CDDRL

WORKING PAPER

January 2015

VOTING AND PEER EFFECTS: EXPERIMENTAL EVIDENCE FROM MOZAMBIQUE

Marcel Fafchamps
Stanford University

Ana Vaz
University of Oxford

Pedro C. Vicente
Universidade Nova de Lisboa

Stanford | CDDRL Center on Democracy, Development, and the Rule of Law

FSI Freeman Spogli Institute for International Studies

CDDRL

ABOUT CDDRL

Since 2002, the Center on Democracy, Development, and the Rule of Law (CDDRL) at Stanford University has collaborated widely with academics, policymakers and practitioners around the world to advance knowledge about the conditions for and interactions among democracy, broad-based economic development, human rights, and the rule of law.

The mission of CDDRL is to understand how countries can overcome poverty, instability, and abusive rule to become prosperous, just, democratic, and well-governed states. This concern for the overall trajectory of national development – and for the intricate links among the economic, political, legal, social, and health dimensions of development – sets CDDRL apart from other research centers.

Center on Democracy, Development, and the Rule of Law
Freeman Spogli Institute for International Studies
Stanford University
Encina Hall
616 Serra Street
Stanford, CA 94305-6055

Voice: 650-723-4610
Fax: 650-724-2996

cddrl.fsi.stanford.edu

Voting and Peer Effects: Experimental Evidence from Mozambique*

Marcel Fafchamps
Stanford University†

Ana Vaz
University of Oxford‡

Pedro C. Vicente
Universidade Nova de Lisboa§

January 2015

Abstract

Voter education campaigns often aim to increase voter participation and political accountability. Randomized interventions were implemented nationwide during the 2009 Mozambican elections using leaflets, text messaging, and a free newspaper. We study the peer effecs triggered by the campaign within households and villages. We investigate whether treatment effects are transmitted through social networks and geographical proximity at the village level. For individuals personally targeted by the campaign, we estimate the reinforcement effect of proximity to other targeted individuals. For untargeted individuals, we estimate how the campaign diffuses as a function of proximity to targeted individuals. We

*We wish to thank Jenny Aker, Oriana Bandiera, Cátia Batista, Paul Collier, Eliana La Ferrara, Ted Miguel, Gerard Padro-i-Miquel, and Imran Rasul for helpful suggestions. Miguel Ferreira provided superb research assistance. We also wish to thank participants in seminars at LSE/UCL, Stanford, ULB, and UNU-MERIT, and at the MOVE, SITE, NEUDC, and North-American Winter Meeting of the Econometric Society conferences for useful comments. We are particularly grateful to newspaper @Verdade, AMODE, Observatório Eleitoral and STAE for fruitful collaboration. Adérito Caldeira, Sheikh Abdul Carimo, Erik Charas, Joshua Haynes, Thomas Kroner, Felisberto Naife, João Pereira, and Carlos Shenga offered crucial inputs to this project for which we are most thankful. Finally, we would like to extend a word to the fantastic work offered by supervisors Egídio Chaimite, Alberto da Cruz, Egídio Guambe, and Aquílcia Samuel, and the group of enumerators with whom we worked: their dedication to this project was critical to its success. We wish to acknowledge financial support from the International Growth Centre and the Foundation Open Society Institute (Zug). Vicente is grateful for a grant from Trinity College Dublin. All errors are our responsibility.

†Freeman Spogli Institute, Encina Hall, 616 Serra St., Stanford, CA 94305, USA. Email: fafchamp@stanford.edu. Tel: +1-650-4974602.

‡Department of Economics, University of Oxford, Manor Road, Oxford OX1 3UQ, UK. Email: ana.vaz@qeh.ox.ac.uk.

§Nova School of Business and Economics, Campus de Campolide, 1099-032 Lisboa, Portugal. Email: pedro.vicente@novasbe.pt. Tel: +351-21-3801601.

find evidence for both effects, similar across treatments and proximity measures. The treatments raise the level of information and interest in the election through networks, in line with the average treatment effect. However, we find a negative network effect of the treatments on voter participation, even though the average effect of the treatments themselves is positive: the effect of treatment on more central individuals is lower and sometimes negative. We interpret this result as a free riding effect, due to the fact that voter participation is costly.

1. Introduction

The rationality of voter turnout in political elections is often questioned: unless a person casts the deciding vote, voting has no effect on the outcome (e.g., *Feddersen, 2004*). This is particularly true in elections where one contender has widespread support and the outcome is fairly certain. If no one votes, however, the electoral outcome is unlikely to reflect the preferences of the electorate. Not voting is therefore equivalent to free riding on other people's electoral participation. As a consequence, voting is often seen as a civic duty. Although some countries (e.g., Belgium, Brazil, Peru) make voting a legal obligation, most do not. The level of electoral participation therefore depends on the probability voters attribute to being pivotal and on the social norms that are in place regarding voting. Peer influence may affect both.

The purpose of this paper is to study peer effects in political participation. A randomized control trial was organized in Mozambique to study the effect of voter education during the 2009 elections. The study of voter education in developing countries has seen recent attention, as electoral problems like clientelism and vote-buying (*Wantchekon, 2003*; *Fujiwara and Wantchekon, 2013*; *Vicente, 2014*), violence (*Collier and Vicente, 2014*), and low accountability (*Banerjee, Kumar, Pande, and Su, 2011*) have been identified to affect the likelihood that elections translate into public policies that produce broad-based development. Specifically, Mozambique has seen a dramatic decrease on political participation since the first democratic elections in 1994, which has accompanied the consolidation of power of the ruling party.

The voter education we study was implemented in collaboration with a free newspaper and a consortium of local NGOs. Its main objective was increasing electoral participation. Three different treatments were administered nationwide across four provinces of the country. The

first is the distribution of the free newspaper, which focuses on neutral information about the elections. The second is a text messaging hotline to which citizens can report electoral problems. The third is civic education delivered via a leaflet and text messages providing information about the elections. All treatments include an appeal to voter participation in the elections. Treatments were allocated randomly across locations. Within locations, a number of randomly selected individuals were directly targeted by the campaign. We refer to them as 'targeted'. There is only one targeted individual per household. We also follow a randomly selected number of individuals who reside in treated locations but are not directly targeted by the campaign. We refer to them as 'untargeted'. Targeted and untargeted individuals are always the head of household or his/her spouse. In each visited household, we collected information on another adult member who, by construction, is not directly targeted by the campaign. We refer to them as 'secondary interviewees'.

Our focus is on estimating the peer effects of the different treatments within the household and the village. At the household level, we estimate the campaign's indirect effect on secondary interviewees living with targeted respondents. At the village level, following *Fafchamps and Vicente (2013)*, we divide peer effects into reinforcement and diffusion effects. Reinforcement refers to the effect of the campaign for targeted individuals who are socially or geographically proximate to other targeted individuals. Diffusion refers to the effect of the treatment for untargeted individuals in treated locations who are close to targeted individuals. *Aker, Collier, and Vicente (2013)* study the direct average treatment effect of the voter education campaign we analyze in this paper. Their results are briefly summarized here to enable comparability with peer effects.

In terms of outcomes variables, we exploit a rich individual dataset that includes survey measures of individual turnout, a behavioral measure of political participation, and measures of information and interest in politics. We also report average treatment effects using official voting records at the polling station level. To estimate reinforcement and diffusion effects, we use detailed measures of social and geographic connectedness between individuals, including measures of chatting, kinship and geographical distance between respondents' houses.

All treatments increase voter turnout at the polling station level, as given by official records. Survey measurements show that turnout increases both among targeted and untargeted individuals. The hotline and the civic education treatments increase the political participation of secondary interviewees. We also document a clear increase in information about the elections among targeted and untargeted individuals, and a weaker effect among secondary interviewees.

Reinforcement and diffusion effects on voter participation are, however, quite different from average treatment effects as they are all negative. This holds for different measures of connectedness, and for both voter turnout and our behavioral measure of political participation. Negative reinforcement and diffusion effects on turnout are particularly strong for the hotline treatment. In contrast, the peer effects on information and interest in politics are positive – and in line with the average effects of the campaign.

We interpret these findings as consistent with a general model of costly political participation. In this framework, voter participation may be induced either by the probability of affecting the electoral process, or by non-instrumental motivations like civic-mindedness. By giving information about the credibility of the elections, the campaign intends to reassure voters about the integrity of the process. So doing, it may also raise civic-mindedness. Both effects are conducive to increased turnout, in line with the average effects of the campaign. However, peer effects can induce free-riding in turnout if more central voters realize that, because of the campaign, turnout will increase and their vote becomes less essential in achieving a politically acceptable turnout rate.

Our estimation of network effects in the context of a randomized field experiment relates to a recent body of work on the role of networks in aid interventions. *Miguel and Kremer (2004)* launched this literature by estimating externalities of a deworming school-based program in Kenya. They estimated the impact of the treatment on control populations. Because their experimental design features program randomization at the school level, it does not allow for an experimental estimation of externalities within treated schools. More recently, *Angelucci and De Giorgi (2009)* extend the study of externalities to a conditional cash transfer program. By exploring a rich set of outcomes at the household level they are able to draw some light into

specific mechanisms of influence of unexposed households. However, these authors do not use explicit network variables. Still in the context of a conditional cash transfer program, *Angelucci, De Giorgi, Rangel, and Rasul (2010)* introduce explicit interaction between households but focus on kinship links. Our analysis of kinship as a measure of social interaction is also related to *Bandiera and Rasul (2006)* who study technology adoption in Mozambique in a non-experimental setting. *Baird, Bohren, McIntosh and Ozler (2014)* study the design of experiments intended specifically to analyze spillover effects.

The experimental literature on voter mobilization was initiated by studies by Gerber and Green. For instance, *Gerber and Green (2000)* studied the impact of a leaflet get-out-the-vote campaign in the U.S. *Dale and Strauss (2009)* introduce text messaging in American get-out-the-vote campaigns and provide evidence that SMS reminders increased the likelihood of voting. The studies by *Nickerson (2008)*, *Fafchamps and Vicente (2013)*, and *Gine and Mansuri (2011)* relate closely to our paper as they analyze peer effects of voter mobilization interventions. The first looks at a door-to-door get-out-the-vote campaign in the U.S. to identify peer-effects in two-member households. The second follows a campaign against political violence in Nigeria to identify reinforcement and diffusion network effects. The third assesses the impact of a voter awareness campaign on female turnout in Pakistan in which peer effects are estimated using geographical distance and friendship. Our paper is, to the best of our knowledge, the first to contrast household and village peer effects of mobilization.

The paper is organized as follows. In Section 2 we describe our testing strategy. We then offer the context of our experiment in Section 3. The treatments are introduced in detail in Section 4. Subsequently, in Section 5 we describe the data including outcome and network variables. In Section 6 we present our empirical results, including balance tests, average effects, peer effects, and robustness. In Section 7 we provide interpretation, including conceptual framework and discussion of alternative explanations for our results. Section 8 concludes.

2. Testing strategy

The combined (i.e., direct and indirect) average effects of the campaign can be estimated as follows. Let y_i be a measure of electoral behavior, information, or interest for individual i. Let $T_i = 1$ if individual i was targeted by the campaign, and 0 otherwise. As we will see when the experiment is described in detail, the campaign takes three distinct forms that we test separately. For the sake of the presentation, here we focus on a single treatment.

Assuming treatment is randomly assigned, the homogeneous (average) effect of the campaign on treated individuals can be estimated using targeted and control observations in a regression of the form:

$$y_i = \alpha + \beta T_i + \varepsilon_i. \tag{2.1}$$

Coefficient β is the average treatment effect on electoral behavior, information, or interest. This regression can also be estimated with village and individual controls. Whenever comparable information about y_i is available at different points in time, a difference-in-differences version can also be estimated in which individual fixed effects net out possible individual unobservables.

We can also estimate the average effect of the campaign on individuals in treated locations who were randomly selected not to be targeted by the campaign. We estimate this average effect using untargeted and control observations in a regression of the form:

$$y_i = \alpha^u + \beta^u T_v + \varepsilon_i \tag{2.2}$$

where $T_v = 1$ if the village was treated. The individual treatment variable T_i drops out since, by design, it is 0 for untargeted individuals. Coefficient β^u is an estimate of the average diffusion effect of the campaign on the electoral behavior, information, or interest of untargeted individuals. Estimations of (2.1) and (2.2) are covered in detail in *Aker, Collier, and Vicente (2013)*.

It is also possible to estimate the average effect of the campaign on individuals who live with targeted subjects. We estimate this average effect using secondary interviewees living in

targeted and control households using a regression of the form:

$$y_i = \alpha^s + \beta^s T_h + \varepsilon_i$$

where $T_h = 1$ if someone else in the household was targeted by the campaign. Coefficient β^s is an estimate of the average diffusion effect of the campaign on the electoral behavior, information, or interest of individuals who cohabit with targeted subjects.

A possible configuration of interest is when $\beta^s = 0$ for electoral information and interest, but $\beta^s > 0$ for electoral behavior. This would suggest that treated individuals – who are the head of household or his spouse – put pressure on dependents to vote, without necessarily providing them with information or without convincing them that voting is a civic duty (*Vaz, 2013).*

Our approach relating to the estimation of peer effects within the village builds on the work of *Fafchamps and Vicente (2013)* who analyze the effect of a campaign against electoral violence in Nigeria. We investigate whether peer effects are stronger for targeted individuals who are socially or geographically close to other targeted individuals. Social proximity is captured by letting $g_{ij} = 1$ if individuals i and j are connected in a relevant social network sense, and 0 otherwise. Geographical proximity is captured by letting g_{ij} be the physical distance between i and j. We estimate a heterogeneous reinforcement effect model of the form:

$$y_i = \alpha + \beta T_i + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \gamma T_i \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \varepsilon_i \tag{2.3}$$

where N is the number of observations in i's village, $V_j = 1$ if j is a targeted individual (i.e., living in a treated location) or lives in a control location, with the value 0 assigned otherwise, and $\frac{1}{N} \sum_{j \neq i} g_{ij} V_j$ is the proportion of targeted or control neighbors to whom i is connected (depending on whether i is in a treated or a control village, respectively). The coefficient of interest is γ.[1] Network reinforcement effects are tested by examining whether the effect of treatment is larger among individuals with more direct links to targeted individuals.[2] Regressor

[1] When estimating the above regression, $\frac{1}{N} \sum_{j \neq i} g_{ij} V_j$ is expressed in difference relative to the mean in the expression $\frac{1}{N} T_i \sum_{j \neq i} g_{ij} V_j$ so as to keep the interpretation of the β's unaffected.

[2] In the event that $\gamma = 0$, we cannot rule out the possibility that social network effects are so strong as to spread evenly to all individuals in treated villages, in which case proximity to treated individuals does not matter.

$\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ is included to control for the possibility that better connected individuals are more likely to vote even in the absence of treatment. Regression (2.3) is estimated in levels using control and targeted individuals only, i.e., excluding untargeted individuals living in treated areas.

One possibility of interest is when $\gamma < 0$ while $\beta > 0$ for turnout. One possible explanation for such configuration is free-riding: treatment raises the likelihood that others vote; this in turn reduces the marginal usefulness of i's vote in achieving a village turnout that is politically or socially acceptable; if i is better connected, i is better able to observe the effect of the campaign on others' intention to vote, and thus i is more aware of the reduced usefulness of his/her vote. We revisit this point more in detail later.

We can investigate the presence of heterogeneous diffusion effects on the untargeted using the same approach:

$$y_i = \alpha^u + \beta^u T_v + \delta^u \frac{1}{N}\sum_{j\neq i} g_{ij}V_j + \gamma^u T_v \frac{1}{N}\sum_{j\neq i} g_{ij}V_j + \varepsilon_i. \tag{2.4}$$

Interpretation is similar to that of heterogeneous reinforcement effects.

We use ordinary least squares in all our main regressions. Since the data we use is clustered by enumeration area (EA), we allow for within-group dependence by clustering standard errors at the EA level.

3. Context

Mozambique, a country with 22.4 million inhabitants, is one of the poorest countries in the world with GDP per capita of 838 USD in 2008 - it ranks 161st in 189 countries (based on latest available years) in terms of GDP per capita. Without prominent natural resources, and with 81 percent of the population directly dependent on agriculture, it is an aid-dependent country with official aid assistance accounting for 22 percent of GNI in 2008.[3]

Mozambique became independent from Portugal in 1975, after which FRELIMO (Frente de Libertação de Moçambique), the independence movement, led a single-party, socialist regime. Beginning in 1977, Mozambique suffered a devastating civil war, fought between FRELIMO

[3] These figures were taken from World Development Indicators, 2009, and CIA World Factbook, 2010.

and RENAMO (Resistência Nacional Moçambicana). RENAMO was supported by Apartheid South Africa and, in the context of the cold war, by the U.S. The civil war finished in 1992 with an agreement to hold multi-party elections. Presidential and parliamentary elections were held in Mozambique in 1994, 1999, 2004, and 2009. FRELIMO and its sponsored presidential candidates won all national elections, with RENAMO as the main contender. More importantly, FRELIMO has been consistently increasing its vote share, while voter turnout has decreased massively from 88 percent in 1994 to just 36 percent in 2004.

Armando Guebuza became FRELIMO's leader and president in 2004, succeeding Joaquim Chissano. Guebuza is a historical figure in FRELIMO. He fought against the Portuguese and was minister of the interior under Samora Machel. He became a wealthy and powerful businessman after the privatization of public companies in the 90s. In the 2009 election that we study he was running for re-election as president of the country. His main opponent, Afonso Dhlakama has been the leader of RENAMO since 1984. He served as guerilla leader during the civil war, and has been RENAMO's presidential candidate at all national elections.

In this paper we focus on the presidential, parliamentary and provincial assembly elections of October 28, 2009. The 2009 elections were relatively calm, with FRELIMO and Guebuza expected to win. The elections were generally unproblematic, with national and international observers considering that the 2009 election followed appropriate international standards, despite many small irregularities. Electoral results attributed 75 percent of the vote to Guebuza at the presidential elections and to FRELIMO at the parliamentary elections.

4. Treatments

The data used in this paper come from a randomized control trial implemented in Mozambique around the time of the 2009 elections. Three treatments are investigated, all geared towards encouraging people to vote. The first treatment is the distribution of an independent newspaper providing electoral information; the second is a campaign to encourage voters to use an SMS-based hotline set up to report electoral problems; and the third is a civic education campaign, which gave information about the election and focused on participation in the election.

The three interventions were designed and conducted with the institutional support and active collaboration of newspaper @Verdade (http://www.verdade.co.mz/) and a consortium of eight Mozambican NGOs, named Observatorio Eleitoral. For more details on these organizations, see *Aker, Collier, and Vicente (2013)*.

Voter education campaigns generally combine one or more of three elements: (i) information – providing information about the electoral process; (ii) nudging – repeatedly reminding people to vote;[4] and (iii) participation – offering voters the opportunity to circulate their observations about the electoral process.

The newspaper treatment combines all three elements, i.e., information, nudging, and participation. This treatment was organized around the distribution of a free newspaper, *@Verdade*, to experimental subjects in selected locations. None of the treated locations had received the newspaper before.[5] The editors of the newspaper took a strictly independent approach to the electoral process, focusing their message on electoral education. The newspaper was distributed for the purpose of the research in the experimental locations from the baseline survey in September 2009 until the post-election survey in November 2009. Over this period, the newspaper included information designed and made available by the electoral commission (CNE/STAE). This information focused on the voting steps on the election day (see middle panel of *Figure A1* in the Online Appendix). The newspaper also advertised a national hotline for reporting electoral problems (see right panel of *Figure A1*). For the distribution of the newspapers to treated villages, priority was given to targeted respondents. 5,000 copies of the newspaper were distributed each week, with a total of 125 for each location.

The hotline treatment emphasizes the information and participation elements. This treatment was organized around the setting-up of two short-code phone numbers contracted with the cell phone operators in Mozambique (Mcel and Vodacom). These short-codes constituted an SMS hotline as they were prepared to receive text messages reporting electoral problems.

[4] See *Dale and Strauss (2009)* for an example of the effects of text messaging nudges on voter turnout in 2006 American elections. The effectiveness of nudging in other fields has also been documented (*Thaler and Sunstein, 2008*; *Pop-Eleches et al., 2011*).

[5] Despite being the highest circulation newspaper in Mozambique (with a minimum of 50,000 certified copies per week), the newspaper was only systematically distributed in the city of Maputo. As all newspaper locations lie outside the city of Maputo, they had never received the newspaper.

Note that this hotline was branded with a different slogan and different short-codes from the newspaper hotline. During the baseline survey, we conducted a door-to-door campaign providing information on the hotline: we distributed 10,000 leaflets (250 per location) primarily directed at targeted respondents, providing basic information about the hotline, i.e., short-codes, examples, format of the reports to be sent,[6] and the name of the sponsors. The leaflet is depicted in *Figure A2*. We promised that the contents of reports would be passed to the media for dissemination, and shared via SMS with all other targeted respondents in hotline treatment locations. Before dissemination each report received on the hotline was verified with local correspondents that were hired in each of the hotline treatment locations. In addition to receiving hotline reports, respondents in hotline areas were sent daily SMS reminders about the existence of the hotline from two weeks prior to the elections until the election day.

The civic education treatment combines information and nudging elements. This treatment was organized around a set of messages providing citizens in selected locations with specific information about the 2009 elections. The intervention started with a door-to-door campaign approximately a month before the elections. This was implemented during the baseline survey and was centered on the distribution of an extended version of the information provided by CNE/STAE through the newspaper. It took the form of a leaflet designed and produced by CNE/STAE. A copy of the leaflet is displayed in *Figure A3*. It explains in detail the voting steps on the election day. 10,000 leaflets were distributed (i.e. 250 per location) primarily to targeted respondents. Moreover, for two weeks prior to the election, subjects in the civic education treatment received five daily text messages on the cell phone number they provided during the baseline survey. The messages focused on the importance of voter participation, as in a 'get-out-the-vote' campaign. Within their 160-character limit, these messages also provide specific information about the electoral process, namely: the scheduled date; the type of elections taking place; the presidential candidates; the parties running for parliament; voter confidentiality; and how to vote.

Given that all three treatments contain an information element, we expect them to all have a positive effect on turnout if lack of information about the electoral process is what turns away

[6] Specifically, ballot location name first, and description of the problem second.

voters. The civic information treatment has the strongest nudging component. If this treatment has a particularly strong effect on turnout, it suggests that nudging can effectively induce people to vote. The hotline has the strongest participation element. A large treatment effect would suggest that turnout can be increased by encouraging voter participation in the electoral process.

5. Data

The project took place in four provinces, Cabo Delgado, Zambezia, Gaza, and Maputo-Province. The sampling base is the 2004 electoral map of the country, and the enumeration area or EA is the area covered by a polling station. Because the use of cell phones is central to all our treatments, we eliminate from the sampling base all polling locations without cell phone coverage. For this purpose, we obtained detailed data from the two cell phone operators on the geographic location of each of their antennae. These were then plotted on a map using their geographical coordinates, with a five-km coverage radius drawn for each. All polling locations outside the covered area were dropped from the sampling base. In 2009, 60 percent of all ballot locations in the country are found to be covered by at least one operator.

From this sampling base, 161 polling locations were selected using two-stage clustered representative sampling – first on provinces, then on EAs. The number of registered voters per polling location is used as sampling weight. Since all registered voters in the sampling frame have the same probability of being sampled, the 161 locations are nationally representative of the voting population of Mozambique that has mobile phone coverage. Of the 161 polling locations selected for our study, 40 were randomly assigned to each of the three treatments, and 41 locations serve as control group, with no treatment administered. The allocation of locations to treatments and control follows a stratified randomization procedure (*Bruhn and McKenzie, 2009*). First, clusters of four similar locations were formed in each province, with similarity based on geography. Within each cluster, locations were then randomly assigned to one of the three treatments or to control. During the baseline survey, in the event that we found no cell phone coverage in a selected location, we replaced it by the closest polling location with cell phone coverage. That happened in seven locations.

In each of the EAs we conducted two face-to-face household surveys, one before the election and treatment, and one after. Sampling in each EA followed standard procedures for household representativeness (n'th house call by enumerators starting from the center of the EA, the polling location, typically a school). Interviews at baseline were aimed at the household head or his/her spouse. Interview and subsequent treatment are conditional on 'having access to a cell phone' for receiving and sending calls and messages. This criterion includes respondents that do not own a cell phone but have access to one via a neighbor or family member nearby. The baseline survey includes 1,766 households/respondents, approximately 11 per EA. It took place from mid-September to mid-October 2009.

In treated EAs, individuals interviewed at baseline were randomly assigned to be targeted or untargeted as follows. Of the average 11 baseline households interviewed in each treated EA, two were, on average, randomly selected not to receive the treatment themselves. They are called 'untargeted'. The other nine were directly targeted for treatment as described in the previous section. This randomization was implemented specifically to study diffusion effects among individuals in treated locations not directly targeted for treatment.

The post-election survey started after the election results were announced in early November. It lasted for about the same duration as the baseline survey. We attempted to re-interview all baseline respondents, and reached 1,154 of them. To check that our results are not an artifact of selective attrition in the post-election survey, we verify, in the next section, whether observable characteristics vary systematically across treatments. More importantly, we also reestimate our main results using a multiple imputation technique to account for missing post-election observations.

In the post-election survey we also interviewed a new sample of respondents aimed to be composed by one additional adult per baseline household. We refer to these individuals as 'secondary interviewees'. Since only one person per targeted household was treated, secondary interviewees were, by construction, not directly targeted by the campaign. The selection of secondary interviewees alternated randomly between the main respondent's spouse, son, daughter, other male members and other female members of the household. Because of work migrations,

not all households include another adult member at the time of the survey.[7] The post-election survey covers 518 secondary interviewees.

5.1. Outcome variables

The outcomes of interest in this paper come mainly from survey and behavioral data collected at the individual level. *Table A1* in the Online Appendix presents a summary of the survey outcome variables. These variables have been grouped into three sets: participation (turnout), information, and interest relating to the elections. We also report on official voting results at the level of the polling station.

We were particularly careful with our measurement of voter turnout. We propose six turnout measures. The first one is self-reported turnout. The second is self-reported turnout adjusted by considering as non-voters those who did not answer correctly questions regarding ballot papers and boxes.[8] The third one is an indicator of whether the respondent showed without hesitation his/her inked finger to the enumerator – dipping one finger in indelible ink was the method used to prevent people from voting multiple times. Turnout index 1 is a composite index measuring how well the respondent answered questions on the sequence of events during the election day. The answer to each question is coded according to how convincing the response is. Turnout index 2 is based on the sub-group of these questions that focuses on knowledge about the polling station (e.g., the number of ballot papers, whether there were photos of the candidates, the number of ballot boxes, whether they were transparent, and whether they were colored). The last measure of turnout is a final enumerator assessment on whether the respondent voted or not. The three last measures take values between 0 and 7 and are thus potentially most informative. To facilitate comparison with the other turnout measures, we normalize them by

[7] The post-election survey took place during the rainy season when most agricultural work ("machambas") occurs. As agricultural plots tend to be located at a fair distance from home (*Sheldon, 1999; De Vletter, 2001*), agricultural workers often migrate during this season. In the survey, the most commonly reported reason for the absence of an adult dependent is agricultural work. Non-farm work and travelling are also frequently reported as reasons for absence in the Maputo province, probably because it is more urbanized and offers more non-farm employment opportunities (*Cungara et al., 2011*).

[8] According to the adjusted turnout, those respondents who have reported to have voted but answered wrongly the questions regarding the number of ballot papers and ballot boxes were considered as not having voted and, thus, assigned a zero.

dividing by 7, so they too range from 0 to 1.[9]

Our proxies of information and interest come, respectively, from questions asking respondents to list presidential candidates and parties running for the 2009 elections, and from questions asking about the interest respondents had in the presidential election, parliamentary election, provincial assemblies' election, and in public matters generally. The latter questions employed a subjective scale. To facilitate analysis and interpretation, we combine all these questions into two indices: one for basic information about the elections; and the other for interest in the political process. The indices are constructed following the approach of *Kling, Liebman and Katz (2007)*: we normalize the survey-indicators using z-scores and we aggregate them using equally weighted averages of the normalized individual variables. The z-scores are calculated by subtracting the control group mean and dividing by the control group standard deviation. Thus, each component of the index has mean 0 and standard deviation 1 for the control group.[10]

A behavioral measure of demand for political accountability, which we refer to as the 'open letter', is obtained as follows. During the post-election survey the enumeration team explained and distributed a leaflet to all survey respondents in all 161 experimental locations, which invited them to send SMS messages proposing policy priorities to the president-elect for his new mandate. We were clear in conveying the limited extent of the initiative (a small number of experimental localities in the whole of Mozambique), and promised that the contents of these messages would reach the president in person (through the newspaper @Verdade). As with the hotline, each message sent by experimental subjects had a small monetary cost. Sending the message therefore represents a costly action. It was observable to us, as all cell phone numbers that sent messages were recorded and matched with those of the experimental subjects. We interpret the sending of an open letter message as an incentive compatible measure of demand for political accountability. The leaflet is depicted in *Figure A4*.

Official voting results at the level of the polling station were made available by the electoral commission of Mozambique. Polling stations are matched with the enumeration areas in our

[9] The correlation between the different turnout measures ranges from 0.50, between the adjusted turnout and the finger measure, to 0.98, between the self-reported turnout and index 1.

[10] Like in *Kling, Liebman, and Katz (2007)*, if an individual has a valid response to at least one component measure of an index, then we impute any missing values for other component measures at the random assignment group mean for the corresponding time period.

experiment which, as mentioned earlier, were based on the polling stations themselves. We employ results for the presidential and parliamentary elections of 2009.

5.2. Network variables

We collected three measures of social and geographical centrality. The first two variables are centrality measures based on social networks.[11] For the first one, which we call 'chatting', a link from i to j exists if i can identify the name of j when prompted, and i stated that he/she talks to j on a regular basis.[12] For the second, which we call 'kinship', a link from i to j exists if i can identify j by name and reports being related to j.[13]

The third variable is a measure of geographical centrality calculated as the average distance to other sampled individuals in the same EA. Each enumerator was asked to locate each respondent on an approximate EA map, and to calculate the distance between interviews. See *Figure A5* for an example. To evaluate the position of each respondent on the map, we construct up-down and left-right coordinates for each of them. The distance between each ij pair is then calculated from these coordinates. Because maps differ in scale, distances are re-scaled to make them comparable across all locations.[14]

6. Empirical results

We start by checking balancedness by treatment on the baseline data. We then summarize the average treatment effects, some of which (though not all) already appear in *Aker, Collier and Vicente (2013)*. We then present our main results in detail.

[11] Because we only observe a fraction of the chatting and kinship networks, we refrain from using other measures of centrality (e.g., Bonacich centrality) that are more sensitive to sample truncation bias (*Chandrasekhar and Lewis, 2012*).

[12] The question asked was 'How frequently do you calmly chat about the day events with the following individuals or members of their households? Not at all, sometimes, or frequently'. We considered a link existed when the individual answered 'sometimes' or 'frequently'.

[13] The exact question used was 'Are the following individuals relatives of yours, i.e. members of your family? Yes-No'.

[14] This is accomplished by using the subset of pairwise distances, i.e., distance between interviews, reported by enumerators.

6.1. Balance

Tables A2 in the Online Appendix present descriptive statistics on demographic traits of the baseline and post-elections samples together with balance tests. Comparisons between treatment and control locations show that the samples are overall balanced. Regarding the sample of targeted respondents at the baseline, only three demographic characteristics are significantly different at the 10 percent level. In the sample of untargeted individuals the number of significant differences is reduced to two. The comparison between control and treated EAs in the follow-up survey is presented in *Table A2d*. We see a similar pattern: in both samples of targeted and untargeted respondents, most household demographics and EA characteristics are not significantly different. Panel attrition seems to have maintained comparability between the treatment groups in terms of observables. Looking at the sample of secondary interviewees, we also find very few differences between those living in treated and control areas (see *Table A2c*).

Social and geographical centrality variables are summarized in *Table A3*. The social centrality variables, chatting and kinship, were collected during the post-election survey and so we only display statistics for the post-election sample.[15] We report average connectedness $\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$, as defined above. Geographical distance is the inverse of connectedness: more central individuals are less distant from other villagers. We do not observe any statistically significant differences across comparison groups.

Finally, we display averages for baseline voting variables at the polling station level. These are voting records from the presidential and parliamentary elections of 2004. Results are presented in *Table A4*. We do not observe any statistically significant differences across comparison groups. Voting variables from the baseline survey are explored in full detail in *Aker, Collier, and Vicente (2013)*: since treated respondents were asked questions on politics after receiving leaflets (for civic education and the hotline) and the newspaper, there may be differences between comparison groups for the targeted due to a first reaction or conformity bias. However, no clear evidence of

[15] As this information could only be collected during the post-elections survey, it raises the concern that the treatments may have affected the networks. The network measure chatting is the most vulnerable to this critique, as it is possible that the interventions trigger conversations between people that ordinarily would not chat. The network measures kinship and geographical distance are much less likely to be susceptible to variations due to the interventions.

such effects is apparent in the data.

6.2. Average effects

We start by summarizing the regression results of the average effects of the campaign in *Aker, Collier, and Vicente (2013)*. We begin with measures of political participation, which is the main intended effect of the campaign. *Table 1* presents the average effects of the three treatments on voter turnout, measured by the inked finger and index 2, and the sending of the open letter, among targeted and untargeted individuals. The average effects on the other turnout measures are presented in *Table A5*. Since this information can only be collected in the post-election survey, all regressions are estimated using post-election data only. For each measure we present one regression with province dummies, and another adding location and individual demographic controls.

We first look at reported values of outcome variables for control individuals. Since self-reported turnout is larger than all other turnout estimates, it suggests that respondents over-report having voted. This is consistent with respondents regarding voting as a civic duty: if respondents see voting as a perilous or controversial activity, we would expect the opposite finding. Still, the lowest turnout measure still puts average turnout close to 70 percent among control respondents. The inked finger measure shows a turnout of 81 percent on average among control individuals. Since panel respondents are either the household head or their spouse, it is probably not surprising that average turnout among them is above the 2009 national average of 44 percent. 15.3 percent of control individuals sent an open letter SMS to the president.

The average effect of the hotline on the turnout of targeted individuals is significant for all six turnout measures and varies between 5 and 14 percentage points, depending on the turnout measure. This is a large effect given the participation rate is already high among control respondents. The civic education treatment is associated with a moderately large increase in turnout; the effect is significant or marginally significant in most cases when we use no controls. For the newspaper treatment we find a small positive effect when we use no controls (i.e., +3.3 percentage points on average across the six measures), but this effect is never statistically

significant. Regarding the open letter, we find a significant positive effect of the newspaper treatment. On average, this treatment increases the probability of sending a SMS proposing policy priorities in 9.9 percentage points (when using controls).

Turning to the average diffusion effect of the campaign on untargeted individuals, we find statistically significant effects for two treatments, newspaper and hotline. Averaged over the six measures (no controls), we find 10.9 and 9.5 percentage point increases in turnout among untargeted households for the newspaper and hotline treatments, respectively. The treatment effect is also large for the civic education treatment, but it is only statistically significant when using the adjusted turnout and no controls. The result on the newspaper is surprising given that the newspaper treatment has no statistically significant effect on turnout among the targeted. This suggests that the diffusion effect of the newspaper treatment is stronger than its direct effect, possibly because of magnification of the importance of the newspaper among individuals without direct access to it. In contrast, the civic education treatment increases turnout among the targeted but not among the untargeted – a finding that suggest that the nudging effect of the treatment is not magnified through social interaction. We find positive average diffusion effects on the open letter for all treatments. However, none of these effects is significant.

We now turn to results that complement those in *Aker, Collier, and Vicente (2013).* We first look at the effects of the treatments on information and interest in politics. These results are displayed in *Table 2.* We find significant positive effects of the hotline on information about the elections. This effect ranges between 0.16 and 0.21 standard deviation units for the targeted; for the untargeted it is equal to 0.24 standard deviation units (in the regression with controls). The newspaper also has a positive impact on information, which is significant for the targeted (0.17 standard deviation units). Neither the hotline nor the newspaper have a clear impact on interest in politics. The civic education treatment does not have any statistically significant effects on any of these outcomes, even though the sign of the effect on the information index is consistently positive.

Now we turn to the diffusion effects of the campaign within households. *Table 3* presents the average diffusion effects of the three treatments on political participation of secondary inter-

viewees living with targeted individuals. In this table we only include two measures of turnout, inked finger and index 2. *Table A6* includes the regressions of the other turnout measures. For each measure we present two sets of estimates: the first has no controls; the second includes provincial dummies and individual characteristics. Within households of targeted individuals, we find statistically significant diffusion effects for hotline and civic education treatments. The civic education has a positive effect on all voter turnout measures, 8.9 percentage points on average across the six measures (when using controls). This effect is significant for almost all turnout measures when we use controls. The hotline has a positive significant effect on the inked finger measure when using no controls and on the interviewer assessment measure. This treatment also has a positive and significant effect on probability of sending an SMS proposing policy priorities.

The intra-household diffusion effects on information and interest in politics are reported in *Table 4*. These results focus on secondary interviewees living with targeted respondents. Within households of targeted respondents, none of the interventions has an effect on secondary interviewees' electoral information. This is in contrast with the direct effects of these treatments on targeted respondents, which are positive. On the other hand, the hotline has a positive and significant effect on interest in politics among adult dependents (0.21 standard deviation units in the regressions with controls).

Finally, we summarize in *Table 5* the average effect of the treatments on actual electoral outcomes from official polling station records (see *Aker, Collier, and Vicente, 2013*, Tables 3a and 3b). All treatments have a strong and significant positive effect on voter turnout. This effect ranges between 5.1 and 5.5 percentage points for the presidential elections, and between 5.2 and 5.7 percentage points for the parliamentary elections. On voting patterns, we observe a positive point estimate of all treatments on voting for the incumbent president and party (FRELIMO) and a negative effect of all treatments on voting for the main challenger candidate and party (RENAMO). However, only civic education is statistically significant in every case. This treatment leads to an increase in the score of FRELIMO and the incumbent president by 3.9 and 4.6 percentage points, respectively. It also reduces the share of votes of RENAMO

and its presidential candidate by 3.8 and 3.2 percentage points, respectively. To summarize, the treatments increased voter turnout, benefited the incumbent, and hurt the chances of the challenger.

6.3. Peer effects on political participation

Next we turn to network effects on political participation, i.e., voter turnout and open letter. In *Table 6* we show our regressions relating to the inked finger measure of voter turnout. We employ the three centrality variables presented earlier: chatting, kinship, and geographic proximity. We start by estimating network reinforcement effects with regression (2.3). Results are displayed in columns (1)-(3). We then estimate network diffusion effects with regression (2.4). Results are displayed in columns (4)-(6). All regressions are estimated using data from the follow-up survey only. We control for provincial dummies, EA characteristics, and individual characteristics. The main focus is the coefficient of $\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ and of the interaction terms, $T_i\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ (reinforcement) and $T_v\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ (diffusion).

We obtain strong positive coefficients for connectedness, particularly when using chatting and kinship as centrality measure, but also when using geographical proximity.[16] This implies that, without treatment, individuals who are more central in their community are more likely to vote. A possible reason is that these individuals feel more pressure to perform their voting duty. From our data, we cannot tell whether centrality causes people to be more civic-minded – e.g., because of social pressure or internalized norms – or whether more civic-minded people become more central – e.g., because they are more sociable.

Turning to the interaction terms, we find negative coefficients for most treatments and centrality measures – though not all are statistically significant. Reinforcement through geographical proximity is consistently negative and statistically significant for all three treatments. A negative and statistically significant effect is also present for diffusion through chatting for the hotline treatment. To get a sense of the magnitude of these peer effects, we present at the bottom of *Table 6* a simple calculation of the difference in predicted voting between a treated

[16] With respect to the variable chatting, it is possible to argue that our testing strategy may be identifying heterogeneous effects by gregariousness instead of network effects. However, given the similar pattern of results for the different network measures, we believe that we are capturing peer effects associated with chatting.

subject with $g_{ij} = 0$ and a treated subject with the average value of network variable g_{ij}. For chatting and kinship, this comparison is between a treated subject with no connections and one with the average number of connections. A negative value represents the reduction in the effect of treatment on the probability of voting that is associated with social proximity. For geographical proximity, the comparison is between a hypothetical treated subjects at distance 0 to others in the village, and a treated subject at the average distance. A positive value indicates that moving from being completely central to being at the average distance from others increases the probability of voting from treatment. The p-values are the same as those for the interaction coefficients. We see that the magnitude of network effects is large, particularly for reinforcement through geographical proximity: relative to a hypothetical centrally located subject, the effect of treatment on the probability of voting for the average subject is 12.5, 12.8 or 10.8 percentage points lower higher, depending on treatment. For diffusion of the hotline treatment, we see that an untargeted subject in a treated village with no network connections is 9.6 percentage point more likely to vote than an untargeted subject with the average level of chatting connections. These are very large figures relative to the direct effect of treatment itself. They show that the ATE hides large variation in treatment effect depending on geographical and network proximity, with more central individuals experiencing a much smaller – and often negative – effect of treatment on their propensity to vote.

Table 7 shows the interaction coefficients for the remaining voter turnout measures. Significant effects are all negative. The hotline stands out as inducing most network reinforcement and diffusion effects. For self-reported turnout and for interviewer assessed turnout, the newspaper induces negative network reinforcement and diffusion through geographical proximity, and civic education induces negative reinforcement effects through kinship and geographical proximity.[17] In Appendix *Table A7* we perform the same counterfactual calculations as we did at the bottom of *Table 6*. These calculations confirm that the magnitude of the network effects is far from negligible, especially given the relatively small average effect of each treatment.

[17] We also estimated average treatment effects for the samples of targeted and untargeted individuals split into the 40 percent above the mean centrality and the 60 percent below the mean centrality. The findings described here are confirmed. Similar – if not stronger – effects are found if we combine baseline and follow-up survey data and apply individual fixed effects.

One of the possible explanations for these negative network effects, as we detail in the final section of this paper, is free-riding: more central individuals are in a better position to realize, due to their centrality in the local network, that others are more likely to vote because of the campaign. They may also realize that the gap between the incumbent and other candidates is likely to increase. Hence their own electoral participation is less necessary, and the likelihood that they turn out to vote decreases.

Results for the open letter are displayed in *Table 8*. We find network reinforcement and diffusion effects that are negative and statistically significant (at the 1 and 5 percent levels) for the civic education treatment when using kinship as measure of social proximity. Here too the magnitude of these effects is large relative to the ATE. For instance, a subject targeted by the civic education treatment is 4.3 percentage point less likely to vote if he/she has the average kinship network than if he/she has no kinship network at all. For untargeted subjects, the difference is 11.2 percentage points. The explanation may be the same as for voter turnout: sending an open letter is a costly action that potentially suffers from free-riding.

We now summarize our results on political participation. We find a strong effect of the hotline treatment on turnout among targeted and untargeted individuals. The effect is largest among individuals who are less central in their community, either socially or geographically. This effect holds whether we consider targeted or untargeted individuals. We also find a clear effect of the newspaper treatment on turnout among untargeted individuals, suggesting a magnification effect through indirect treatment. A stronger effect of the newspaper on less central individuals (geographically) is also identified (for both reinforcement and diffusion). In contrast, the civic education treatment only affects turnout among targeted individuals. We do, however, find a stronger effect of this treatment on targeted individuals who are less central socially or geographically. Regarding the open letter, although we find a significant effect of the newspaper on targeted individuals, we do not find any other effects of the treatments on targeted or untargeted individuals. Still, we find that less central individuals, in the social network sense, are more likely to send an open letter. Overall we find positive direct effects of the treatments on political participation, but negative network effects of similar magnitude.

6.4. Peer effects on information and interest in politics

We now seek to identify the channels through which the treatments affected political participation. We have already noted that the treatments had a direct positive effect on information about the election among targeted and untargeted individuals. However, we could not find statistically significant effects for interest in politics. We now examine network diffusion and reinforcement effects on information and interest in politics. We want to know whether information and interest are transmitted across networks, and, if yes, whether the effects are negative like for political participation. As in the previous section, we employ the three centrality variables presented – chatting, kinship, and geographic proximity – and we estimate network reinforcement effects with regression (2.3) and network diffusion effects with regression (2.4). All regressions are estimated using post-election data only. We control for provincial dummies, EA characteristics, and individual characteristics. The main focus is the coefficient of $\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ and the interaction terms $T_i\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ (reinforcement) and $T_v\frac{1}{N}\sum_{j\neq i} g_{ij}V_j$ (diffusion).

Table 9 shows the results for our index of information about the elections. We only find one statistically significant interaction effect: network reinforcement through kinship when analyzing the impact of the newspaper treatment. The estimated coefficient implies that the average respondent in the control group (in terms of kinship) sees an increase in the index of information of 0.079 standard deviation units through network reinforcement for the newspaper treatment. This is a relatively small effect, but unlike the ones observed for political participation, it is positive.

Results for our index of interest in politics are presented in *Table 10*. We find a larger number of significant interaction effects, all of which are positive. In contrast, ATE's are all non-significant, suggesting that all the effect of treatment is due to peer effects. The most robust network effects are found for the newspaper treatment. Results indicate that both reinforcement and diffusion are channeled through kinship and chatting, and that only diffusion is channeled through geographical proximity. Almost all peer effects of the newspaper treatment are significant at the 1 percent level. As for *Tables 6* and *8*, we report at the bottom of *Table 10* the difference in outcome between a treated subject with $g_{ij} = 0$, and one with an average value of

g_{ij}. The difference is measured in terms of standard deviation units of the dependent variable. Taking the newspaper treatment, for instance, we see that, relative to someone with no chatting network, a subject with an average network is 0.27 (reinforcement) or 0.30 (diffusion) standard deviation units more interested in elections. For kinship, the correspondings figures are 0.14 and 0.21. We find slightly lower network reinforcement effects for the hotline through chatting and kinship. Slightly lower reinforcement and diffusion network effects are also present for the civic education treatment, but only through chatting.

To summarize, for information and interest in politics, direct treatment effects and network effects are all positive – even though the direct treatment effect on interest in politics is not statistically significant. These findings suggest that information and interest in politics are transmitted across networks, possibly because the transmission of information and interest in politics across peers does not entail large costs and, therefore, does not lead to free riding.

6.5. Robustness check: multiple imputation

Although balance tests do not indicate that panel attrition significantly affects the comparability of treatment and control groups, we nevertheless test how sensitive our results are to missing post-elections observations. We use the multiple imputation method to replace the missing values of outcome and control variables; and we re-estimate the average and network effects on political participation using the full sample of baseline respondents. We employ multivariate normal regressions.[18] In the imputation model we include the variables that we use in our empirical analysis, other characteristics of the household and of the respondents, characteristics of the enumeration area, and interactions between the interventions and characteristics of the household and respondents.

Recalculated estimates of the average treatment effects on political participation are similar to the ones obtained earlier. In *Table A8a* in the Online Appendix we present the average effect of each of the three treatments on the political participation of targeted respondents. This table

[18] Given than most variables are categorical, we considered using chained equations. However, it was very difficult to find a model that would include all the relevant variables and converge. In addition, *Schafer and Graham (2002)* argue that normal imputation models have a good performance for linear regressions, even when the variables are non-normal.

is to be compared with *Table 1* and *Table A5*. We find a very similar pattern of significant effects. When looking at the results relating to untargeted respondents in *Table A8b,* the differences to the main results are even smaller.[19]

For network effects, the coefficients of the network interaction terms remain negative for most treatments and network measures. We find a similar pattern of significant results, although with a smaller magnitude. *Table A9* displays the estimates of peer effects on turnout measure index 1 using imputed data. Comparing these results with the ones displayed in *Table 7* we see that most of the significant network effects remain, although they have smaller magnitudes. The same can be said when we compare the estimates of the interaction terms with other turnout measures using imputed data (*Table A10*) and using the original data (*Tables 6 and 7*). Overall, we conclude that using multiple imputation to correct for attrition corroborates our findings.

7. Discussion

Having summarized our empirical results, we now discuss what we have learnt about the channel through which the treatments affected outcomes. Learning about channels of influence is important for the external validity of our findings. We begin by describing potential channels by which our three treatments can affect turnout. We also relate to our empirical results.

A first possibility is that credible information about the electoral process increases voter confidence and induces discouraged voters, namely opposition supporters, to vote. We would then expect an increase in information and interest about the electoral process in treated EAs, as well as a higher share of ballots going to the opposition. Because information is likely to diffuse along social networks, we expect a positive average diffusion effect on untargeted individuals.[20] Relating to our empirical results, we find some but limited effects of the treatments on information and interest: even though we can report some impact on information and interest of our panel respondents, we do not find clear effects on information and interest of secondary interviewees. Still, we report clear diffusion effects on the turnout of untargeted individuals

[19] The same can be reported about the effects of the treatments on secondary interviewees (tables available upon request).

[20] See for instance *Montgomery and Casterline (1996)* on social learning.

(namely those of the newspaper). Crucially, we do not find positive effects of the treatments on voting for the opposition.

A second possibility is that people may vote because the treatments inflame partisan passions and people want to 'support their team'. This channel of influence does not require that people become more knowledgeable about the details of the electoral process. Since people vote not so much to affect the electoral outcome but to show support for a party or candidate, it does not matter if they do not expect to be pivotal voters. In this case, we therefore expect treatments to induce high participation rates, and, in a context dominated by the incumbent, more votes for the ruling party. Because this channel of influence relies on herding behavior, we expect to observe both diffusion and reinforcement effects. More of our findings are consistent with a support-your-team effect: namely the limited effects on information and the clear turnout effects. Treatments increase voting for the dominant party, a finding that is difficult to reconcile with the idea that treatment increased turnout by reassuring opposition voters to cast their vote. We also find that the hotline treatment has the strongest positive effect on turnout among targeted and untargeted, perhaps because SMS messages about electoral abuse can be used to rally others.

A third possibility is that our treatments affect voting through social pressure – either directly through treatment nudging, or indirectly through peer-to-peer diffusion and reinforcement effects. This channel is likely to be most relevant when the act of voting is seen as a civic duty. Civic education is expected to have the strongest direct treatment effect in this case because it is focused on nudging. If this is the channel through which treatments increase turnout, we do not necessarily expect treated individuals to be more knowledgeable about the electoral process, or to be more interested in the electoral outcome. Some of our findings are consistent with a social pressure/civic duty interpretation, notably the robust direct effect of the civic education treatment on turnout and the positive turnout effect on secondary interviewees without any simultaneous effect on information or interest about the elections. The latter is suggestive that treatments raise turnout among dependents not because they are better informed or more interested in elections, but because they have been told to vote by someone who has some authority

over them.

This leaves one systematic effect that, at first glance, seems to contradict all the above interpretations – namely, the negative reinforcement and diffusion effects associated with the treatments. This means that individuals who are more central in the village network are less induced to vote by treatment than less connected individuals. We now develop a conceptual framework that is consistent with this finding.

7.1. Conceptual framework

To help us interpret our empirical results, we introduce a general framework for the analysis of voter participation. The focus is primarily on turnout, even though other forms of political participation (e.g., open letter) largely follow the same logic. The starting point of our effort is the idea that an educational campaign about elections raises the information level of voters. This affects their beliefs – e.g., in the fairness and transparency of the electoral process – and hence the interest they have in the voting process. People then adjust their behavior to reflect their new levels of information and interest. The campaign may also trigger various forms of peer effects which can impact the same outcomes.

We formalize this general idea as follows, building on numerous sources summarized by *Dhillon and Peralta (2002)* and *Feddersen (2004)*. Let us assume that an individual i decides a political participation vector x_i (e.g., casting a vote, voting for a specific candidate, sending text messages with political content) to maximize the following payoff function:

$$\max_{x_i} E_{\Omega_i} U(G(x_i, x_{-i}), x_i) - C(x_i) \tag{7.1}$$

where $G(x_i, x_{-i})$ is the outcome of the electoral process, x_{-i} is the combined action of individuals other than i, Ω denotes i's information set, and $C(x_i)$ is the total material cost of the action (e.g., transport cost, opportunity cost of time, cost of text messaging). To capture non-instrumental motivations – e.g., support-your-team motive, or civic-mindedness – we allow x_i to enter the function U independently from the outcome of the voting process G.

The first order condition:

$$E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}+\frac{\partial U}{\partial x_i}\right]=\frac{dC}{dx_i}$$

illustrates how a voter education campaign can influence turnout. First, the campaign can change voters' information set Ω_i. Distributing information about the electoral process may convince voters of the integrity of the electoral process, thereby raising $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]$. Second, the campaign may increase the non-instrumental motivation $\partial U/\partial x_i$ either through a support-your-team effect or by raising civic-mindedness. All these channels are conducive to increasing voter participation: that is the direct impact we expect from the voter education campaign that we study.

Focusing on turnout, difficulties arise when non-instrumental motivations are absent, i.e., when $\frac{\partial U}{\partial x_i}=0$. Optimal turnout requires $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]=\frac{dC}{dx_i}$. Since a single vote has little effect on the electoral outcome – $\partial G(x_i,x_{-i})/\partial x_i$ is small – voting is not individually rational unless the marginal cost of participation is zero, which is clearly not the case in our study area. This paradox dates back at least to *Downs (1957)*.[21] Introducing non-instrumental motives for voting eliminates the problem: the first order condition $E_{\Omega_i}\left[\frac{\partial U}{\partial x_i}\right]=\frac{dC}{dx_i}$ can be satisfied for an interior x_i even when $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]=0$.

As already discussed, the diffusion and reinforcement of treatment effects can in principle operate through the circulation of information which affects voters' information set Ω_i. They can also operate via social pressure to support one's team or to vote because it is a civic duty. For social pressure to arise spontaneously, however, voters must somehow derive a subjective benefit that increases in the total number of votes cast, not just their own. Otherwise they would have no reason to spend time and effort encouraging others to vote. Let total turnout in location v be denoted x_v. We now assume that $U(G(x_i,x_{-i}),x_i,x_v)$ with $\partial U/\partial x_v\geq 0$ – people value a higher

[21] A lively debate has followed. Using a game-theoretic voting game with two candidates, *Palfrey and Rosenthal (1983)* find a high turnout equilibrium generated by a high probability of being pivotal. This stems from having nearly identical numbers of voters supporting each candidate. This result was short-lived: the same authors (*Palfrey and Rosenthal, 1985*) demonstrate that the introduction of incomplete information and a large population eliminates the possibility that high turnout arises in equilibrium. Recently, *Myatt (2012)* recovered the idea that $\partial G(x_i,x_{-i})/\partial x_i$ depends on the perceived competitiveness of the election. Myatt considers a two-candidate election in which there is aggregate uncertainty about the popularity of each candidate. Despite an underdog effect through which higher turnout from the underdog compensates the advantage of the frontrunner, Myatt finds that turnout is high and that it peaks in elections that are expected to be close.

turnout.[22] We also assume that $\partial^2 U/\partial x_v^2 < 0$ to capture the intuitive idea that individuals care about achieving a minimum target level of electoral participation, beyond which utility rises less. By voting themselves, individuals also increase x_v. Ignoring instrumental motivations, the new first order condition becomes:

$$E_{\Omega_i}\left[\frac{\partial U}{\partial x_i} + \frac{1}{N}\frac{\partial U}{\partial x_v}\right] = \frac{dC}{dx_i}$$

where N denotes the number of eligible voters in the polling station. Simple comparative statics yields:

$$\frac{dx_i}{dE_{\Omega_i}[x_v]} = -\frac{\frac{1}{N}E_{\Omega_i}\left[\partial^2 U/\partial x_v^2\right]}{SOC} < 0$$

An individual who expects a higher turnout by others is less likely to vote himself or herself. This free riding effect can arise in the support-your-team environment: if others turn up in large numbers to support their favorite politician, it is less essential that i turns up as well. It can also arise when social pressure is motivated by conformism to a civic norm: if the village community as a whole performs its duty, one member can more easily be excused. This kind of reasoning probably explains why voting is more prevalent among heads of household and their spouse than among dependent adults: if mom and dad have voted, I do not need to.

Treatment affects turnout, but who in the village is more likely to realize this? Presumably those individuals who are better connected and can discuss turnout intentions with many people. To illustrate how this can affect voting, imagine that each individual i expects turnout to be the same as in the last election, except for those whose voting intentions that i can observe or infer. Let N_i be the subset of voters that i observes, and N_{-i} be the rest. Let T denote treatments as before. We have $x_v = \sum_{N_i} x_j + \sum_{N_{-i}} x_j$ and $\frac{\partial E_{\Omega_i}[x_v]}{\partial T} = \frac{N_i}{N}E_{\Omega_i}\left[\frac{\partial x_j}{\partial T}\right]$. Since the effect of treatment on turnout is positive, i.e., $\frac{\partial x_j}{\partial T} > 0$, it follows that individuals with a larger N_i increase their expectation $E_{\Omega_i}[x_v]$ more than people with a small N_i – and thus are less likely

[22] There may be many reasons for this, some subjective, others material. To illustrate the latter, suppose that one party has overwhelming support and that it values the legitimacy provided by the electoral process, as is reasonable to assume in the case of the 2009 Mozambique elections. In this case, the election may turn into a turnout contest across locations for clientelistic benefits: the incumbent can look at turnout per ballot station and reward those with high turnout after the election. In this setting, a voter education campaign is likely to mobilize additional incumbent supporters.

to vote. This formally demonstrates that treatment can potentially increase free-riding among better connected individuals.

7.2. Free-riding vs. saturation

Demonstrating that treatment can theoretically increase voter free-riding is not the same thing as showing that free-riding is behind our result. There may be other explanations. One particularly threatening candidate is the possibility that the negative coefficient of the $T_i \frac{1}{N} \sum_{j \neq i} g_{ij} V_j$ term reflects a voter saturation effect rather than free-riding. Because individuals with a larger social network vote with a high probability on average, it may be more difficult for them to further increase their likelihood of voting. This, and not free-riding, could explain why the effect of the treatment on these individuals is weaker than on individuals with a smaller social network. To show this formally, let P_i represent individual i's propensity to vote in the absence of treatment. We now assume that voter turnout among the targeted follows the following model:

$$y_i = P_i + \beta T_i + \gamma T_i \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \pi P_i T_i + \varepsilon_i \tag{7.2}$$

where γ captures free-riding as before and a significantly negative π coefficient indicates voter saturation. A bias in the estimation of γ arises if P_i is correlated with network size $\frac{1}{N} \sum_{j \neq i} g_{ij} V_j$. To demonstrate, let $P_i = \alpha + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} V_j$ and replace P_i in (7.2):

$$y_i = \alpha + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + (\beta + \alpha\pi) T_i + (\gamma + \delta\pi) T_i \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \varepsilon_i \tag{7.3}$$

Comparing (7.3) with (2.3) it is immediately apparent that voter saturation – a negative π – can be misinterpreted as free-riding – a negative γ – when estimating regression (2.3).

The solution we propose is to estimate $\widehat{P_i}$ using individuals in untreated locations, and use it as a control function to obtain separate estimates of γ and π. We obtain $\widehat{P_i}$ by regressing, using control individuals only, y_i on network size $\frac{1}{N} \sum_{j \neq i} g_{ij} V_j$ and other characteristics known to affect turnout, such as gender and age. Because treatment is assigned randomly, $\widehat{P_i}$ is a consistent predictor of treated individuals' propensity to vote in the absence of treatment. We

can thus estimate (7.2) on targeted individuals using $\widehat{P}_i$ in lieu of P_i. Since by design $T_i = 1$ for the targeted, the estimated regression boils down to:[23]

$$y_i = \beta + (1 + \pi)\widehat{P}_i + \gamma \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \varepsilon_i \tag{7.4}$$

Voter saturation $\pi < 0$ requires that the coefficient of $\widehat{P}_i$ be less than 1. Coefficient γ in regression (7.4) is estimated free of voter saturation bias.

Equation 7.4 is estimated with and without provincial dummies and individual controls, for different turnout measures y_i, and separately for each treatment and each network measure. The results for the individuals targeted by the hotline are displayed in *Table A11.* The different panels in the table relate to chatting, kinship, and geographical proximity, respectively. Point estimates of γ are negative in most regressions and they are significant in many of them. Furthermore, the point estimates of the coefficient of $\widehat{P}_i$ are in most cases larger than one. Similar results are obtained for the civic education treatment, but the evidence is less clear for the newspaper treatment, possibly because it has a smaller treatment effect to start with. Overall, this evidence corroborates the hypothesis of a free-riding effect in the hotline and civic education treatments.[24]

7.3. Endogenous peer effects

Since *Manski (1993)*, the literature distinguishes two types of peer effects: exogenous and endogenous. Exogenous or contextual effects refer to situations in which an individual's behavior depends on the exogenous characteristics or circumstances of his/her peers. Endogenous effects refer to situations in which the behavior of the individual depends on the behavior of his/her peers.

The reduced form equations (2.3) and (2.4) estimate the combined exogenous and endogenous peer effects. In the setting of our experiment, endogenous peer effects correspond to the situation where subjects are less likely to vote when they know that more of their social or geographical

[23] It is easy to verify that including the control individuals as well does not affect the results, given the way $\widehat{P}_i$ is constructed. So control individuals can be ignored.

[24] We repeated the analysis on untargeted individuals in treated locations. We do not find significant effects, possibly because of the small sample size.

neighbors will vote. In contrast, an exogenous peer effect arises when the treatment of individual j has a direct effect on the voting behavior of individual i. Distinguishing between the two types of peer effects may help confirm free riding.

To throw some light on the issue, we estimate a model that incorporates endogenous peer effects but assumes away exogenous effects. This alternative model is written as follows (for the case of targeted respondents):[25]

$$y_i = \alpha + \beta T_i + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} V_j + \lambda \frac{1}{N} \sum_{j \neq i} g_{ij} V_j y_j + \varepsilon_i. \tag{7.5}$$

If peer effects estimated in regressions (2.3) and (2.4) are driven exclusively by exogenous peer effects, then we should find no evidence of endogenous peer effects, i.e., we should observe $\lambda = 0$.

Regression (7.5) cannot be estimated using OLS due to the reflection problem (e.g., *Manski, 1993*). Therefore, we instrument the voting behavior of i's neighbors ($\frac{1}{N} \sum_{j \neq i} g_{ij} y_j$) with their treatment ($\frac{1}{N} \sum_{j \neq i} g_{ij} T_j$).[26] The equation is estimated using two-stages least squares. The coefficient λ measures endogenous network effects. A negative coefficient ($\lambda < 0$) indicates that participants are less likely to vote when more of their social or geographical neighbors are expected to vote as a result of treatment.

The full results for turnout measure index 1 are presented in *Table A12*. The estimates of regression (7.5) are displayed in columns (1)-(3). The estimates of a similar regression for untargeted respondents are displayed in columns (4)-(6). All regressions are estimated using follow-up observations only. We control for provincial dummies, EA characteristics, and individual characteristics. The coefficient of $\frac{1}{N} \sum_{j \neq i} g_{ij} V_j y_j$ is negative in all regressions, significantly so for reinforcement effects through geographical proximity. This offers some additional support to our free-riding interpretation.

[25] Given our experimental design, it is in principle possible to estimate endogenous and exogenous peer effects simultaneously by using the treatment of i's neighbors as instrument for the behavior of i's neighbors (see *Bramoulle, Djebbari and Fortin, 2009*). We tried this approach as well. Unfortunately the small sample size in each location precluded this approach: because of overlap in distance-2 neighborhoods, there is not enough variation in the instrument to identify endogenous and exogenous effects separately.

[26] Here $\frac{1}{N} \sum_{j \neq i} g_{ij}$ is computed using all respondents in the EA for whom information on y_j is available, to correct for possible missing observations.

8. Concluding remarks

Using a large-scale field experiment, we have investigated how voter education treatments affected political participation in the 2009 elections in Mozambique. Three types of interventions were tested: distribution of an independent newspaper; access to a text message hotline; and a civic education campaign. The interventions are shown to increase voter turnout and the electoral knowledge of targeted and untargeted individuals in treated locations; and to increase voter turnout of individuals living with targeted subjects. Using several measures of network centrality based on social and geographical connectedness, we estimate reinforcement and diffusion network effects. We find that peer effects on political participation are consistently negative, i.e., individuals connected to many treated individuals are less likely to vote than similar individuals with fewer connections. This is particularly clear for the hotline treatment. At the same time, information and interest in politics are positively transmitted across peers.

We interpret these findings in the context of a voter participation framework where voter education can affect information and interest in politics, and, hence, change voter behavior. We argue that the sign of the network effects suggests free-riding: a smaller treatment effect on central individuals for electoral participation seems to result from their realization that the campaign is driving more people to vote, making their own turnout less essential.

These results have implications for the design of voter education campaigns. While social networks tend to magnify cheap information and interest effects, they tend to produce free-riding when costly behavior is elicited. However, we must emphasize that these findings may be specific to countries similar to Mozambique. The 2009 elections pitted against each other the two main protagonists of the civil war that followed independence. The voter education may have brought back memories of the war and, so doing, may have raised partisanship. This is in agreement with recent experimental evidence showing that civil war increases in-group egalitarianism but reduces it across groups (*Bauer, Cassar, Chytilova, and Henrich, 2012*). Since FRELIMO had an overwhelming dominance in these elections, this raised partisanship may have mobilized FRELIMO voters disproportionately. This may have helped the sense that political competition had decreased, hence leading to the free-riding peer effects we document.

9. References

Aker, J., P. Collier, and P. C. Vicente (2013): "Is Information Power? Using Cell Phones and Free Newspapers during an Election in Mozambique," Universidade Nova de Lisboa, mimeograph.

Angelucci, M., and G. De Giorgi (2009): "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?," *American Economic Review*, 99(1): 486–508.

Angelucci, M., G. De Giorgi, M. Rangel, and I. Rasul (2010): "Family Networks and School Enrollment: Evidence from a Randomized Social Experiment," *Journal of Public Economics*, 94(3-4), 197–221.

Baird, S., J. A. Bohren, C. McIntosh, and B. Ozler (2014): "Designing Experiments to Measure Spillover Effects", Penn Institute for Economic Research, Working Paper No. 14-006.

Bandiera, O., and I. Rasul (2006): "Social Networks and Technology Adoption in Northern Mozambique," *Economic Journal*, 116(514), 862–902.

Banerjee, A. V., S. Kumar, R. Pande, and F. Su (2011): "Do Informed Voters Make Better Choices? Experimental Evidence from Urban India," Harvard University, mimeograph.

Bauer, M., A. Cassar, J. Chytilova, and J. Henrich (2012): "Warfare Experience during Ontogeny Increases Egalitarian and Parochial Motivations," Charles University in Prague, mimeograph.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of Peer Effects through Social Networks," *Journal of Econometrics*, 150(1), 41-55.

Bruhn, M. and D. McKenzie (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *American Economic Journal: Applied Economics*, 1(4), 200-32.

Chandrasekhar, A. G., and R. Lewis (2012): "Econometrics of Sampled Networks," Massachusetts Institute of Technology, mimeograph.

Collier, P., and P. C. Vicente (2014): "Votes and Violence: Evidence from a Field Experiment in Nigeria," *Economic Journal*, 124(574), F327-355.

Cungara, B., G. Fagilde, J. Garrwtt, R. N. Uaiene, and D. Headey (2011): "Growth Without

Change: The Elusiveness of Agricultural and Economics Transformation in Mozambique," Paper presented at "Dialogue on Promoting Agricultural Growth in Mozambique", 21 July 2011. Available at: http://fsg.afre.msu.edu/mozambique/caadp/

Dale, A., and A. Strauss (2009): "Don't Forget to Vote: Text Message Reminders as a Mobilization Tool," *American Journal of Political Science*, 53(4), 787–804.

De Vletter, F. (2001): "Coping with Extreme Poverty through Traditional Skills: The Case of the Xirundzu Basket Makers of Mozambique," International Labour Organization, Southern African Multidisciplinary Advisory Team (ILO/SAMAT), Discussion Paper No. 18.

Dhillon, A., and S. Peralta (2002): "Economic Theories of Voter Turnout," *Economic Journal*, 112, F332–F352.

Downs, Anthony (1957): *An Economic Theory of Democracy*, New York: Harper and Row.

Fafchamps, M., and P. C. Vicente (2013): "Political Violence and Social Networks: Experimental Evidence from a Nigerian Election," *Journal of Development Economics*, 101, 27-48.

Feddersen, T. (2004): "Rational Choice Theory and the Paradox of Not Voting," *Journal of Economic Perspectives*, 18(1): 99-112.

Fujiwara, V., and L. Wantchekon (2013): "Can Informed Public Deliberation Overcome Clientelism? Experimental Evidence from Benin," *American Economic Journal: Applied Economics*, 5(4), 241-255.

Gerber, A. and D. Green (2000): "The Effect of a Nonpartisan Get-out-the-Vote Drive: An Experimental Study," *Journal of Politics*, 62, 846–57.

Gine, X., and G. Mansuri (2011): "Together We Will: Experimental Evidence on Female Voting Behavior in Pakistan," World Bank, Policy Research Working Paper No. 5692.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007): "Experimental Analysis of Neighborhood Effects," *Econometrica*, 75(1), 83–119.

Miguel, E., and M. Kremer (2004): "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, 72(1), 159–217.

Manski, Charles F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60, 531-42

Montgomery, M. R. and J. B. Casterline (1996): "Social Learning, Social Influence, and New Models of Fertility," *Population and Development Review*, 22, Supplement: Fertility in the United States: New Patterns, New Theories, 151-175.

Myatt, D. P. (2012): "A Rational Choice Theory of Voter Turnout," London Business School, mimeograph.

Nickerson, D. W. (2008): "Is Voting Contagious? Evidence from Two Field Experiments," *American Political Science Review*, 102(1), 49–57.

Palfrey, T., and H. Rosenthal (1983): "A Strategic Calculus of Voting," *Public Choice*, 41(1), 7-53.

Palfrey, T., and H. Rosenthal (1985): "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79(1), 62-78.

Pop-Eleches, C, H. Thirumurthy, J. P. Habyarimana, J. G. Zivin, M. P. Goldstein, D. De Walque, L. MacKeen, J. Haberer, S. Kimaiyo, J. Sidle, D. Ngare, D. R. Bangsberg (2011): "Mobile Phone Technologies Improve Adherence to Antiretroviral Treatment in a Resource-limited Setting: a Randomized Controlled Trial of Text Message Reminders," *AIDS*, 25, 825–34.

Schafer, J. L. and J. W Graham (2002): "Missing Data: Our View of the State of the Art," *Psychological Methods*, 7(2), 147–177.

Sheldon, K. (1999): "Machambas in the City: Urban Women and Agricultural Work in Mozambique," *Lusotopie*, 121-140.

Thaler, R. H., and C. R. Sunstein (2009): *Nudge: Improving Decisions about Health, Wealth and Happiness*, New York: Pinguin Books.

Vaz, A. (2013): "Interpersonal Influence and Network Effects on Voting Behavior: Experimental Evidence from Mozambique," D.Phil. Thesis, Oxford University, mimeograph.

Vicente, P. C. (2014): "Is Vote-Buying Effective? Evidence from a Field Experiment in West Africa," *Economic Journal*, 124(574), F356-387.

Wantchekon, L. (2003): "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin," *World Politics*, 15(3), 399-422.

Table 1: Average effect of the three treatments on respondents' political participation

	Homogeneous effects on targeted respondents						Homogeneous effects on untargeted respondents					
	Turnout				Open letter		Turnout				Open letter	
	Finger measure		Index 2				Finger measure		Index 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.014	0.013	0.033	0.036	0.088*	0.099**	0.143***	0.122***	0.092***	0.075**	0.065	0.088
	(0.040)	(0.036)	(0.033)	(0.029)	(0.050)	(0.050)	(0.036)	(0.036)	(0.034)	(0.035)	(0.066)	(0.065)
Hotline	0.063**	0.049*	0.078***	0.070***	-0.036	-0.017	0.090**	0.074*	0.085**	0.078**	0.007	0.022
	(0.026)	(0.027)	(0.024)	(0.025)	(0.035)	(0.033)	(0.043)	(0.042)	(0.033)	(0.033)	(0.047)	(0.044)
Civic education	0.055*	0.046	0.050**	0.050**	0.043	0.053	0.057	0.046	0.059	0.047	0.106	0.108
	(0.029)	(0.029)	(0.026)	(0.025)	(0.048)	(0.046)	(0.048)	(0.051)	(0.039)	(0.039)	(0.070)	(0.068)
Mean dep. variable (control)	0.807	0.805	0.757	0.756	0.153	0.151	0.807	0.811	0.757	0.758	0.153	0.151
Adjusted R-squared	0.014	0.027	0.036	0.069	0.013	0.029	0.015	0.031	0.039	0.058	0.015	0.047
No. of observations	953	943	953	943	973	957	437	431	437	431	449	442
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions (1) to (6) include observations for targeted (in treated locations) and control respondents. Regressions (7) to (12) include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. In the second column for each outcome we control for demographic characteristics and enumeration area characteristics. All regressions include province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Average effect of the three treatments on respondents' information and interest

	Homogeneous effects on targeted respondents				Homogeneous effects on untargeted respondents			
	Basic electoral information		Interest in elections		Basic electoral information		Interest in elections	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Newspaper	0.115	0.168**	-0.011	-0.027	0.097	0.120	0.004	-0.007
	(0.086)	(0.082)	(0.079)	(0.080)	(0.125)	(0.110)	(0.144)	(0.146)
Hotline	0.162**	0.214***	0.113	0.090	0.181	0.239**	0.082	0.109
	(0.078)	(0.080)	(0.077)	(0.073)	(0.115)	(0.112)	(0.104)	(0.099)
Civic education	0.037	0.043	0.074	0.060	0.159	0.177	-0.011	0.007
	(0.089)	(0.090)	(0.059)	(0.060)	(0.121)	(0.113)	(0.133)	(0.123)
Adjusted R-squared	0.037	0.055	0.117	0.126	0.057	0.179	0.185	0.225
No. of observations	976	975	976	975	453	453	454	454
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: All regressions are OLS and use only second-round data. The dependent variables are indices. In the second column for each outcome we control for demographic characteristics and enumeration area characteristics. All regressions include province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Average effects of the treatments on secondary interviewees' political participation

	Turnout				Open letter	
	Finger measure		Index 2			
	(1)	(2)	(3)	(4)	(5)	(6)
Newspaper	0.035	0.017	0.042	0.030	-0.008	-0.007
	(0.055)	(0.058)	(0.049)	(0.049)	(0.029)	(0.032)
Hotline	0.100**	0.072	0.063	0.051	0.070*	0.057
	(0.049)	(0.050)	(0.042)	(0.041)	(0.036)	(0.037)
Civic education	0.093*	0.082	0.074*	0.088**	0.030	0.022
	(0.054)	(0.055)	(0.044)	(0.042)	(0.031)	(0.032)
Mean dep. variable (control)	0.771	0.773	0.764	0.755	0.054	0.058
Adjusted R-squared	0.006	0.032	0.003	0.056	0.007	0.021
No. of observations	426	397	426	397	459	427
Controls	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for secondary interviewees living with targeted and control respondents. All regressions are OLS and use only second-round data. In the second column for each outcome we control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Average effect of the treatments on secondary interviewees' information and interest

	Basic electoral information		Interest in elections	
	(1)	**(2)**	**(3)**	**(4)**
Newspaper	-0.072	-0.114	-0.036	0.061
	(0.140)	(0.118)	(0.131)	(0.105)
Hotline	0.060	0.005	0.175	0.211**
	(0.134)	(0.112)	(0.127)	(0.104)
Civic education	-0.042	-0.037	0.099	0.170
	(0.129)	(0.117)	(0.141)	(0.108)
Adjusted R-squared	-0.004	0.076	0.003	0.221
No. of observations	459	427	459	427
Controls	No	Yes	No	Yes

Note: Regressions include observations for secondary interviewees living with targeted and control respondents. All regressions are OLS and use only second-round data. The dependent variables are indices. In the second column for each outcome we control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Average effect of the three treatments on the official electoral results (ballot-station level)

	Presidential elections			Parliamentary elections		
	Turnout	**% votes in Guebuza**	**% votes in Dhlakama**	**Turnout**	**% votes FRELIMO**	**% votes RENAMO**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.055**	0.040*	-0.015	0.057**	0.034	-0.020
	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.016)
Hotline	0.051**	0.025	-0.015	0.053**	0.023	-0.017
	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.015)
Civic education	0.053**	0.046**	-0.032**	0.052**	0.039*	-0.038**
	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.015)
Mean dep. variable (control)	0.440	0.723	0.114	0.438	0.722	0.136
Adjusted R-squared	0.389	0.673	0.582	0.377	0.666	0.637
No. of observations	161	161	161	161	161	161
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include ballot stations in control and treated locations. All regressions are OLS. We control for enumeration area characteristics and province dummies. Standard errors reported.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Network effects on turnout measure based on inked finger

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.016	0.004	0.005	0.167***	0.163***	0.161***
	(0.042)	(0.043)	(0.050)	(0.040)	(0.040)	(0.044)
Hotline	0.059*	0.052*	0.074**	0.076	0.074	0.070
	(0.033)	(0.031)	(0.035)	(0.053)	(0.052)	(0.056)
Civic education	0.048	0.041	0.055	0.081	0.074	0.022
	(0.035)	(0.034)	(0.041)	(0.052)	(0.051)	(0.064)
Network	0.268**	0.233**	0.082*	0.341***	0.302***	0.093*
	(0.128)	(0.112)	(0.043)	(0.131)	(0.112)	(0.049)
Network x Newspaper	**0.031**	**-0.207**	**-0.105***	**-0.106**	**-0.088**	**-0.033**
	(0.152)	(0.186)	(0.058)	(0.166)	(0.189)	(0.046)
Network x Hotline	**-0.125**	**-0.215**	**-0.107****	**-0.446****	**-0.234**	**-0.103**
	(0.155)	(0.162)	(0.053)	(0.222)	(0.172)	(0.082)
Network x Civic education	**-0.116**	**-0.171**	**-0.091****	**-0.075**	**0.145**	**-0.107**
	(0.173)	(0.154)	(0.043)	(0.162)	(0.231)	(0.071)
constant	0.709***	0.753***	0.881***	0.567***	0.622***	0.722***
	(0.070)	(0.071)	(0.082)	(0.095)	(0.100)	(0.112)
Network Effect - Newspaper	0.007	-0.026	0.125*	-0.023	-0.009	0.039
Network Effect - Hotline	-0.029	-0.027	0.128**	-0.096**	-0.025	0.123
Network Effect - Civic education	-0.027	-0.021	0.108**	-0.016	0.015	0.127
Adjusted R-squared	0.032	0.015	0.017	0.027	0.008	0.015
No. of observations	845	845	721	379	379	324
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions on targeted vs. control include observations for targeted (in treated locations) and control respondents; regressions on untargeted vs. control include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics, enumeration area characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level. In the case of chatting and kinship, the network effect corresponds to the difference between the average treatment effect on individuals with average size networks and the average treatment effect on individuals with no network at all. In the case of geographic proximity, the network effect corresponds to the difference between the average treatment effect on individuals with average proximity and the average treatment effect on individuals with maximum proximity.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Estimates of the interaction coefficients in turnout regressions using the remaining survey proxies

		reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
		chatting	kinship	proximity	chatting	kinship	proximity
		(1)	(2)	(3)	(4)	(5)	(6)
Self-reported	Network x Newspaper	-0.057	-0.135	-0.081	-0.124	-0.066	-0.073*
		(0.128)	(0.145)	(0.055)	(0.153)	(0.186)	(0.043)
	Network x Hotline	-0.220*	-0.286**	-0.093*	-0.301**	-0.363**	-0.163***
		(0.123)	(0.124)	(0.051)	(0.141)	(0.170)	(0.063)
	Network x Civic education	-0.197	-0.274*	-0.079*	-0.158	0.019	-0.079
		(0.144)	(0.145)	(0.044)	(0.132)	(0.216)	(0.065)
Adjusted	Network x Newspaper	-0.237	-0.024	-0.011	-0.056	0.197	-0.099
		(0.152)	(0.162)	(0.080)	(0.203)	(0.232)	(0.089)
	Network x Hotline	-0.275*	-0.219	-0.020	-0.368**	-0.445*	0.038
		(0.153)	(0.160)	(0.079)	(0.178)	(0.229)	(0.128)
	Network x Civic education	-0.167	-0.132	-0.006	-0.032	0.396	0.125
		(0.167)	(0.159)	(0.064)	(0.166)	(0.292)	(0.077)
Index 1	Network x Newspaper	-0.053	-0.119	-0.060	-0.116	-0.071	-0.051
		(0.118)	(0.129)	(0.051)	(0.135)	(0.175)	(0.040)
	Network x Hotline	-0.200*	-0.255**	-0.069	-0.307**	-0.305*	-0.110*
		(0.113)	(0.110)	(0.046)	(0.137)	(0.163)	(0.063)
	Network x Civic education	-0.181	-0.254*	-0.061	-0.112	0.116	-0.040
		(0.136)	(0.131)	(0.041)	(0.118)	(0.198)	(0.060)
Index 2	Network x Newspaper	-0.078	-0.107	-0.058	-0.126	-0.044	-0.053
		(0.116)	(0.123)	(0.053)	(0.136)	(0.164)	(0.044)
	Network x Hotline	-0.193*	-0.239**	-0.072	-0.271**	-0.281*	-0.079
		(0.110)	(0.111)	(0.049)	(0.138)	(0.160)	(0.068)
	Network x Civic education	-0.131	-0.207	-0.048	-0.092	0.189	-0.029
		(0.137)	(0.130)	(0.043)	(0.114)	(0.201)	(0.059)
Interviewer assessment	Network x Newspaper	-0.017	-0.202	-0.104**	-0.061	-0.010	-0.048
		(0.135)	(0.159)	(0.052)	(0.155)	(0.220)	(0.049)
	Network x Hotline	-0.155	-0.256*	-0.107**	-0.201	-0.247	-0.102
		(0.131)	(0.137)	(0.049)	(0.178)	(0.180)	(0.067)
	Network x Civic education	-0.223	-0.316*	-0.092**	-0.132	-0.131	-0.060
		(0.160)	(0.173)	(0.039)	(0.173)	(0.205)	(0.067)

Note: Regressions on targeted vs. control include observations for targeted (in treated locations) and control respondents; regressions on untargeted vs. control include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics, enumeration area characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Network effects on behavior measure open letter

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.110**	0.107**	0.107*	0.078	0.095	0.076
	(0.053)	(0.054)	(0.056)	(0.067)	(0.075)	(0.070)
Hotline	-0.002	-0.004	-0.005	0.037	0.023	0.012
	(0.038)	(0.039)	(0.041)	(0.053)	(0.054)	(0.058)
Civic education	0.064	0.060	0.054	0.103	0.075	0.138
	(0.054)	(0.055)	(0.062)	(0.073)	(0.066)	(0.089)
Network	0.032	0.091	0.010	0.014	0.099	0.063
	(0.102)	(0.156)	(0.045)	(0.109)	(0.172)	(0.044)
Network x Newspaper	**-0.020**	**-0.156**	**0.004**	**0.196**	**0.395**	**0.086**
	(0.150)	(0.184)	(0.064)	(0.259)	(0.553)	(0.081)
Network x Hotline	**0.092**	**-0.160**	**0.015**	**0.100**	**0.263**	**0.040**
	(0.152)	(0.193)	(0.053)	(0.277)	(0.316)	(0.084)
Network x Civic education	**-0.212**	**-0.368****	**0.069**	**-0.191**	**-1.043*****	**-0.022**
	(0.146)	(0.179)	(0.061)	(0.199)	(0.391)	(0.090)
constant	0.133	0.127	0.106	0.154	0.134	0.165*
	(0.086)	(0.086)	(0.090)	(0.096)	(0.093)	(0.097)
Network Effect - Newspaper	-0.005	-0.018	-0.005	0.042	0.042	-0.103
Network Effect - Hotline	0.021	-0.019	-0.018	0.022	0.028	-0.048
Network Effect - Civic education	-0.049	-0.043**	-0.083	-0.041	-0.112***	0.026
Adjusted R-squared	0.014	0.015	0.020	0.005	0.025	0.003
No. of observations	817	817	699	386	386	332
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions on targeted vs. control include observations for targeted (in treated locations) and control respondents; regressions on untargeted vs. control include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics, enumeration area characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level. In the case of chatting and kinship, the network effect corresponds to the difference between the average treatment effect on individuals with average size networks and the average treatment effect on individuals with no network at all. In the case of geographic proximity, the network effect corresponds to the difference between the average treatment effect on individuals with average proximity and the average treatment effect on individuals with maximum proximity.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Network effects on the index of basic electoral information

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.194**	0.194**	0.192**	0.185	0.158	0.196
	(0.088)	(0.084)	(0.095)	(0.118)	(0.135)	(0.122)
Hotline	0.233***	0.236***	0.231**	0.231*	0.269**	0.232*
	(0.082)	(0.082)	(0.096)	(0.119)	(0.109)	(0.135)
Civic education	0.047	0.046	0.044	0.186	0.180	0.299**
	(0.099)	(0.099)	(0.118)	(0.128)	(0.128)	(0.123)
Network	-0.074	-0.394	0.125	-0.051	-0.358*	0.108
	(0.246)	(0.275)	(0.129)	(0.244)	(0.215)	(0.112)
Network x Newspaper	**0.278**	**0.641***	**-0.206**	**0.091**	**-0.124**	**-0.270**
	(0.368)	(0.376)	(0.145)	(0.566)	(1.171)	(0.181)
Network x Hotline	**-0.126**	**0.073**	**-0.110**	**-0.697**	**-0.747**	**-0.119**
	(0.314)	(0.382)	(0.127)	(0.633)	(0.773)	(0.182)
Network x Civic education	**-0.271**	**-0.163**	**-0.177**	**0.273**	**0.612**	**-0.207**
	(0.360)	(0.353)	(0.141)	(0.482)	(0.658)	(0.132)
constant	-0.210	-0.184	-0.062	0.165	0.224	0.255
	(0.162)	(0.151)	(0.166)	(0.184)	(0.182)	(0.224)
Network Effect - Newspaper	0.064	0.079*	0.245	0.020	-0.013	0.322
Network Effect - Hotline	-0.029	0.009	0.131	-0.150	-0.080	0.142
Network Effect - Civic education	-0.063	-0.020	0.211	0.059	0.066	0.246
Adjusted R-squared	0.062	0.066	0.074	0.176	0.184	0.182
No. of observations	865	865	741	395	395	340
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions on targeted vs. control include observations for targeted (in treated locations) and control respondents; regressions on untargeted vs. control include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics, enumeration area characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level. In the case of chatting and kinship, the network effect corresponds to the difference between the average treatment effect on individuals with average size networks and the average treatment effect on individuals with no network at all. In the case of geographic proximity, the network effect corresponds to the difference between the average treatment effect on individuals with average proximity and the average treatment effect on individuals with maximum proximity.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Network effects on the index of interest in elections

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	-0.033	-0.051	-0.076	-0.013	0.038	-0.056
	(0.079)	(0.079)	(0.091)	(0.141)	(0.142)	(0.148)
Hotline	0.102	0.094	0.118	0.063	0.082	0.076
	(0.071)	(0.073)	(0.087)	(0.098)	(0.095)	(0.116)
Civic education	0.006	0.006	0.005	-0.040	-0.083	0.021
	(0.060)	(0.061)	(0.065)	(0.129)	(0.152)	(0.142)
Network	-0.262	-0.344	0.010	-0.316*	-0.206	-0.054
	(0.178)	(0.270)	(0.064)	(0.166)	(0.315)	(0.087)
Network x Newspaper	**1.159*** **	**1.110*** **	**0.021**	**1.371*** **	**1.925*** **	**0.444** **
	(0.266)	(0.399)	(0.091)	(0.483)	(0.675)	(0.181)
Network x Hotline	**0.838*** **	**0.809** **	**-0.118**	**-0.044**	**-0.014**	**0.077**
	(0.234)	(0.380)	(0.110)	(0.440)	(0.650)	(0.225)
Network x Civic education	**0.444***	**0.135**	**-0.059**	**1.225*** **	**-0.045**	**0.041**
	(0.255)	(0.432)	(0.072)	(0.453)	(1.029)	(0.141)
constant	0.052	0.046	0.021	0.246	0.170	0.183
	(0.121)	(0.126)	(0.141)	(0.180)	(0.183)	(0.212)
Network Effect - Newspaper	0.269***	0.136***	-0.025	0.295***	0.206***	-0.529**
Network Effect - Hotline	0.194***	0.099**	0.141	-0.009	-0.001	-0.092
Network Effect - Civic education	0.103*	0.017	0.070	0.264***	-0.005	-0.049
Adjusted R-squared	0.130	0.120	0.091	0.236	0.225	0.229
No. of observations	865	865	741	396	396	341
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions on targeted vs. control include observations for targeted (in treated locations) and control respondents; regressions on untargeted vs. control include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics, enumeration area characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level. In the case of chatting and kinship, the network effect corresponds to the difference between the average treatment effect on individuals with average size networks and the average treatment effect on individuals with no network at all. In the case of geographic proximity, the network effect corresponds to the difference between the average treatment effect on individuals with average proximity and the average treatment effect on individuals with maximum proximity.

* significant at 10%; ** significant at 5%; *** significant at 1%.