of these found the accused guilty and sanctioned him, and the remaining were either acquitted or have been appealed. Apart from the cases in which decisions have been made, more than 12% are still in intermediate stages, either being investigated or being prepared for trial. Over 85% of them have been archived, either because they are not of the AG’s direct competence, do not have enough evidence about the occurrence of the irregularities or are duplicated. Only 0.08% were archived due to expired terms. As expected, more recent reports are less likely to have been decided: whereas for 1.75% reports made in the period 2016-2018 a decision has been made, this number goes up to 2.85% for the ones made in the period 2010-2015.

While these numbers should be taken with caution for the reasons outlined earlier, a conservative conclusion one gets from this analysis is that reports made to AG are, at least to some extent, a useful way of sanctioning irregularities since at least some of the electoral irregularities reported are ultimately sanctioned. Knowing how many worthy reports go without sanctioning or how many unworthy reports are included in the statistics presented is impossible without having an objective measure of the quality of reports, which is unfortunately not available in this context.

A.2 Candidates in the 2018 Presidential Elections

In this section we give more details about the background of the candidates participating in the 2018 Presidential elections and we discuss some limitations of our definition of ‘traditional’ and ‘non-traditional’ candidates.

We begin by discussing each of the five candidates participating in these elections. Ivan Duque came into the first round of elections as the candidate with the highest intention to vote according to virtually every poll. Duque was supported mainly by *Centro Democrático*, a right-wing party founded in 2013 by ex-president, Álvaro Uribe, and smaller parties which, jointly, possessed over 21% of seats in Congress. Gustavo Petro, the forerunner in most intention polls, was, in contrast, supported by a small coalition of socialist and communist parties which held less than 4% of seats in the recent Congressional elections. The candidate to come in third in most opinion polls was Sergio Fajardo, a center-left candidate mainly supported by the Green Party (later called *Alianza Verde*), a young party that constituted one of most popular opposition parties in the past couple of elections, and the *Polo Democrático Alternativo*, a left wing party established in the early 2000s. Fajardo’s supporting parties reached a historic 9% of seats in the 2018 congressional elections.

Candidate Germán Vargas held the largest coalition of traditional parties belonging mostly to the center and the center-right. Former president, Juan Manuel Santos’ party, the *Partido de la U*, as well as Vargas’ own party, *Cambio Radical*, the Conservative Party, and a handful of smaller parties backed Vargas. They jointly held approximately 44% of Congress seats. Finally, Humberto De la Calle came in as the main candidate to support the peace deal achieved between Ex-President Santos’ government and the FARC guerrilla group, backed mainly by a centrist coalition of the Liberal Party, one of the oldest parties in the country, and a couple of smaller parties which gathered over 17% of seats in Congress.

The distinction between traditional candidates and non-traditional candidates that we have made here also broadly aligns closely with the distinction between center and right wing candidates, which are supported by more historically powerful parties, and left-leaning candidates, which have been backed by smaller and younger parties.

A.3 Ad Delivery and Budget Details for Both Experiments

In both experiments, the Facebook ad campaign was set to use the maximum budget per municipality and per day, so that ads were delivered throughout the intervention period in roughly even pattern. Similarly, the delivery optimization strategy was set to maximize reach, in an attempt to approximate a uniform distribution of people receiving the ad within the Facebook user population in each municipality.⁵⁰

For the first experiment, we set the budget allotted to each municipality so that it was directly proportional to the logarithm of the population in each municipality. The exact rule to allocate the budget to each municipality was calibrated using the population of a sub-sample of municipalities to predict the costs that Facebook Ad Manager forecasted to reach half of the population of users in each municipality.⁵¹ For the second experiment, we instead used the data from our first experiment to calculate the average cost of the advertisements in each bin defined by the population deciles in the sample of municipalities. We then estimated how much it would cost to reach a third of the users in each municipality and used this as our budget.

⁵⁰The Facebook Ad Manager documentation is notoriously obscure in disclosing how the ad delivery optimization algorithms are designed. However, the reach maximization setting seems to be the best at guaranteeing that the ad is seen by the largest population possible. According to the online documentation, “[t]he reach objective maximizes the number of people who see your ads and how often they see them”, which can be compared to other options such as ‘traffic objective’, which is described as targeting “[...] people in your audience who are most likely to click the ads” and would thus disproportionately show the ad more to users which have been shown to be more likely to click on such links in the past (Facebook, 2019).

⁵¹As we show in the results section in the first experiment the ad campaign did not reach our initial target of half of the user population, but this was expected as Facebook’s estimates are only approximate and the election period might have increased the demand for ads.

A.4 Covariates Included in the Analysis

The covariates included in the analysis for our first two experiments can be broadly categorized in three groups:

Past reports: We include the number of reports made to the MOE in the 2015 local elections and the 2018 congressional elections, as well as the reports made to the AG in the 2018 congressional elections, as a way to control both for previous experience with reporting channels and the prevalence of electoral irregularities.

Socioeconomic characteristics: As geographical and demographic variables we use the municipal population in 2018, the population density, the proportion of rural population and the municipalities' altitude. As measures of economic activity and development we use GDP per capita, the % of poor population⁵² and the distance from the nearest wholesale market. All of these variables were taken from the the *National Department of Statistics* (DANE) except for the last one which was taken from the Municipal Characteristics database held by the *Centro de Estudios sobre Desarrollo Económico* (CEDE). From the CEDE's database we also used the municipal homicide rate and inflow and outflow of people displaced as violence proxies. Finally, we also included Facebook's penetration rate (defined as the number of active Facebook users divided by total population), the number of protests against the government gathered by CINEP between 2014 and May 26, 2018, and an indicator of whether each municipality suffered connectivity problems in their public wi-fi spots.⁵³

Political preferences: In order to get a rich set of political characteristics for each municipality we used the turnout and the vote share for each major party in the 2018 congressional elections, the vote share for each candidate in the second round of the 2014 presidential elections and the winning margin of the elected Mayor in the 2015 local elections. All of these variables were constructed from the official records held by the *Registraduría Nacional*.

A.5 Effects of the 2018 Intervention on each Candidate's Vote Share

In Table A.10 we report the effect of the intervention on each candidates' vote shares separately using administrative data. Consistent with the previous results, the ad campaign had the unambiguous effect of increasing non-traditional candidates' vote share and decreasing the vote share for traditional candidates, although many of the estimates lose statistical significance. The ad campaign had the largest effects on Duque and Petro, who were the main contenders according to the polls up to one week before the elections (see Section 3), as well as on Vargas. The effect of the pooled ad treatments on these candidates was to increase Petro's vote share by approximately

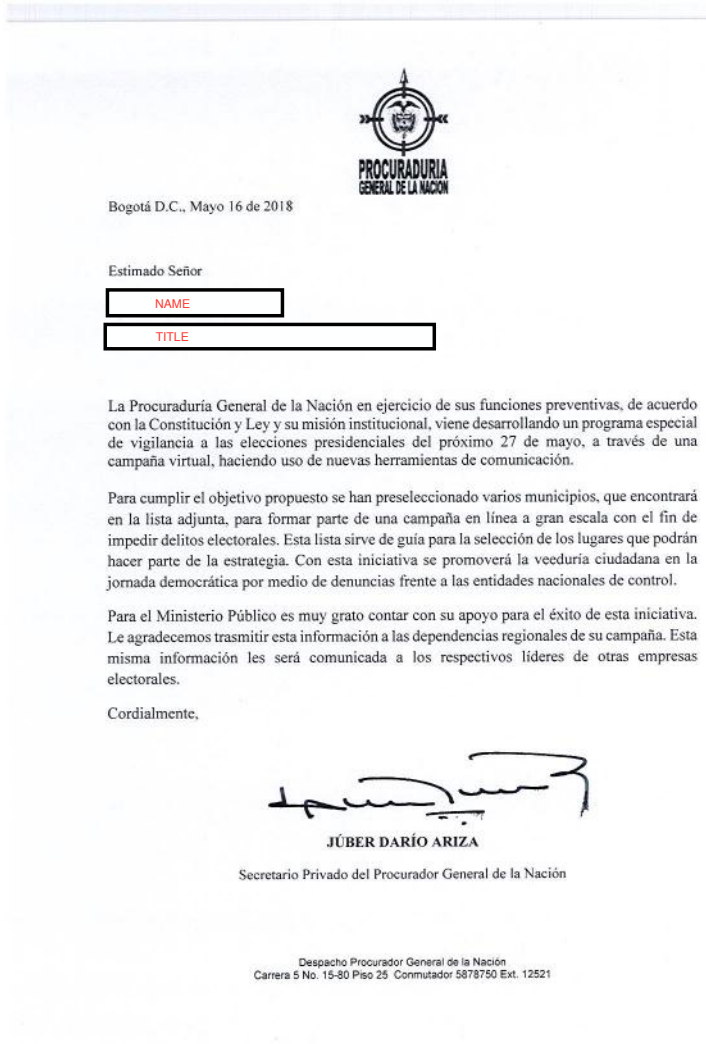
⁵²This is defined by DANE as the % of people living without a set of basic needs.

⁵³These connectivity issues occurred due to delays in payments from local governments to internet providers. They took place throughout our intervention period and limited the access to public wi-fi spots in slightly less than a half of the municipalities in our sample.

1.4% (p-value ≈ 0.1), and to reduce Duque's by 0.7% (p-value > 0.1) and Vargas' by 0.9% (p-value ≈ 0.1). As before, these effects are in general largest for the 100% exposure treatments and similar for the Basic and Efficiency ads.

A Figure Appendix

Figure A.1: Letter sent to politicians for the 1st experiment

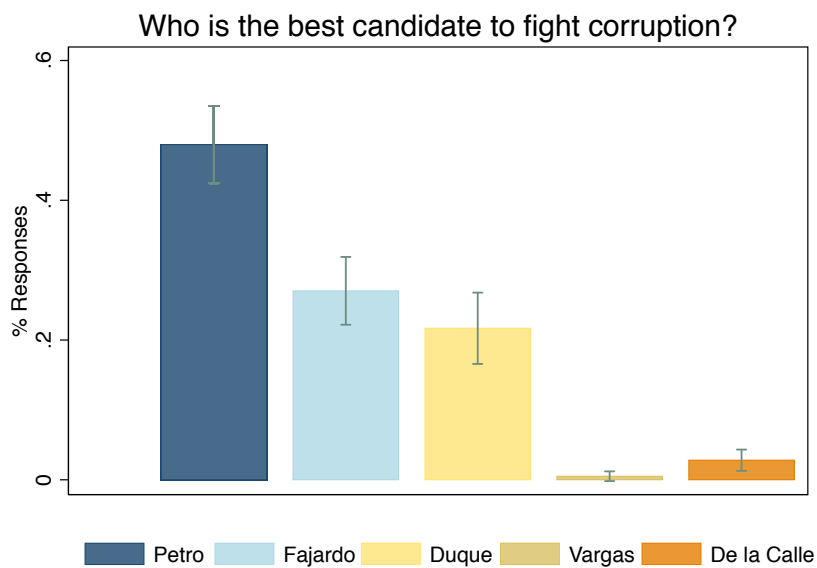


Translation:

Dear Mr. (NAME), (TITLE)
The Attorney General of the Nation, in the exercise of its preventive functions stipulated in the Constitution and its Institutional Mission, is implementing a special program to watch over the forthcoming presidential elections of May 27 through an online campaign that makes use of the new communication tools.
For this purpose, several municipalities that you will find in an attached list have been pre-selected to participate in an online large-scale campaign to prevent electoral irregularities. This list acts indicates which municipalities might be included in our strategy. This initiative will promote the civilian oversight of the elections by encouraging reports made to the national watchdog institutions.
The Public Ministry welcomes your help in the success of this initiative. We ask you to spread this information to your campaigns' regional offices. This same information will be communicated to the leaders of the other campaigns.[...]

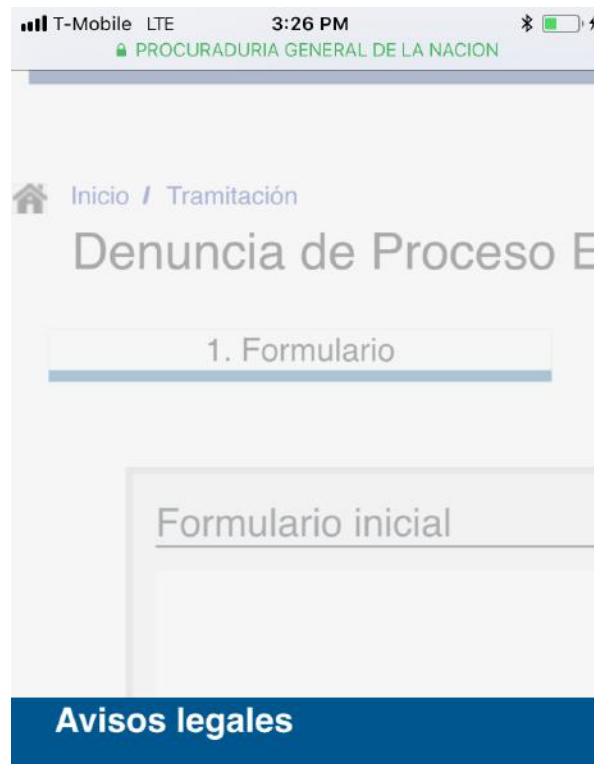
Notes: An example of an actual letter sent to politicians in the first experiment is shown on the left. On the right is a translation to English of the text contained in the letter.

Figure A.2: Survey Responses: Best Candidate to Fight Corruption



Notes: This figure plots the responses to the question “Who is the best candidate to fight corruption?” among survey participants in the first experiment. Details about the survey performed in the first experiment are found in Section 4.3.1. The sample is restricted to only respondents coming from control municipalities ($N = 392$) so that responses are not affected by the different treatment conditions. 95% confidence intervals for the average response to each mentioned reason are shown in gray.

Figure A.3: Illustration of Cellphone Compatibility Issue

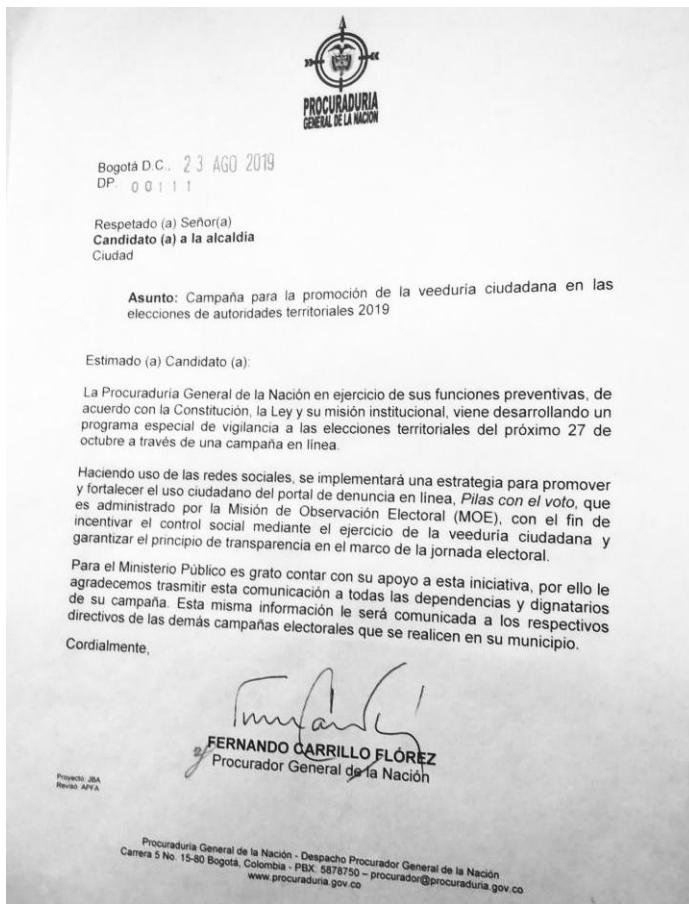


u responsabilidad, que los datos
umple con los requisitos estable
a la pretensión realizada.

en este formulario pasarán a f

Notes: This figure displays a screen shot of how the AG's landing page looks like on a cellphone.

Figure A.1: Letter sent to politicians allowing for signal jamming - Second experiment



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.

Making use of social media, 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. 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: An example of an actual letter sent to politicians allowing for signal jamming in the second experiment is shown on the left. On the right is a translation to English of the text contained in the letter.

A Tables Appendix

Table A.1: Summary Statistics of Survey vs Full Samples - 2018 Experiment

	(1) Mean Full Sample	(2) Mean Survey Sample	(3) Difference (2)-(1)
Population 2018 (Thousands)	24.69	32.89	8.20*** (1.61)
Facebook Penetration 2018	0.41	0.49	0.08*** (0.02)
Population density (per km)	106.18	155.46	49.28* (19.88)
Reports to AG 2018	0.14	0.19	0.05 (0.04)
Reports to MOE 2018	0.46	0.63	0.17* (0.07)
Reports to MOE 2015	3.15	3.79	0.64* (0.29)
Per Capita GDP 2016 (Millions of Pesos)	14.20	15.51	1.31 (1.57)
% Rural Population 2017	51.03	43.01	-8.02*** (1.48)
% Poor 2005	43.56	38.04	-5.52*** (1.32)

Notes: This table reports the mean value of a group of select group of variables for the full sample of municipalities in the first experiment in 2018 (column 1), the sample of municipalities with respondents in the post-treatment survey (column 2), and their difference (column 3). Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A.2: Summary Statistics of Post-Treatment Survey vs Full Samples - 2019 Experiment

	(1) Mean Full Sample	(2) Mean Survey Sample	(3) Difference (2)-(1)
Population 2018 (Thousands)	27.56	28.36	0.80 (1.34)
Facebook Penetration 2018	0.41	0.44	0.02 (0.02)
Population density (per km)	108.08	108.65	0.58 (11.93)
Reports to AG 2018	0.14	0.15	0.01 (0.03)
Reports to MOE 2018	0.52	0.53	0.01 (0.06)
Reports to MOE 2015	3.47	3.51	0.03 (0.25)
Per Capita GDP 2016 (Millions of Pesos)	14.07	14.45	0.38 (1.17)
% Rural Population 2017	52.27	51.80	-0.47 (1.30)
% Poor 2005	44.96	43.93	-1.03 (1.16)

Notes: This table reports the mean value of a group of select group of variables for the full sample of municipalities in the second experiment in 2019 (column 1), the sample of municipalities with respondents in the post-treatment survey (column 2), and their difference (column 3). Robust standard errors are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A.3: Covariate Balance - 2018 Experiment

	Full Sample			Post-Treatment Survey Sample		
	(1) Control Group Mean	(2) Facebook Ad Treatment vs Control Difference	(3) Letter to Politicians vs No Letter Difference	(4) Control Group Mean	(5) Facebook Ad Treatment vs Control Difference	(6) Letter to Politicians vs No Letter Difference
Panel A. Previous Reports Covariates						
Reports to AG 2018	0.137	0.038 (0.039)	-0.009 (0.042)	0.259	0.172* (0.092)	0.064 (0.104)
Reports to MOE 2018	0.802	0.019 (0.083)	-0.080 (0.077)	0.922	-0.121 (0.221)	-0.492*** (0.187)
Reports to MOE 2015	3.829	-0.149 (0.323)	-0.276 (0.315)	5.174	-1.104 (1.054)	0.460 (0.954)
Panel B. Socioeconomic Covariates						
Population 2018	37,157.175	632.091 (1,678.290)	277.196 (1,632.440)	49,752.977	-3,493.443 (6,340.578)	-7,625.608 (6,226.537)
Population density (per km)	156.120	-6.942 (21.701)	14.405 (18.408)	330.304	-189.168 (145.493)	154.083 (120.164)
GDP pc 2016 (Ms of Pesos)	13.722	0.838 (1.442)	-1.918 (1.610)	16.458	1.028 (2.996)	-1.594 (3.269)
% Poor 2005	45.009	-0.884 (1.709)	0.635 (1.581)	30.115	2.359 (3.101)	-0.194 (2.983)
% Rural Population 2017	55.882	0.269 (1.884)	0.783 (1.753)	32.735	3.670 (4.129)	3.467 (3.787)
Homicide Rate 2017	23.458	2.819 (2.476)	-1.714 (2.486)	24.960	4.625 (3.344)	-1.304 (3.446)
Displaced People 2017	67.071	-2.661 (12.622)	15.629 (12.855)	45.779	4.271 (12.197)	-9.078 (11.516)
Displaced People Received 2017	65.521	4.146 (8.909)	11.224 (10.167)	39.255	-2.407 (9.360)	-5.373 (8.150)
Protests 2014-2018	2.763	0.104 (0.225)	0.141 (0.216)	2.940	1.275* (0.752)	0.010 (0.856)
Facebook Penetration 2018	0.398	-0.008 (0.024)	-0.025 (0.022)	0.530	0.032 (0.043)	-0.047 (0.041)
Distance to Main Mkt (Kms)	130.091	-1.367 (7.680)	-5.579 (7.344)	94.510	13.017 (12.896)	3.924 (12.801)
Altitude (meters)	1,139.089	75.935 (91.380)	31.316 (101.004)	1,181.616	52.325 (189.116)	16.755 (172.354)
Disconnected Wi-fi spots (=1)	0.475	0.053 (0.041)	0.049 (0.039)	0.534	0.118 (0.093)	-0.066 (0.089)
Panel C. Political Covariates						
Santos Vote Share 2014 (%)	47.759	2.092 (1.761)	0.356 (1.672)	48.075	1.613 (2.509)	0.335 (2.461)
Zuluaga Vote Share 2014 (%)	49.535	-2.125 (1.708)	-0.286 (1.623)	48.032	-1.418 (2.336)	-0.344 (2.296)
Mayor Wining Margin 2015	14.079	1.298 (0.907)	-0.557 (0.876)	16.826	-1.059 (4.522)	-4.707 (3.716)
Turnout for Congress 2018	0.504	-0.001 (0.007)	0.003 (0.007)	0.517	-0.008 (0.011)	0.003 (0.010)
Liberals Vote Share 2018	15.481	-1.129 (0.919)	-0.392 (0.863)	13.488	1.660 (1.439)	0.642 (1.441)
Cambio Radical Vote Share 2018	15.243	1.431 (0.982)	0.910 (0.953)	13.570	-0.994 (1.593)	-2.839* (1.461)
Centro Dem Vote Share 2018	13.094	-1.639* (0.900)	-0.531 (0.825)	15.007	-3.109** (1.493)	0.138 (1.400)
Partido de la U Vote Share 2018	14.656	-0.065 (0.938)	0.030 (0.894)	14.916	0.355 (1.666)	2.784* (1.633)
Liberals Vote Share 2018	15.481	-1.129 (0.919)	-0.392 (0.863)	13.488	1.660 (1.439)	0.642 (1.441)
Green Party Vote Share 2018	6.002	0.064 (0.557)	-0.328 (0.552)	4.960	-0.305 (1.031)	-0.392 (0.919)
Polo Vote Share 2018	2.731	-0.110 (0.245)	-0.390* (0.208)	3.437	0.629 (0.441)	-0.678 (0.502)
Decentes Vote Share 2018	1.205	0.122 (0.107)	-0.118 (0.094)	1.708	-0.072 (0.355)	-0.310 (0.287)
Share Blank Votes 2018	4.596	0.249 (0.354)	-0.623* (0.335)	6.110	-0.018 (1.277)	-1.276 (1.102)

Notes: This table presents the balance checks for the first experiment in 2018. Columns (1)-(3) use the full sample of 652 municipalities, while columns (4)-(6) use the sample of 328 municipalities from which we collected responses in the post-treatment survey. Clustered standard errors at the municipality level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A.4: Covariate Balance - 2019 Experiment

	Full Sample					Post-Treatment Survey Sample				
	(1) Control Group Mean	(2) Information vs Control Difference	(3) Salience vs Control Difference	(4) Info + Salience vs Control Difference	(5) Letter to Politicians vs No Letter	(6) Control Group Mean	(7) Information vs Control Difference	(8) Salience vs Control Difference	(9) Info + Salience vs Control Difference	(10) Letter to Politicians vs No Letter
Panel A. Previous Reports Covariates										
Reports to AG 2018	0.141	0.022 (0.062)	0.005 (0.061)	0.044 (0.070)	-0.136** (0.066)	0.169	0.013 (0.092)	-0.018 (0.087)	0.017 (0.092)	-0.148** (0.067)
Reports to MOE 2018	0.523	0.082 (0.112)	-0.048 (0.106)	0.014 (0.109)	-0.078 (0.101)	0.599	0.135 (0.154)	-0.109 (0.137)	0.074 (0.158)	-0.168 (0.138)
Reports to MOE 2015	3.474	0.318 (0.484)	-0.271 (0.453)	0.083 (0.495)	0.095 (0.427)	3.620	0.608 (0.635)	-0.418 (0.501)	-0.300 (0.506)	-0.357 (0.514)
Panel B. Socioeconomic Covariates										
Population 2018	27,564.517	638.202 (2,749.538)	-2,723.354 (2,364.434)	-1,811.492 (2,449.003)	-1,110.077 (2,292.049)	30,161.693	1,553.515 (3,385.845)	-3,252.330 (2,932.392)	-2,406.861 (2,911.201)	-3,766.398 (2,899.253)
Population density (per km)	108.076	31.476 (26.970)	-9.380 (21.170)	-8.974 (20.233)	13.037 (16.454)	110.553	42.573 (32.818)	-3.403 (23.373)	-15.534 (22.303)	7.938 (21.027)
GDP pc 2016 (Ms of Pesos)	14.069	-1.257 (2.175)	-2.310 (1.985)	0.524 (2.603)	2.175 (1.447)	15.624	-1.552 (3.933)	-2.975 (3.425)	2.282 (4.671)	4.553** (2.193)
% Poor 2005	44.958	1.744 (2.261)	1.771 (2.227)	0.685 (2.123)	-3.480 (2.127)	42.656	1.280 (2.702)	-1.886 (2.259)	0.182 (2.361)	-4.223* (2.535)
% Rural Population 2017	52.275	-5.917** (2.566)	-2.614 (2.389)	-0.535 (2.467)	0.578 (2.347)	49.358	-4.483 (2.922)	-1.703 (2.441)	2.204 (2.692)	1.680 (2.588)
Homicide Rate 2017	27.132	-3.830 (3.113)	0.208 (3.573)	-2.123 (3.179)	1.119 (2.848)	27.149	-2.000 (3.298)	0.730 (3.637)	-1.944 (3.213)	0.360 (3.058)
Displaced People 2017	65.098	-21.705* (12.305)	22.879 (25.763)	-15.479 (11.930)	-6.614 (17.453)	59.232	-18.196 (11.172)	-1.004 (19.198)	-12.702 (11.502)	-11.072 (13.594)
Displaced People Received 2017	40.797	-0.151 (10.726)	28.044 (21.969)	-0.771 (9.123)	-8.497 (15.444)	37.848	-0.592 (9.877)	11.802 (16.053)	3.352 (10.342)	-9.517 (12.674)
Protests 2014-2018	1.846	0.150 (0.323)	0.115 (0.343)	0.003 (0.288)	-0.152 (0.304)	1.976	0.342 (0.553)	0.183 (0.552)	0.011 (0.355)	-0.376 (0.563)
Facebook Penetration 2019	0.415	0.028 (0.032)	0.032 (0.031)	0.001 (0.030)	0.005 (0.029)	0.473	0.024 (0.040)	0.042 (0.041)	0.012 (0.037)	0.039 (0.035)
Distance to Main Mkt (Kms)	123.063	0.650 (8.845)	5.493 (8.721)	-0.813 (10.031)	-9.132 (9.405)	122.342	-3.270 (10.126)	-7.187 (9.993)	-3.950 (10.988)	-7.693 (10.162)
Altitude (meters)	1,042.467	118.296 (182.761)	-5.873 (92.148)	25.114 (91.560)	202.730* (117.346)	1,018.079	91.791 (141.068)	118.457 (104.464)	98.560 (107.200)	200.310* (109.671)
Disconnected Wi-fi spots (=1)	0.479	0.066 (0.058)	0.057 (0.057)	0.041 (0.058)	-0.010 (0.055)	0.526	0.023 (0.068)	0.001 (0.066)	0.032 (0.067)	0.029 (0.063)
Panel C. Political Covariates										
Santos Vote Share 2014 (%)	49.379	2.445 (2.261)	1.035 (2.268)	-0.162 (2.314)	1.824 (2.025)	47.047	2.450 (2.640)	-0.355 (2.531)	-2.947 (2.477)	0.707 (2.350)
Zuluaga Vote Share 2014 (%)	47.902	-2.535 (2.192)	-1.214 (2.200)	0.070 (2.251)	-1.963 (1.961)	50.081	-2.569 (2.556)	0.048 (2.452)	2.740 (2.417)	-0.944 (2.279)
Mayor Wining Margin 2015	13.250	0.415 (1.326)	-0.505 (1.135)	-0.169 (1.175)	1.143 (1.101)	13.465	0.306 (1.702)	-0.969 (1.362)	-0.000 (1.386)	0.533 (1.387)
Turnout for Congress 2018	0.500	0.020** (0.010)	0.000 (0.009)	0.009 (0.010)	-0.003 (0.009)	0.498	0.015 (0.012)	0.002 (0.010)	-0.000 (0.010)	-0.010 (0.011)
Liberals Vote Share 2018	15.364	-0.913 (1.125)	0.508 (1.199)	0.417 (1.140)	0.664 (1.062)	14.706	-0.267 (1.292)	0.411 (1.311)	0.636 (1.394)	0.645 (1.274)
Cambio Radical Vote Share 2018	15.129	0.761 (1.356)	-1.116 (1.335)	-0.217 (1.294)	-1.435 (1.229)	14.675	1.361 (1.704)	-1.166 (1.413)	-0.517 (1.365)	-1.813 (1.478)
Centro Dem Vote Share 2018	12.594	-1.634 (1.088)	-1.654 (1.093)	-0.829 (1.123)	0.184 (0.957)	14.089	-1.480 (1.322)	-0.978 (1.334)	0.869 (1.388)	0.604 (1.197)
Partido de la U Vote Share 2018	15.174	1.010 (1.176)	-0.058 (1.205)	0.431 (1.167)	-0.590 (1.106)	14.753	-0.890 (1.401)	-0.741 (1.330)	-1.285 (1.341)	-0.681 (1.141)
Liberals Vote Share 2018	15.364	-0.913 (1.125)	0.508 (1.199)	0.417 (1.140)	0.664 (1.062)	14.706	-0.267 (1.292)	0.411 (1.311)	0.636 (1.394)	0.645 (1.274)
Green Party Vote Share 2018	4.587	0.278 (0.608)	0.585 (0.738)	-0.337 (0.548)	-0.238 (0.660)	4.684	-0.054 (0.672)	-0.161 (0.713)	-0.580 (0.631)	-0.224 (0.633)
Polo Vote Vote Share 2018	2.756	-0.248 (0.273)	0.229 (0.403)	-0.307 (0.263)	0.132 (0.316)	2.848	-0.284 (0.341)	0.287 (0.426)	-0.310 (0.342)	0.439 (0.275)
Decentes Vote Share 2018	1.302	0.120 (0.124)	0.027 (0.156)	0.169 (0.150)	0.098 (0.143)	1.410	0.147 (0.148)	0.116 (0.187)	0.270 (0.202)	0.063 (0.207)
Share Blank Votes 2018	4.470	0.313 (0.482)	-0.210 (0.446)	0.056 (0.442)	-0.088 (0.436)	4.806	0.008 (0.597)	-0.079 (0.540)	0.442 (0.561)	0.374 (0.549)
Number of candidates 2019	4.849	-0.114 (0.204)	-0.183 (0.213)	-0.333 (0.203)	-0.177 (0.201)	4.937	-0.081 (0.238)	-0.267 (0.232)	-0.320 (0.235)	-0.316 (0.213)

Notes: This table presents the balance checks for the second experiment in 2019. Columns (1)-(5) use the full sample of 681 municipalities, while columns (6)-(10) use the sample of 621 municipalities from which we collected responses in the post-treatment survey. Clustered standard errors at the municipality level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A.5: Demographic Covariate Balance for Post-Treatment Survey Respondents - 2019 Experiment

	(1) Control Group Mean	(2) Information vs Control Difference	(3) Saliency vs Control Difference	(4) Info + Saliency vs Control Difference	(5) Letter to Politicians vs No Letter
Female (=1)	0.516	0.008 (0.029)	0.038 (0.025)	-0.000 (0.026)	0.017 (0.027)
Age	33.501	-0.918 (0.676)	0.238 (0.748)	-0.128 (0.643)	0.581 (0.667)
Less Than High School (=1)	0.099	0.007 (0.016)	-0.008 (0.015)	0.007 (0.016)	0.023* (0.014)
High School (=1)	0.357	0.025 (0.028)	-0.009 (0.027)	-0.034 (0.026)	-0.033 (0.026)
More than High School	0.104	-0.017 (0.016)	-0.000 (0.017)	0.015 (0.017)	0.015 (0.015)

Notes: This table presents the balance checks for the demographic characteristics of the respondents of the post-treatment survey conducted for the second experiment in 2019. Clustered standard errors at the municipality level are shown in parentheses; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.

Table A.6: Reports Made to the AG by Status of Case

Period:	2010-2018	2010-2015	2016-2018
Intermediate Case Stages	820 (12.41%)	265 (5.84%)	555 (26.80%)
Archived	5616 (85.01%)	4136 (91.20%)	1480 (71.46%)
Archived (overdue)	5 (0.08%)	5 (0.11%)	0 (0%)
Decision made - Sanction	43 (0.65%)	35 (0.77%)	8 (0.39%)
Decision made - Acquittal	75 (1.14%)	61 (1.35%)	14 (0.68%)
Decision made - Other	47 (0.71%)	33 (0.73%)	14 (0.68%)
Total	6606	4535	2071

Notes: This table displays the reports about electoral irregularities received by the AG that were not redirected to other agencies broken down by the status of the cases opened and by the period in which the reports were received by the AG. The definition for each of the status categories is presented in the main text. The percentage share of reports with respect with total reports in each period is shown in parenthesis.

Table A.7: Summary of Presidential Candidates in 2018

<i>Candidate</i>	<i>Position in political spectrum</i>	<i>Main parties supporting the candidate</i>	<i>% of seats in Congress held by supporting parties</i>
Duque	Right wing	Centro Democrático	21.1%
Vargas	Center	Cambio Radical, Partido de la U and Conservative Party	43.7%
De la Calle	Center	Liberal Party	17.6%
Petro	Left wing	Colombia Humana, Alternative Indigenous and Social Movement, Unión Patriótica, Communist Party	3.6%
Fajardo	Center-left	Alianza Verde, Polo Democrático Alternativo	9.0%

Notes: This table summarizes the position in the political spectrum of each candidate running in the 2018 presidential elections, the main parties supporting them and whether we classify them as traditional or not, using the definition provided in the main text. The last column shows the percentage of seats in Congress, combining both the House of Representatives and the Senate, held by the coalition of parties supporting each candidate.

Table A.8: Timing of Irregularities According to Reports

Election Type and Year:	Local 2015		Congress 2018	
	Six or more days before election	After	Six or more days before election	After
Vote Buying	372 (42%)	514 (58%)	119 (24%)	380 (76%)
Public Servants Campaigning	565 (77%)	172 (23%)	115 (60%)	76 (40%)
Fraud in Voter Registration	604 (88%)	81 (12%)	27 (84%)	5 (16%)
Voter Intimidation	224 (66%)	114 (34%)	108 (49%)	114 (51%)
Electoral Fraud	10 (10%)	86 (90%)	3 (4%)	73 (96%)

Notes: This table compares the number of reports made to the MOE about electoral irregularities occurring six days or more before elections to those occurring after this date broken down by the five main types of irregularities. The first two columns display the reports for the 2015 local elections and the last two columns display the results for the 2018 congressional elections. The percentage share of reports of each type of irregularity is shown in parenthesis.

Table A.9: Effects on Voting as Reported in Survey - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Voted (=1)		Non-traditional vote (=1)		Traditional vote (=1)		Blank votes (=1)	
[TF] Facebook Ad Treatment	-0.024	-0.021	0.076**	0.059*	-0.083**	-0.070**	-0.014*	-0.012*
	(0.019)	(0.019)	(0.032)	(0.032)	(0.035)	(0.033)	(0.007)	(0.007)
	[0.248]	[0.320]	[0.020]	[0.092]	[0.008]	[0.050]	[0.006]	[0.020]
[TL] Letter to Politicians	0.011	0.014	-0.006	0.003	0.034	0.024	-0.001	-0.001
	(0.020)	(0.019)	(0.031)	(0.030)	(0.032)	(0.031)	(0.005)	(0.006)
	[0.606]	[0.518]	[0.798]	[0.912]	[0.326]	[0.448]	[0.932]	[0.834]
Control Mean	0.90	0.90	0.65	0.65	0.29	0.29	0.02	0.02
Sample Size	1029	1029	928	928	928	928	928	928
Selected Controls		X		X		X		X

Notes: The outcome in columns 1-2 is an indicator for whether the respondent said he voted in the first round of presidential elections. In columns 3-4 and 5-6 it is indicator for whether the respondent said he voted for non-traditional and traditional candidates, respectively. In columns 7-8 it is an indicator for whether the respondent voted blank. For a discussion on how traditional candidates were defined see the main text. Specifications in even-numbered columns include the covariates selected using the method method described in Chernozhukov et al., 2015 and Belloni et al., 2014. Clustered standard errors at the municipality 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 A.10: Effects on Vote Shares of Each Candidate - 2018 Experiment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Non-traditional candidates				Traditional candidates					
	Petro votes (%)		Fajardo votes (%)		Duque votes (%)		Vargas votes (%)		De la Calle votes (%)	
[TF] Facebook Ad Treatment	3.223**	1.426	0.369	0.668	-2.945**	-0.746	-0.529	-0.905	-0.164	-0.132
	(1.505)	(0.884)	(0.777)	(0.484)	(1.423)	(0.634)	(0.699)	(0.567)	(0.161)	(0.147)
	[0.034]	[0.114]	[0.612]	[0.146]	[0.030]	[0.242]	[0.418]	[0.124]	[0.258]	[0.292]
[TL] Letter to Politicians	-0.835	-1.021	-0.345	0.690	1.336	1.029*	-0.059	-0.566	0.103	0.133
	(1.478)	(0.813)	(0.754)	(0.456)	(1.356)	(0.585)	(0.631)	(0.512)	(0.139)	(0.125)
	[0.602]	[0.224]	[0.624]	[0.130]	[0.332]	[0.110]	[0.920]	[0.282]	[0.464]	[0.284]
Control Mean	24.06	24.06	13.08	13.08	47.49	47.49	10.53	10.53	2.03	2.03
Selected Controls		X		X		X	X	X		

Notes: The outcomes in this table are the vote shares of each of the five main candidates in the first round of presidential elections. Columns 1-4 use the vote share for non-traditional candidates. Columns 5-10 display use the vote share for traditional candidates. Specifications in even-numbered columns include the covariates selected using the method 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 A.11: Effects on Legal Campaigning - Second Experiment (2019)

	(1)	(2)	(3)	(4)	(5)
	Campaigning Index	Saw Public Speeches by Candidate (=1)	Received Fliers About Candidate (=1)	Heard Radio Ads About Candidate (=1)	Received Social Media Ads About Candidate (=1)
[TI] Info Message	0.070 (0.058) [0.030]	0.006 (0.022) [0.374]	0.031 (0.026) [0.042]	0.030 (0.028) [0.256]	0.016 (0.026) [0.176]
[TS] Salience Message	0.050 (0.056) [0.826]	0.026 (0.024) [0.086]	0.005 (0.023) [0.754]	0.006 (0.028) [0.286]	0.018 (0.025) [0.946]
[TB] Info and Salience Messages	0.020 (0.059) [0.016]	0.016 (0.023) [0.506]	0.003 (0.026) [0.460]	0.007 (0.031) [0.004]	0.012 (0.024) [0.038]
[TL] Letter to Politicians	-0.049 (0.047) [0.046]	0.006 (0.018) [0.530]	-0.022 (0.019) [0.040]	-0.015 (0.024) [0.218]	-0.021 (0.020) [0.078]
Control Mean	-0.00	0.34	0.45	0.59	0.65
Sample Size	9615	11375	12248	10952	11904
Test $TI = TS$, p-value	0.70	0.37	0.25	0.37	0.92
Test $TB = TS$, p-value	0.59	0.67	0.94	0.97	0.79
Test $TB = TI$, p-value	0.38	0.64	0.25	0.41	0.88
Test $TL + \frac{TI+TS+TB}{3} = 0$, p-value	0.97	0.21	0.62	0.96	0.76
Selected Controls	X	X	X	X	X

Notes: This table displays the effects of the different interventions performed in the second experiment in 2019 on different metrics of the extent of candidates' legal campaigning, as measured by citizen's self-reported witnessing of different types of campaigning. The outcome in columns 2-5 are indicators for whether survey respondents witnessed each type of campaigning at least once in the week prior to the elections by each candidate. In column 1 the outcome variable is an index composed of the indicators in the following columns. 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.

Table A.12: Effects on Additional Survey Outcomes - 2018 Experiment

	(1)	(2)	(3)	(4)
	Effectiveness of reporting (z-score)		Easiness of reporting (z-score)	
[TF] Facebook Ad Treatment	-0.026 (0.070) [0.702]	-0.032 (0.070) [0.662]	-0.069 (0.068) [0.312]	-0.043 (0.063) [0.466]
[TL] Letter to Politicians	0.110 (0.068) [0.096]	0.086 (0.067) [0.256]	-0.056 (0.064) [0.442]	-0.024 (0.062) [0.672]
Control Mean	-0.00	-0.00	-0.00	-0.00
Sample Size	966	966	966	966
Selected Controls		X		X

Notes: The outcome in columns 1-2 is a measure of how effective the respondent thinks reporting is on a scale of 1 to 7, where higher numbers represent the perception that reports are easier to make, that has been standardized using the mean and standard deviation in the control group. In columns 3-4 it is a measure of how easy the respondent thinks reporting is on a scale of 1 to 7, where higher numbers represent the perception that reporting is easier, and standardized as before. Specifications in even-numbered columns include the covariates selected using the method described in Chernozhukov et al., 2015 and Belloni et al., 2014. Clustered standard errors at the municipality level are shown in parentheses and random inference p-values are shown in square brackets. Two-sided p-values of tests of coefficient equality are shown at the end of each panel; *** p-value < 0.01, ** p-value < 0.05, * p-value < 0.1.