Voting and Peer Effects: Experimental Evidence from Mozambique*

Marcel Fafchamps
Stanford University†

Ana Vaz
University of Oxford‡

Pedro C. Vicente
Universidade Nova de Lisboa§

December 2016

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 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 individuals in our sample. For untargeted individuals, we estimate how the campaign diffuses as a function of proximity to others in the sample. We find evidence for both effects, similar across treatments and proximity measures. The treatments raise the

*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@economics.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.

level of 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 consistent with free-riding through pivotal reasoning and we provide additional evidence to support this claim.

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'. 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.

Our focus is on estimating the peer effects of the different treatments within the household and the village. 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 sampled individuals. Diffusion refers to the effect of the treatment for untargeted individuals in treated locations who are close to other sampled individuals. *Aker, Collier, and Vicente (2016)* 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. We also report a clear increase in information about the elections among targeted and untargeted individuals.

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 a reduction in turnout if central voters realize that, because of the campaign, turnout will increase and their vote becomes less essential in achieving a politically acceptable electoral result.

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 Alan Gerber and Donald 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.

The paper is organized as follows. In Section 2 we introduce the context of our experiment. The treatments are introduced in detail in Section 3. Subsequently, in Section 4 we describe the data, including outcome and network variables. In Section 5 we report on average effects, including balance tests. Peer effects are presented in Section 6, together with robustness analysis. In Section 7 we discuss various possible interpretations of our peer effect estimates and we introduce additional tests to disentangle them. Section 8 concludes.

2. 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 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.[1]

Mozambique became independent from Portugal in 1975, after which FRELIMO (Frente de Libertação de Moçambique), the independence movement, led a single-party, socialist regime.

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

Beginning in 1977, Mozambique suffered a devastating civil war, fought between FRELIMO 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.

3. 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 cam-

paign, 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 Observatório Eleitoral. For more details on these organizations, see *Aker, Collier, and Vicente (2016).*

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;[2] 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 locations in the experiment had received the newspaper before.[3] 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

[2] 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*).

[3] 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.

an SMS hotline as they were prepared to receive text messages reporting electoral problems. 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 reports to be sent,[4] 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.[5]

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

[5] The experimental protocol for the three treatments was specifically that all targeted respondents were to be given the newspaper and the leaflets (for the hotline and civic education), and that no untargeted respondents

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

4. 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 experimental locations are nationally representative of the voting population of Mozambique that has mobile phone coverage. 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

were to be given these materials.

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.[6]

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, the 'targeted', were directly treated 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.[7]

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.[8] 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. We also reestimate our main results using a multiple imputation technique to account for missing post-election observations.

[6] One control substitute location was sampled but found not to be needed. It was added to our sample but has no impact on the results.

[7] Note that the relatively low number of untargeted respondents in treated locations has implications for relatively low statistical power to identify these diffusion effects.

[8] 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*).

4.1. 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. The third one is an indicator of whether the respondent showed without hesitation his/her index finger to the enumerator when asked about which finger was marked after voting (shows finger) – dipping one finger in indelible ink was the method used to prevent people from voting multiple times. Turnout index 1 is a composite index incorporating information on how well the respondent answered questions on the sequence of events during the election day (including the one on the inked finger). 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 (the number of ballot papers, whether there were photos of the candidates, the number of ballot boxes, whether they were transparent, whether they were colored, and whether the respondent showed his/her index finger). 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 dividing by 7, so they too range from 0 to 1.[9] In the text of the paper, even though we show evidence for all measures of individual turnout, we tend to focus our attention on our preferred measure of individual turnout, index 2, provided it includes the most important set of factual adjustments to self-reported turnout we have available in our design.

Our proxy of electoral information is built from survey questions asking respondents about

[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.

which type of elections took place in 2009, the duration of a presidential mandate, the lists of presidential candidates and parties running for the 2009 elections, and the meaning of electoral abstention. Our proxy for interest in the elections is based on 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 EAs in our experiment which, as mentioned earlier, were based on the polling stations themselves. We employ results for the

[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.

presidential and parliamentary elections of 2009.

4.2. Network variables

We collected three measures of social and geographical centrality. The first two are based on degree centrality in social networks.[11] Let i and j be two individuals in the selected sample of targeted and untargeted participants in EA v and let the EA sample size be $N_v + 1$. We ask each i whether he/she can identify j by his or her name. If this is the case, we then ask i whether he/she is related to j[12] and whether he/she talks to j on a regular basis.[13] We call the first type of social connection 'kinship' and the second 'chatting'. Let $g_{ijv} = 1$ if i has a social connection to j, and 0 otherwise. The social network of i is then defined as $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$ where N_v is the number of respondents other than i in EA v. Dividing by N_v serves to net out slight differences in sample size across EAs. To illustrate, suppose that g_{ijv} represents kinship. Then $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$ is the proportion of participants in EA v (other than i) that is related to i. Similarly, if g_{ijv} represents chatting, then $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$ is the proportion of participants in EA v (other than i) to whom i reports talking on a regular basis.

The third variable captures how close i is to other individuals in the sample. Since the sample is randomly selected, individuals who live close to the geographical center of the EA are closer to other participants than individuals located at the outskirts of the EA. Geographical centrality can thus be proxied by the (negative of the) average distance from i to others in the EA sample. Formally, let g_{ijv} be the physical distance between individuals i and j in EA v.[14] The geographical centrality of i is defined as $\frac{-1}{N_v}\sum_{j\neq i} g_{ijv}$: the higher, i.e., less negative, this

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

[13] 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'.

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

number is, the more central is i. With these definitions, social centrality increases in each of the measures and is normalized by the size of the EA sample.

5. Average treatment effects

We begin by summarizing the average treatment effects, some of which (though not all) already appear in *Aker, Collier and Vicente (2016)*. Peer effects are presented in the next Section.

5.1. Testing strategy

The combined (i.e., direct and indirect) average effects of the campaign are estimated as follows. Let y_{iv} be a measure of electoral behavior, information, or interest for individual i in village v. Let $T_v = 1$ if village v was treated, and 0 otherwise. Even though we test below the three distinct forms of voter education 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 only targeted and control observations in a regression of the form:

$$y_{iv} = \alpha + \beta T_v + \varepsilon_{iv} \tag{5.1}$$

Coefficient β is the average treatment effect on electoral behavior, information, or interest. This regression can also be estimated with village and individual controls.

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 only untargeted and control observations in a regression of the same form showed above. Coefficient β is then an estimate of the average diffusion effect of the campaign on the electoral behavior, information, or interest of untargeted individuals. Estimations of (5.1) constitute the focus of *Aker, Collier, and Vicente (2016)*.

5.2. Balance

Before we show average treatment effects, we check balance by treatment on the baseline data. *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 locations in the follow-up survey yields a similar pattern: in both samples of targeted and untargeted respondents, most household demographics are not significantly different. Panel attrition seems to have maintained comparability between the treatment groups in terms of observables. We also do not see statistically significant differences for EA characteristics, with the exception of existence of an health center for newspaper areas compared to control areas.

Social and geographical centrality variables are summarized in *Tables 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 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. Individual outcome variables from the baseline survey are explored in full detail in *Aker, Collier, and Vicente (2016)*: 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 such effects is apparent in the data.

[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.

5.3. Results

We start by summarizing the regression results estimating the average effects of the campaign. 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 variables shows finger and index 2, and the sending of the open letter, separately for targeted and untargeted individuals. The average effects on the remaining turnout measures are presented in *Table A5* of the Online Appendix. 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 index 2 measure shows turnout of 76 percent on average among control individuals. Since panel respondents are either the household head or their spouse, it is not surprising that average turnout among them is above the 2009 national average of 45 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. For the newspaper treatment we find a small positive effect when we use no controls (i.e., +3.2 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. This treatment increases the probability of sending a SMS proposing policy priorities by 8.5 percentage points (when using controls).

Turning to the average diffusion effect of the campaign on untargeted individuals, we find some statistically significant coefficients for all three treatments. Averaged over the six measures (without controls), we find 10.6, 9.8, and 5.7 percentage point increases in turnout among untargeted households for the newspaper, hotline, and civic education treatments, respectively. The treatment effect for the civic education is not always statistically significant, though. The result on the newspaper is surprising given that it has no statistically significant effects on turnout among the targeted. We find a positive average diffusion effect on the open letter for all treatments, but this effect is never significant.

We now turn to the effects of the treatments on information and interest in politics. These results are displayed in *Table 2*. We find significant positive effects of all treatments on information about the elections. This effect is largest for the hotline as it ranges between 0.15 and 0.17 standard deviation units for the targeted, and between 0.19 and 0.25 standard deviation units for the untargeted. The effects for the newspaper and civic education are 0.16 standard deviation units for the targeted, and respectively 0.17 and 0.16 standard deviation units for the untargeted (in the regressions with controls). We do not observe statistically significant effects of the treatments on interest in politics. However, the hotline and civic education interventions approach significance, with positive treatment effects for the targeted.

Finally, we summarize in *Table 3* the average effect of the treatments on actual electoral outcomes from official polling station records (see *Aker, Collier, and Vicente, 2016*). All treatments have a strong and significant positive effect on voter turnout. This effect ranges between 5.3 and 5.4 percentage points for the presidential elections, and between 5.2 and 5.6 percentage points for the parliamentary elections. On voting patterns, we observe a positive point estimate of all treatments on voting for the incumbent president (Guebuza) and party (FRELIMO), and a negative effect of all treatments on voting for the main challenger candidate (Dhlakama) 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 for RENAMO and its presidential candidate by 3.8 and 3.2 percentage points, respectively. To summarize, the

treatments increased voter turnout – which was their objective. But they also benefited the incumbent and hurt the chances of the main challenger.

6. Peer effects

6.1. Testing strategy

Drawing inspiration from *Fafchamps and Vicente (2013)*, who analyze the peer effects 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 individuals in treated EAs. We estimate a reinforcement effect model of the form:

$$y_{iv} = \alpha + \beta T_v + \delta \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \gamma T_v \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \varepsilon_{iv}, \tag{6.1}$$

where we employ observations from targeted and control individuals only, i.e., where we exclude untargeted individuals living in treated EAs. Regressor $\frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ is included as control variable to capture the effect that network centrality has on y_{iv} in the absence of treatment: coefficient δ measures the predictive effect of social or geographical centrality on y_{iv}. The main coefficient of interest is γ.[16] It captures how the effect of a treatment varies with social or geographical proximity to others in the same EA.[17]

We also investigate the presence of diffusion effects using the same specification above while employing observations from untargeted and control individuals only, i.e., excluding targeted individuals living in treated EAs. Interpretation is similar to that of reinforcement effects.

Based on earlier studies, we expect social links and geographical proximity to magnify the effect of treatment – e.g., because the information content of treatment spreads more readily to central nodes and thus leads to a stronger impact of treatment. It is also conceivable, however, that peer effects are strategic complements rather than strategic substitutes. For instance, in

[16] As is well known, when estimating regression (6.1), the coefficient of the treatment variable β is mechanically affected when we add any regressor interacted with treatment T_v. To ensure comparability with ATE estimates reported earlier, we express $\frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ in deviation from its sample mean. This method leaves γ unaffected, but ensures that β still estimates the ATE. This approach is followed throughout this section, i.e., all regressors interacted with T_v are always demeaned.

[17] 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.

our setting this would arise if $\gamma < 0$ but $\beta > 0$. This coefficient configuration indicates that treatment increases y_{iv}, but less so for more central individuals. This could arise if behavior y_{iv} is beneficial for the group but individually costly, and central individuals free-ride on the effect that treatment has on others. We revisit this point more in detail later.

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

6.2. Peer effects on political participation

We first apply the above testing strategy to our main focus of interest, namely political participation, measured by voter turnout and open letter. We start by showing in *Table 4* the regressions using index 2 as voter turnout measure. We employ the three centrality variables presented earlier: chatting, kinship, and geographic proximity. Network reinforcement effects estimated are displayed in columns (1)-(3); network diffusion effects 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 on the δ and γ coefficients in specification (6.1).

Regarding δ, we note that more central individuals have a higher turnout propensity in control EAs: estimated coefficients for $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$ are strongly positive, particularly when using chatting and kinship as centrality measure, but also when using geographical proximity. This means that, without treatment, individuals who are more central in their community are more likely to vote. From these results alone, 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 coefficients γ, they are negative for almost all treatments and centrality measures – though not always statistically significant. Reinforcement through geographical proximity is consistently negative and statistically significant for the hotline and civic education. Negative and statistically significant effects are also present for diffusion for the newspaper and hotline treatments.

To get a sense of the magnitude of these peer effects, we present towards the bottom of *Table 4* a simple calculation of the difference in predicted voting between a treated subject with extreme $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$ and a treated subject with the average value of $\frac{1}{N_v}\sum_{j\neq i} g_{ijv}$. For chatting and kinship, this comparison is between a treated subject with no connections, i.e., $g_{ijv} = 0$ for all j, and one with the average number of connections. A negative value means that moving from no connections to average centrality decreases the probability of voting induced by treatment. For geographical proximity, the comparison is between a hypothetical treated subject at distance 0 to others in the EA (i.e., most central), and a treated subject at the average distance. A negative value indicates that moving from being at the average distance from others to being completely central reduces the probability of voting induced by treatment. The p-values are the same as those for the interaction coefficients.

We see that the magnitude of peer effects is large, particularly through geographical proximity. For instance, relative to a hypothetical centrally located subject, the effect of treatment on the probability of voting for the average subject is 6.3 (civic education) and 15.7 (newspaper) percentage points higher, in terms of reinforcement and diffusion respectively. We also observe clear impacts through chatting and kinship, particularly for the hotline treatment. These are large figures relative to the direct effect of treatment itself. They show that the ATE hides large variation depending on geographical and network centrality: more central individuals experience a much smaller – and often negative – effect of treatment on their propensity to vote.

Table 5 summarizes the results on interaction coefficients for the remaining turnout measures. Significant effects are all negative. The hotline stands out as inducing the strongest reinforcement and diffusion effects. 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.[18] In Online Appendix *Table A6* we perform the same calculations as we did in second panel of *Table 4*. They confirm that the magnitude of the network effects is far from negligible, especially given the relatively small average effect of each treatment. Our preferred explanation for the negative network effects we

[18] 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.

find is free-riding through pivotal reasoning: 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 to achieve a sufficient win gap to remain in the good favors of the government, and the likelihood that they turn out to vote decreases. We explore this – and other – explanations for negative peer effects more in detail in the final section of this paper.

Results for the open letter are displayed in *Table 6*. 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 3.4 percentage point less likely to send the open letter if he/she has the average kinship network than if he/she has no kinship network at all. For untargeted subjects, the difference is 8.7 percentage points. The explanation may be the same as for voter turnout: individuals with a large network realize others will send the open letter as a result of treatment, and feel that their participation is less essential.

6.3. 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 elections among targeted and untargeted individuals. But we could not find a statistically significant treatment effect on interest in politics. We now examine reinforcement and diffusion effects on information and interest in politics. We want to know whether information and interest are transmitted socially or geographically, and, if yes, whether the effects are negative as was the case for political participation. Regression model (6.1) is the same as before. All regressions are estimated using post-election data. As before we control for provincial dummies, EA characteristics, and individual characteristics.

Table 7 shows the results for information about the elections. We find some largely positive

interaction effects but none is statistically significant. Unlike in the case of political participation, we do not see significant negative network effects, and observe that the largest of these effects in absolute value are positive.

Results for interest in politics are presented in *Table 8*. We find a large number of significant interaction effects (nine), all being positive with one exception. In contrast, ATE's are all not significantly different from zero, suggesting that treatment affects participants positively or negatively depending on their social or geographical centrality. The most robust peer effects are found for the newspaper treatment. Results indicate that chatting, kinship, and geographical proximity are all channels for reinforcement and diffusion effects. Almost all peer effects of the newspaper treatment are statistically significant, a majority of which at the 1 percent level. As for *Tables 4, 6, and 7*, we report in the second panel of *Table 8* estimates of the magnitude of network effects. For the newspaper treatment, we see that, relative to someone with no chatting links, a subject with an average chatting centrality is 0.21 (reinforcement) or 0.28 (diffusion) standard deviation units more interested in elections. For kinship, the corresponding figures are 0.11 and 0.19. We find lower reinforcement effects for the hotline and civic education treatments.

To summarize, direct treatment effects and network effects are generally positive – although the direct treatment effects on interest in politics and the network effects on information about politics are not statistically significant. These findings suggest that at least interest in politics is transmitted across networks, possibly because transmission across peers does not entail large costs and, therefore, is not subject to free-riding through pivotal reasoning.

6.4. Robustness checks

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-election 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.[19] 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 EA, 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 A7* in the Online Appendix we present the average effect of each of the three treatments on the political participation of targeted respondents. This table is to be compared with *Table 1*. We find a very similar pattern of significant effects, particularly for the results on diffusion employing untargeted respondents.

For networks interacted with the treatment effects, the coefficients remain negative for most treatments and network measures. We find a similar pattern of significant results, although with a smaller magnitude. *Table A8* displays the estimates of interaction effects on turnout index 2 using imputed data. Comparing these results with the ones observed in *Table 4* we see that almost all of the significant network effects remain, although they have smaller magnitudes. The same can be said when we estimate the interaction terms for other turnout measures using imputed data (*Table A9*) and compare them to the original results (*Table 5*). Overall, we conclude that using multiple imputation to correct for attrition corroborates our findings.

We also investigate whether peer effects vary depending on whether social and geographical proximity is to targeted or untargeted individuals. To this effect, we estimate two complementary models of the form:

$$y_{iv} = \alpha_1 + \beta_1 T_v + \delta_1 \frac{1}{N_v^{TARG}} \sum_{j \neq i; j \in TARG} g_{ijv} + \gamma_1 T_v \frac{1}{N_v^{TARG}} \sum_{j \neq i; j \in TARG} g_{ijv} + \varepsilon_{iv} \tag{6.2}$$

$$y_{iv} = \alpha_2 + \beta_2 T_v + \delta_2 \frac{1}{N_v^{UNTARG}} \sum_{j \neq i; j \in UNTARG} g_{ijv} + \gamma_2 T_v \frac{1}{N_v^{UNTARG}} \sum_{j \neq i; j \in UNTARG} g_{ijv} + \varepsilon_{iv} \tag{6.3}$$

where $TARG$ and $UNTARG$ refer to targeted and untargeted groups of individuals. We apply

[19] 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.

the above regression models to both reinforcement and diffusion. If peer effects occur solely through proximity to targeted individuals, we should observe $\gamma_1 \neq 0$ and $\gamma_2 = 0$. In this particular configuration, regression model (6.3) can then be seen as a falsification test of (6.2). In contrast, if γ_1 and γ_2 are similar in magnitude and significance, we should conclude that γ_1 and γ_2 capture systematic variation in the effect of treatment on central and non-central individuals, irrespective of whether they are close to targeted or untargeted individuals.

Estimates for these models are presented in *Table A10* in the Online Appendix. We do not find strong evidence that γ_1 and γ_2 coefficients vary systematically. If anything, we obtain many significant γ_2 estimates in spite of the fact that the number of untargeted individuals in each EA is much smaller than the number of targeted individuals – and hence power should be smaller in regression model (6.3). From this we conclude that the benchmark model we have estimated is the most informative for the data we have collected.

7. Discussion

In order to assess the external validity of our findings, we need some understanding of the channels by which treatments affect outcomes. To this purpose, we look for a coherent narrative that can account for the whole body of evidence that we have gathered, with a particular focus on turnout which is our main outcome of interest.

A first possible channel of influence is that credible information about the electoral process increases voter confidence and induces discouraged voters, namely opposition supporters, to vote. If this were true, we would 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 often diffuses along social networks, we would also expect positive diffusion and reinforcement effects.[20] Is this narrative supported by the empirical results? On the positive side, we find some (limited) effects of the treatments on information about voting and on in interest in elections. But if anything the treatments have increased voting for the incumbent and reduced voting for the opposition, and the negative peer effects we document on turnout

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

are hard to reconcile with this narrative.

A second possibility is that the treatments inflame partisan passions and people vote 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. 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 interest about the elections and the clear average effect on turnout. Treatments increase voting for the dominant party, a finding that is difficult to reconcile with the idea that treatment reassured 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. But the negative peer effects are again difficult to reconcile with this narrative.

A third possibility is that our treatments affect voting through social pressure – either directly through treatment nudging, or indirectly through peer-to-peer reinforcement and diffusion 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 this social pressure/civic duty interpretation, notably the robust direct effect of the civic education treatment on turnout. But once again negative peer effects are difficult to explain with this narrative.

This leaves one systematic empirical finding that contradicts all the above interpretations: the negative reinforcement and diffusion effects associated with the treatments. We now propose a conceptual framework that can account for this finding.

7.1. Conceptual framework

To articulate – and check the internal consistency of – our proposed narrative, we introduce a model of voter participation combining several of the features discussed in the literature. The focus is on turnout.[21] We start by making sure that our model incorporates the implicit belief underlying our treatment: namely, that an educational campaign about elections raises the information level of voters; this affects their belief in the fairness and transparency of the electoral process; their interest in the voting process rises as a result; and people increase turnout to reflect their heightened level of information and interest. This causal chain naturally extends to the diffusion of treatment to individuals untargeted by the campaign, and for reinforcement effects among the targeted – i.e., information circulates among people, interest in elections rises, and turnout increases.

To formalize this general idea in a compact manner, we build on the 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 a 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 for individual i (e.g., transport cost, opportunity cost of time, cost of text messaging). To capture non-instrumental motivations 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

[21] Other forms of political participation such as the open letter follow the same logic.

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 effects increase voter participation.

Well-known conceptual difficulties arise when non-instrumental motivations are absent, i.e., when $\frac{\partial U}{\partial x_i} = 0$. Optimal turnout then requires $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right] = \frac{dC}{dx_i}$. When a single vote has little effect on the electoral outcome, as is conceivable for large elections, then $\partial G(x_i, x_{-i})/\partial x_i$ is small and voting is not individually rational unless the marginal cost of participation is close to zero. This paradox dates back at least to *Downs (1957)*.[22] Introducing non-instrumental motives for voting alleviates 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 $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right] = 0$.

The implicit assumption that motivates our treatments is that circulating information affects voters' information set Ω_i and, so doing, increases $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i} + \frac{\partial U}{\partial x_i}\right]$ and induces higher turnout. The model shows that information can increase turnout in two ways: by increasing the probability that i's vote is pivotal ($E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]$); and by strengthening i's non-instrumental motivation for voting ($E_{\Omega_i}\left[\frac{\partial U}{\partial x_i}\right]$). The model also shows that more information can reduce turnout if it lowers the probability of being pivotal, i.e., if it lowers $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]$. For instance, a citizen may decide not to vote if he/she learns that his/her preferred politician is guaranteed to be elected.

Being pivotal is usually understood as influencing who wins the election. This definition of being pivotal makes sense in advanced democracies. Voters in other countries may care about other electoral outcomes, such as turnout or win gap. This is particularly true in electoral autocracies where the winner of the election is often known in advance, and where a low turnout or win gap is seen as disapproving the government. Political disapproval can then be punished through various means, including a lower supply of local public goods. In such political envi-

[22] 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 (2015)* 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. Crucially, Myatt finds that turnout is high under reasonable conditions.

ronment, being pivotal does not mean casting the ballot that determines who wins the election; it means bringing the turnout or win gap above the threshold below which the community faces reprisals. Some prima facie support for this approach can be found in our data: the treatments increased turnout by individuals who voted for the incumbent party.

In the traditional meaning of pivotal voter, the deciding ballot is a precisely defined concept; the only uncertainty is about whether *i*'s vote is decisive. This limits the range of beliefs consistent with voting (e.g., *Myatt, 2015*). In contrast, a pivot based on turnout or win gap is subject to additional uncertainty regarding the minimum level needed to avoid political reprisal. This implies that the proportion of voters who can rationally believe to be pivotal is larger. The pivotal logic remains, however: i is more (less) likely to vote if i receives information that raises (lowers) the likelihood of being pivotal.[23]

Pivotal reasoning predicts that an information treatment changes the level of turnout depending on location-specific beliefs about turnout or win gap. In EAs where one party won the 2004 election by a large margin, voters may rationally expect the win gap to be large in 2009 as well – and thus may have a lower expectation of being pivotal. These beliefs are of course victim to a fallacy of composition, which some voters may realize as a consequence of treatment. We test this prediction in *Table 9*. Results reported in the first column confirm that control EAs where the win gap was large in 2004 have statistically-significant lower average turnout in 2009. We also observe positive interaction effects between the treatments and the win gap in 2004 that are statistically significant: the treatments contributed to reversing the pattern of behavior in control locations, i.e., they led to increases in voter turnout in EAs with larger win gaps. The second column of *Table 9* replaces the 2004 win gap with the 2004 voter turnout. The effects have the same sign but are not statistically significant, suggesting that voters care mainly about the win gap.

Pivotal reasoning implies that an information treatment can affect people differently depending on how aware they are about others' turnout intentions. To illustrate, take $N + 1$ voters

[23] The pivotal logic can take an instrumental interpretation at the individual level. But it can also work at the collective level: since reaching a target turnout or win gap generates a local public good, voting can be regarded as contributing to that public good. Social pressure can then be applied to induce individual contributions so that the collective target is reached.

arranged in a star shaped network. The center of the star represents a subject who is more central in a social or geographical sense. Let us assume that each voter observes the voting intentions of its immediate neighbor. The star center thus observes the voting intentions of the N other voters who, in contrast, only observe the intention of the star center. The impact of the treatments in this network (assuming for simplicity that all nodes are targeted) depends on whether each node believes to be pivotal. If the star center is more aware of the positive average effects of the campaign on the win gap, he/she is less likely to believe being pivotal than the spokes believe themselves to be.[24] Hence, free-riding through pivotal reasoning implies that the treatment should result in a lower turnout propensity for the star center than for spoke voters. Note that the same effect could be expected if the star center finds, from his/her privileged position and corresponding information, independently of the behavior by others, that he/she was simply overestimating the probability of being pivotal.

This prediction is consistent with our findings. In control EAs, subjects that are more central are more likely to vote – possibly because they realize that the large 2004 win gap has disincentivized spoke voters to turn out. This pattern is by and large reversed in treated EAs: treatment induces the average voter to turn out to vote, and possibly as a result, well-informed central subjects need not increase their own turnout as much as in control EAs.

7.2. Pivotal reasoning vs. saturation

Demonstrating that treatment can theoretically reduce turnout among central subjects due to free-riding through pivotal reasoning is not the same as showing that this is behind our result. There may be other explanations. One particularly threatening candidate is the possibility that the negative coefficient of the $T_v \frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ term reflects a voter saturation effect rather than pivotal reasoning. 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,

[24] This example can easily be generalized as follows. Let N_{iv} be the subset of other voters that i observes in village v, and N_{-iv} be the rest. Let T_v denote treatments as before. We have $x_v = \sum_{j \in N_{iv}} x_j + \sum_{j \in N_{-iv}} x_j$ and $\frac{\partial E_{\Omega_i}[x_v]}{\partial T_v} = \frac{N_{iv}}{N_v} E_{\Omega_i} \left[\frac{\partial x_j}{\partial T_v} \right]$, as the observed change in behavior. Since the effect of treatment on turnout is positive, i.e., $\frac{\partial x_j}{\partial T_v} > 0$, it follows that individuals with a larger N_{iv} increase their expectation $E_{\Omega_i}[x_v]$ more than people with a small N_{iv}. Hence they are less likely to see themselves are pivotal, and thus to vote.

and not pivotal reasoning, 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_{iv} represent individual i's propensity to vote in village v in the absence of treatment. We now assume that voter turnout among the targeted follows the following model:

$$y_{iv} = P_{iv} + \beta T_v + \gamma T_v \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \pi P_{iv} T_v + \varepsilon_{iv} \tag{7.2}$$

where γ captures pivotal reasoning as before, and a significantly negative π coefficient indicates voter saturation. A bias in the previous estimation of γ arises if P_{iv} is correlated with network size $\frac{1}{N_v} \sum_{j \neq i} g_{ijv}$. To demonstrate this, let $P_{iv} = \alpha + \delta \frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ and replace P_{iv} in (7.2):

$$y_{iv} = \alpha + \delta \frac{1}{N_v} \sum_{j \neq i} g_{ij} + (\beta + \alpha\pi) T_v + (\gamma + \delta\pi) T_v \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \varepsilon_{iv} \tag{7.3}$$

Comparing (7.3) with (6.1) it is immediately apparent that voter saturation – a negative π – can be misinterpreted as pivotal reasoning – a negative γ – when estimating regression (6.1).

The solution we propose is to estimate $\widehat{P}_{iv}$ using individuals in untreated locations, and use it as a control function to obtain separate estimates of γ and π. We obtain $\widehat{P}_{iv}$ by estimating a Logit model using control individuals only, while regressing y_{iv} on network size $\frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ and other characteristics known to affect turnout, such as gender and age. Because treatment is assigned randomly, $\widehat{P}_{iv}$ 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 $\widehat{P}_{iv}$ in lieu of P_{iv}. Since by design $T_i = 1$ for the targeted, the estimated regression boils down to:[25]

$$y_{iv} = \beta + (1 + \pi) \widehat{P}_{iv} + \gamma \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \varepsilon_{iv} \tag{7.4}$$

Voter saturation $\pi < 0$ requires the coefficient of $\widehat{P}_i$ to 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

[25] 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.

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 10*. 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 some of them. Furthermore, the point estimates of the coefficient of $\widehat{P}_{iv}$ are in many 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 indicates that saturation is not the explanation for our results, thereby corroborating the hypothesis that free-riding through pivotal reasoning affects the impact of the hotline and civic education treatments.[26]

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 employing specification (6.1) estimate 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 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 the relevance of free-riding through pivotal reasoning as discussed above.

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:

[26] We repeated the analysis for untargeted individuals in treated locations. We do not find significant effects, possibly because of the smaller sample size.

$$y_{iv} = \alpha + \beta T_v + \delta \frac{1}{N_v} \sum_{j \neq i} g_{ijv} + \lambda \frac{1}{N_v} \sum_{j \neq i} g_{ijv} y_{jv} + \varepsilon_{iv}. \tag{7.5}$$

and it is applied to both reinforcement and diffusion effects.[27] If peer effects estimated in regressions (6.1) are driven exclusively by exogenous peer effects, 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, $\frac{1}{N_v} \sum_{j \neq i} g_{ijv} y_{jv}$, with their treatment, $\frac{1}{N_v} \sum_{j \neq i} g_{ijv} T_{jv}$, where T_{jv} is j's condition as directly targeted or untargeted by the treatment.[28] The equation is estimated using two-stage 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 2 are presented in *Table 11*. 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 $\frac{1}{N_v} \sum_{j \neq i} g_{ijv} y_{jv}$ is typically negative, significantly so for reinforcement effects through geographical proximity. This offers some additional support for free-riding through pivotal reasoning.

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

[27] 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 *Bramoullé, 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.

[28] Here $\frac{1}{N_v} \sum_{j \neq i} g_{ijv}$ is computed using all respondents in the EA for whom information on y_{jv} is available, to correct for possible missing observations.

a civic education campaign. The interventions are shown to increase voter turnout and the electoral knowledge of targeted and untargeted individuals in treated locations. 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 with many connections to other surveyed subjects are less likely to vote than similar individuals with fewer connections. This is particularly clear for the hotline treatment. At the same time, interest in politics is 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 through pivotal reasoning: a smaller treatment effect on turnout among central individuals results from realizing 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 treatment effects on soft outcomes such as interest in elections, they can attenuate turnout by circulating information about voting intentions, thereby triggering free-riding through pivotal reasoning. However, we must emphasize that these findings may be specific to electoral autocracies such as 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, 2014*). Since FRELIMO had an overwhelming dominance in these elections, raised partisanship may have mobilized FRELIMO voters disproportionately, hence leading to the peer effects we document.

9. References

Aker, J., P. Collier, and P. C. Vicente (2016): "Is Information Power? Using Cell Phones and Free Newspapers during an Election in Mozambique," *Review of Economics and Statistics*,

forthcoming.

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 (2014): "War's Enduring Effects on the Development of Egalitarian Motivations and In-group Biases," Psychological Science, 25(1), 47-57.

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. Garrett, 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. (2015): "A 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: Penguin 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 treatment effect on political participation

		Homogeneous effects on targeted respondents						Homogeneous effects on untargeted respondents					
		Turnout				Open letter		Turnout				Open letter	
		Shows finger		Index 2				Shows finger		Index 2			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	coefficient	0.014	0.013	0.033	0.036	0.088*	0.085*	0.143***	0.190***	0.092***	0.120***	0.065	0.068
	standard error	(0.040)	(0.036)	(0.033)	(0.029)	(0.050)	(0.050)	(0.036)	(0.031)	(0.034)	(0.031)	(0.066)	(0.064)
Hotline	coefficient	0.063**	0.049*	0.078***	0.070***	-0.036	-0.030	0.090**	0.080*	0.085**	0.083**	0.007	0.023
	standard error	(0.026)	(0.027)	(0.024)	(0.025)	(0.035)	(0.034)	(0.043)	(0.045)	(0.033)	(0.035)	(0.047)	(0.046)
Civic education	coefficient	0.055*	0.046	0.050**	0.050**	0.043	0.048	0.057	0.102**	0.059	0.085**	0.106	0.100
	standard error	(0.029)	(0.029)	(0.026)	(0.025)	(0.048)	(0.047)	(0.048)	(0.047)	(0.039)	(0.039)	(0.070)	(0.074)
Mean dep. variable (control)		0.807	0.805	0.757	0.756	0.153	0.151	0.807	0.805	0.757	0.756	0.153	0.151
R-squared adjusted		0.014	0.027	0.036	0.069	0.013	0.025	0.015	0.077	0.039	0.078	0.015	,037
No. of observations		953	943	953	943	973	962	437	430	437	430	449	441
h0: Newspaper = Hotline	F-stat p-value	0.206	0.300	0.163	0.234	0.012	0.019	0.282	0.023	0.870	0.360	0.434	0.520
h0: Newspaper = Civic education	F-stat p-value	0.310	0.386	0.610	0.657	0.454	0.537	0.111	0.067	0.472	0.410	0.649	0.715
h0: Hotline = Civic education	F-stat p-value	0.774	0.921	0.266	0.407	0.094	0.079	0.569	0.697	0.563	0.975	0.199	0.329
h0: all treatments = 0	F-stat p-value	0.078	0.236	0.016	0.039	0.055	0.062	0.001	0.000	0.012	0.001	0.390	0.480
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. All regressions include province dummies. In the second column for each outcome we control for demographic characteristics (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area) and enumeration area characteristics (has a post office, has a health center). Standard errors are clustered at the enumeration area level.

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

Table 2: Average treatment effect on electoral 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	coefficient	0.122*	0.159**	-0.011	-0.014	0.125	0.174**	0.005	-0.075
	standard error	(0.067)	(0.064)	(0.078)	(0.080)	(0.099)	(0.085)	(0.143)	(0.143)
Hotline	coefficient	0.151**	0.167***	0.111	0.100	0.185**	0.254***	0.081	0.040
	standard error	(0.062)	(0.056)	(0.076)	(0.073)	(0.085)	(0.084)	(0.102)	(0.105)
Civic education	coefficient	0.066	0.159**	0.073	0.098	0.116	0.156*	-0.011	-0.096
	standard error	(0.066)	(0.063)	(0.058)	(0.064)	(0.102)	(0.081)	(0.131)	(0.127)
Mean dep. variable (control)		0.000	0.000	-0.000	0.006	0.000	0.000	-0.000	0.006
R-squared adjusted		0.080	0.243	0.117	0.126	0.082	0.299	0.184	0.214
No. of observations		976	965	976	965	453	445	454	446
h0: Newspaper = Hotline	F-stat p-value	0.647	0.886	0.156	0.162	0.589	0.429	0.645	0.500
h0: Newspaper = Civic education	F-stat p-value	0.395	0.996	0.233	0.120	0.940	0.859	0.930	0.902
h0: Hotline = Civic education	F-stat p-value	0.163	0.881	0.572	0.966	0.541	0.334	0.555	0.378
h0: all treatments = 0	F-stat p-value	0.090	0.019	0.310	0.220	0.145	0.011	0.881	0.802
Controls		No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions (1) to (4) include observations for targeted (in treated locations) and control respondents. Regressions (5) to (8) include observations for untargeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. The dependent variables are indices. All regressions include province dummies. In the second column for each outcome we control for demographic characteristics (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area) and enumeration area characteristics (has a post office, has a health center). Standard errorsare clustered at the enumeration area level.

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

Table 3: Average treatment effect on official electoral results at the ballot-station level

		Presidential elections			Parliamentary elections		
		Turnout	% votes for Guebuza	% votes for Dhlakama	Turnout	% votes FRELIMO	% votes RENAMO
		(1)	(2)	(3)	(4)	(5)	(6)
Newspaper	coefficient	0.054**	0.041**	-0.016	0.055**	0.036*	-0.021
	standard error	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.016)
Hotline	coefficient	0.053**	0.022	-0.012	0.056**	0.020	-0.015
	standard error	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.016)
Civic education	coefficient	0.053**	0.046**	-0.032**	0.052**	0.039*	-0.038**
	standard error	(0.025)	(0.020)	(0.016)	(0.025)	(0.021)	(0.016)
Mean dep. variable (control)		0.440	0.723	0.114	0.438	0.722	0.136
Adjusted R-squared		0.383	0.670	0.569	0.373	0.665	0.627
No. of observations		161	161	161	161	161	161
h0: Newspaper = Hotline	F-stat p-value	0.989	0.233	0.236	0.875	0.370	0.155
h0: Newspaper = Civic education	F-stat p-value	0.981	0.807	0.362	0.887	0.888	0.318
h0: Hotline = Civic education	F-stat p-value	0.970	0.358	0.803	0.990	0.462	0.693
h0: all treatments = 0	F-stat p-value	0.079	0.094	0.272	0.071	0.223	0.123

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

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

Table 4: Network effects on turnout index 2

		Reinforcement effect (targeted vs. control)			Diffusion effect (untargeted vs. control)		
		Chatting	**Kinship**	**Proximity**	**Chatting**	**Kinship**	**Proximity**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	coefficient	0.041	0.035	0.022	0.124***	0.122***	0.114***
	standard error	(0.028)	(0.029)	(0.033)	(0.032)	(0.031)	(0.031)
Hotline	coefficient	0.073***	0.069***	0.070***	0.086**	0.088**	0.086**
	standard error	(0.025)	(0.025)	(0.026)	(0.036)	(0.036)	(0.034)
Civic education	coefficient	0.052**	0.048*	0.050*	0.092**	0.093**	0.066
	standard error	(0.025)	(0.025)	(0.028)	(0.038)	(0.037)	(0.043)
Network	coefficient	0.248***	0.197**	0.053*	0.233***	0.203***	0.081**
	standard error	(0.075)	(0.078)	(0.029)	(0.074)	(0.074)	(0.032)
Network x Newspaper	coefficient	**-0.120**	**-0.153**	**-0.058**	**-0.175**	**-0.148**	**-0.134****
	standard error	(0.098)	(0.118)	(0.042)	(0.120)	(0.145)	(0.060)
Network x Hotline	coefficient	**-0.195****	**-0.235****	**-0.044**	**-0.260****	**-0.324****	**-0.073**
	standard error	(0.099)	(0.100)	(0.034)	(0.121)	(0.135)	(0.051)
Network x Civic education	coefficient	**-0.138**	**-0.205***	**-0.051***	**-0.146**	**0.056**	**-0.021**
	standard error	(0.114)	(0.117)	(0.031)	(0.118)	(0.149)	(0.058)
Magnitude of network effects (a):							
Newspaper		-0.026	-0.018	-0.071	-0.038	-0.016	-0.157**
Hotline		-0.043**	-0.027**	-0.054	-0.057**	-0.035**	-0.085
Civic education		-0.030	-0.024*	-0.063*	-0.032	0.006	-0.025
Mean dep. variable (control)		0.756	0.756	0.755	0.756	0.756	0.755
Adjusted R-squared		0.080	0.070	0.077	0.088	0.080	0.095
No. of observations		943	943	800	430	430	364
Control variables		Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions (1) to (3) include observations for targeted (in treated locations) and control respondents; regressions (4) to (6) 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 (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

(a) For chatting and kinship, the reported network effect is the difference in average treatment effect between an individual with an average size network and an individual with no network at all. For geographical proximity, the reported network effect is the difference in average treatment effect between an individual with maximum (i.e., zero) proximity and an individual with average proximity.

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

Table 5: Summary of interaction coefficients for the remaining turnout regressions

			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**	coefficient	-0.057	-0.183	-0.077*	-0.153	-0.183	-0.126**
		standard error	(0.111)	(0.123)	(0.041)	(0.123)	(0.148)	(0.060)
	Network x Hotline	coefficient	-0.203*	-0.278***	-0.045	-0.274**	-0.396***	-0.090**
		standard error	(0.106)	(0.099)	(0.035)	(0.117)	(0.126)	(0.037)
	Network x Civic education	coefficient	-0.144	-0.212*	-0.052	-0.081	-0.057	-0.042
		standard error	(0.121)	(0.114)	(0.033)	(0.143)	(0.171)	(0.060)
Adjusted	**Network x Newspaper**	coefficient	-0.234	0.014	-0.058	-0.101	0.241	-0.250***
		standard error	(0.157)	(0.173)	(0.064)	(0.210)	(0.219)	(0.095)
	Network x Hotline	coefficient	-0.220	-0.140	-0.043	-0.341	-0.240	0.013
		standard error	(0.171)	(0.176)	(0.061)	(0.212)	(0.209)	(0.115)
	Network x Civic education	coefficient	-0.117	-0.108	-0.074	-0.303	0.125	0.097
		standard error	(0.184)	(0.195)	(0.052)	(0.240)	(0.226)	(0.099)
Shows finger	**Network x Newspaper**	coefficient	-0.004	-0.199	-0.098*	-0.192	-0.171	-0.095
		standard error	(0.128)	(0.170)	(0.052)	(0.129)	(0.148)	(0.062)
	Network x Hotline	coefficient	-0.108	-0.209	-0.064*	-0.372**	-0.272*	-0.062
		standard error	(0.148)	(0.160)	(0.038)	(0.183)	(0.159)	(0.067)
	Network x Civic education	coefficient	-0.124	-0.167	-0.088**	-0.098	-0.059	-0.110
		standard error	(0.149)	(0.147)	(0.037)	(0.148)	(0.181)	(0.075)
Index 1	**Network x Newspaper**	coefficient	-0.088	-0.156	-0.058	-0.173	-0.200	-0.131**
		standard error	(0.100)	(0.124)	(0.041)	(0.118)	(0.136)	(0.057)
	Network x Hotline	coefficient	-0.196*	-0.243**	-0.042	-0.261**	-0.307**	-0.096**
		standard error	(0.102)	(0.100)	(0.032)	(0.122)	(0.135)	(0.046)
	Network x Civic education	coefficient	-0.168	-0.224*	-0.058**	-0.113	0.026	-0.026
		standard error	(0.115)	(0.117)	(0.030)	(0.119)	(0.157)	(0.055)
Interviewer assessment	**Network x Newspaper**	coefficient	-0.068	-0.225	-0.078*	-0.151	-0.054	-0.111*
		standard error	(0.112)	(0.146)	(0.043)	(0.127)	(0.128)	(0.061)
	Network x Hotline	coefficient	-0.161	-0.245**	-0.047	-0.206	-0.325**	-0.084
		standard error	(0.116)	(0.122)	(0.034)	(0.148)	(0.151)	(0.054)
	Network x Civic education	coefficient	-0.229	-0.325*	-0.062*	-0.147	-0.230	-0.059
		standard error	(0.140)	(0.166)	(0.032)	(0.140)	(0.180)	(0.070)

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 (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

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

Table 6: Network effects on posting the 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	coefficient	0.085*	0.084*	0.087	0.069	0.083	0.051
	standard error	(0.050)	(0.051)	(0.054)	(0.064)	(0.066)	(0.065)
Hotline	coefficient	-0.030	-0.032	-0.028	0.020	0.011	-0.012
	standard error	(0.034)	(0.035)	(0.037)	(0.047)	(0.048)	(0.052)
Civic education	coefficient	0.049	0.047	0.048	0.099	0.082	0.123
	standard error	(0.047)	(0.047)	(0.052)	(0.073)	(0.066)	(0.085)
Network	coefficient	0.022	0.064	0.046	0.047	0.076	0.085***
	standard error	(0.116)	(0.141)	(0.033)	(0.106)	(0.145)	(0.032)
Network x Newspaper	coefficient	**-0.034**	**-0.078**	**-0.026**	**0.211**	**0.437**	**0.079**
	standard error	(0.157)	(0.188)	(0.064)	(0.259)	(0.484)	(0.083)
Network x Hotline	coefficient	**0.120**	**-0.047**	**-0.015**	**0.117**	**0.250**	**0.023**
	standard error	(0.153)	(0.198)	(0.037)	(0.203)	(0.248)	(0.059)
Network x Civic education	coefficient	**-0.134**	**-0.289***	**0.035**	**-0.205**	**-0.790****	**-0.016**
	standard error	(0.144)	(0.167)	(0.058)	(0.188)	(0.322)	(0.118)
Magnitude of network effects (a):							
Newspaper		-0.008	-0.009	-0.031	0.047	0.048	0.093
Hotline		0.026	-0.005	-0.018	0.026	0.027	0.027
Civic education		-0.029	-0.034*	0.042	-0.045	-0.087**	-0.018
Mean dep. variable (control)		0.151	0.151	0.158	0.151	0.151	0.158
Adjusted R-squared		0.024	0.025	0.026	0.034	0.047	0.046
No. of observations		962	962	817	441	441	373
Controls		Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions(1) to (3) include observations for targeted (in treated locations) and control respondents; regressions (4) to (6) 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 (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

(a) For chatting and kinship, the reported network effect is the difference in average treatment effect between an individual with an average size network and an individual with no network at all. For geographical proximity, the reported network effect is the difference in average treatment effect between an individual with maximum (i.e., zero) proximity and an individual with average proximity.

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

Table 7: Network effects on the electoral information index

		Reinforcement effect (targeted vs. control)			Diffusion effect (untargeted vs. control)		
		Chatting	**Kinship**	**Proximity**	**Chatting**	**Kinship**	**Proximity**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	coefficient	0.159**	0.161**	0.142**	0.174**	0.167**	0.184*
	standard error	(0.065)	(0.064)	(0.071)	(0.084)	(0.085)	(0.096)
Hotline	coefficient	0.166***	0.168***	0.144**	0.261***	0.265***	0.240**
	standard error	(0.056)	(0.055)	(0.066)	(0.081)	(0.082)	(0.100)
Civic education	coefficient	0.159**	0.161***	0.147**	0.161**	0.162**	0.187**
	standard error	(0.062)	(0.061)	(0.071)	(0.081)	(0.082)	(0.085)
Network	coefficient	-0.093	-0.402*	-0.023	-0.064	-0.398	0.020
	standard error	(0.224)	(0.244)	(0.066)	(0.233)	(0.244)	(0.069)
Network x Newspaper	coefficient	**0.173**	**0.518**	**-0.088**	**-0.002**	**0.203**	**-0.102**
	standard error	(0.294)	(0.347)	(0.081)	(0.375)	(0.571)	(0.140)
Network x Hotline	coefficient	**0.090**	**0.425**	**0.084**	**-0.140**	**0.028**	**0.014**
	standard error	(0.269)	(0.308)	(0.077)	(0.406)	(0.434)	(0.125)
Network x Civic education	coefficient	**-0.054**	**0.047**	**-0.029**	**0.345**	**0.693**	**-0.078**
	standard error	(0.263)	(0.290)	(0.074)	(0.415)	(0.432)	(0.136)
Magnitude of network effects (a):							
Newspaper		0.038	0.060	-0.108	-0.001	0.022	-0.120
Hotline		0.020	0.049	0.102	-0.031	0.003	0.017
Civic education		-0.012	0.005	-0.036	0.076	0.075	-0.092
Mean dep. variable (control)		0.000	0.000	0.009	0.000	0.000	0.009
Adjusted R-squared		0.242	0.247	0.248	0.295	0.302	0.297
No. of observations		965	965	820	445	445	377
Controls		Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions (1) to (3) include observations for targeted (in treated locations) and control respondents; regressions (4) to (5) 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 (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

(a) For chatting and kinship, the reported network effect is the difference in average treatment effect between an individual with an average size network and an individual with no network at all. For geographical proximity, the reported network effect is the difference in average treatment effect between an individual with maximum (i.e., zero) proximity and an individual with average proximity.

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

Table 8: 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	coefficient	-0.002	-0.011	-0.018	-0.071	-0.035	-0.127
	standard error	(0.078)	(0.078)	(0.084)	(0.136)	(0.136)	(0.149)
Hotline	coefficient	0.106	0.103	0.131	0.048	0.050	0.068
	standard error	(0.071)	(0.071)	(0.082)	(0.103)	(0.102)	(0.120)
Civic education	coefficient	0.102	0.104*	0.133*	-0.084	-0.101	0.010
	standard error	(0.062)	(0.061)	(0.068)	(0.125)	(0.134)	(0.135)
Network	coefficient	-0.232	-0.374	0.131*	-0.159	-0.141	0.097
	standard error	(0.174)	(0.249)	(0.076)	(0.170)	(0.289)	(0.073)
Network x Newspaper	coefficient	**0.977*****	**0.984****	**-0.152**	**1.255*****	**1.711*****	**0.336***
	standard error	(0.296)	(0.418)	(0.102)	(0.430)	(0.606)	(0.193)
Network x Hotline	coefficient	**0.579****	**0.732****	**-0.107**	**0.099**	**0.053**	**-0.185**
	standard error	(0.237)	(0.347)	(0.140)	(0.451)	(0.687)	(0.195)
Network x Civic education	coefficient	**0.513****	**0.195**	**-0.190****	**0.651**	**-0.345**	**0.016**
	standard error	(0.220)	(0.374)	(0.086)	(0.439)	(0.839)	(0.157)
Magnitude of network effects (a):							
Newspaper		0.214***	0.114**	-0.186	0.276***	0.186***	0.397*
Hotline		0.127**	0.085**	-0.131	0.022	0.006	-0.218
Civic education		0.112**	0.023	-0.232**	0.143	-0.038	0.019
Mean dep. variable (control)		0.006	0.006	-0.011	0.006	0.006	-0.011
Adjusted R-squared		0.163	0.158	0.137	0.222	0.219	0.233
No. of observations		965	965	820	446	446	378
Controls		Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions (1) to (3) include observations for targeted (in treated locations) and control respondents; regressions (4) to (6) 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 (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

(a) For chatting and kinship, the reported network effect is the difference in average treatment effect between an individual with an average size network and an individual with no network at all. For geographical proximity, the reported network effect is the difference in average treatment effect between an individual with maximum (i.e., zero) proximity and an individual with average proximity.

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

Table 9: Treatment effect on turnout index 2 by 2004 win gap and turnout

		Win gap	Turnout
		(1)	(2)
Newspaper	coefficient	0.044	0.037
	standard error	(0.028)	(0.028)
Hotline	coefficient	0.077***	0.069***
	standard error	(0.025)	(0.025)
Civic education	coefficient	0.059**	0.046*
	standard error	(0.026)	(0.025)
2004 turnout/win gap between Frelimo and Renamo	coefficient	-0.044**	-0.018
	standard error	(0.022)	(0.019)
2004 turnout/win gap x Newspaper	coefficient	0.058*	0.036
	standard error	(0.031)	(0.034)
2004 turnout/win gap x Hotline	coefficient	0.046*	0.002
	standard error	(0.026)	(0.021)
2004 turnout/win gap x Civic education	coefficient	0.035	-0.006
	standard error	(0.021)	(0.029)
Mean dep. variable (control)		0.756	0.756
Adjusted R-squared		0.070	0.069
No. of observations		943	943
Controls		Yes	Yes

Note: Regressions include observations for targeted (in treated locations) and control respondents. All regressions are OLS and use only second-round data. We control for demographic characteristics (sex, age, single, divorced, protestant, retail informal sector, commerce, professional, teacher, domestic worker, household has enough food, owned house, chitsua ethnic group, lomue language, time living in the enumeration area), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level. The win gap is the difference in 2004 vote shares between the Frelimo and Renamo presidential candidates at the EA level.

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

Table 10: Network effects of the hotline, controlling for saturation

		Panel A: Chatting											
		Turnout measures											
		Self-reported		**Adjusted**		**Shows finger**		**Index 1**		**Index 2**		**Interviewer**	
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Network	coefficient	-0.109	-0.128	-0.167	-0.230	0.089	0.057	-0.280	-0.306	-0.293	-0.311	-0.143	-0.178
	standard error	(0.116)	(0.125)	(0.287)	(0.265)	(0.320)	(0.310)	(0.215)	(0.228)	(0.212)	(0.214)	(0.188)	(0.196)
Predicted propensity to vote	coefficient	0.565	0.601	0.860	0.707	0.350	0.240	1.461	1.543	1.633	1.584	1.059	1.326
	standard error	(0.646)	(0.698)	(0.679)	(0.659)	(0.992)	(0.985)	(1.217)	(1.306)	(1.201)	(1.221)	(0.969)	(1.046)
constant	coefficient	0.472	0.385	0.268	0.358	0.567	0.551	-0.334	-0.442	-0.507	-0.482	-0.033	-0.393
	standard error	(0.554)	(0.609)	(0.422)	(0.415)	(0.749)	(0.742)	(1.025)	(1.105)	(1.013)	(1.037)	(0.819)	(0.890)
Adjusted R-squared		0.004	-0.004	0.019	0.058	0.011	0.018	0.031	0.027	0.041	0.044	0.011	0.051
No. of observations		230	229	230	229	230	229	230	229	230	229	230	229
		Panel B: Kinship											
		Turnout measures											
		Self-reported		**Adjusted**		**Shows finger**		**Index 1**		**Index 2**		**Interviewer**	
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Network	coefficient	-0.032	-0.080	-0.146	-0.282	-0.151	-0.248	-0.215	-0.259	-0.272	-0.292	-0.084	-0.190
	standard error	(0.117)	(0.136)	(0.307)	(0.309)	(0.324)	(0.319)	(0.213)	(0.237)	(0.201)	(0.213)	(0.192)	(0.212)
Predicted propensity to vote	coefficient	0.055	0.052	0.880	0.721	0.715	0.666	0.767	0.760	1.056	0.904	0.504	0.802
	standard error	(0.430)	(0.479)	(1.026)	(1.054)	(0.938)	(0.939)	(0.896)	(0.985)	(0.857)	(0.869)	(0.726)	(0.811)
constant	coefficient	0.899**	0.840**	0.232	0.324	0.306	0.238	0.226	0.185	-0.045	0.059	0.421	0.031
	standard error	(0.370)	(0.417)	(0.681)	(0.690)	(0.738)	(0.741)	(0.753)	(0.830)	(0.723)	(0.735)	(0.613)	(0.689)
Adjusted R-squared		-0.009	-0.015	0.011	0.057	-0.002	0.013	0.004	-0.001	0.014	0.020	-0.004	0.032
No. of observations		230	229	230	229	230	229	230	229	230	229	230	229
		Panel C: Geographical proximity											
		Turnout measures											
		Self-reported		**Adjusted**		**Shows finger**		**Index 1**		**Index 2**		**Interviewer**	
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Network	coefficient	-0.082	-0.083	-0.145	-0.148	0.055	0.062	-0.143	-0.170	-0.152	-0.218**	-0.257**	-0.175*
	standard error	(0.068)	(0.074)	(0.183)	(0.189)	(0.102)	(0.107)	(0.094)	(0.103)	(0.099)	(0.106)	(0.104)	(0.103)
Predicted propensity to vote	coefficient	1.556	1.525	1.571	1.337	-0.784	-0.790	2.251*	2.449*	2.343	2.921**	3.355**	2.498*
	standard error	(1.076)	(1.121)	(1.377)	(1.439)	(1.434)	(1.527)	(1.361)	(1.470)	(1.431)	(1.427)	(1.398)	(1.418)
constant	coefficient	-0.514	-0.558	-0.433	-0.377	1.569	1.444	-1.252	-1.502	-1.370	-1.963	-2.354*	-1.667
	standard error	(0.997)	(1.038)	(1.160)	(1.204)	(1.254)	(1.318)	(1.273)	(1.363)	(1.341)	(1.330)	(1.314)	(1.324)
Adjusted R-squared		0.020	0.011	0.025	0.100	-0.003	-0.001	0.023	0.030	0.020	0.056	0.020	0.026
No. of observations		192	192	192	192	192	192	192	192	192	192	192	192
Controls		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions only include observations for respondents targeted by the hotline. All regressions are OLS and use only second-round data. In the second column for each outcome we control for demographic characteristics (married or in an union, completed 5 years of primary education and household size) and province dummies. Predicted propensity to vote was obtained by regressing the respective turnout measure on the individual's network size, gender and age, using only the sample of individuals in control areas and using a logit specification. Panel A presents the regressions with chatting measure. Panel B presents the regressions with kinship measure. Panel C presents the regressions with geographical proximity. Standard errors were obtained using bootstrapping.

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

Table 11: Endogenous network effects on turnout index 2

		Reinforcement effect (targeted vs. control)			Diffusion effect (untargeted vs. control)		
		Chatting	**Kinship**	**Proximity**	**Chatting**	**Kinship**	**Proximity**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	coefficient	0.046	0.039	0.018	0.135***	0.121***	0.103**
	standard error	(0.042)	(0.036)	(0.024)	(0.049)	(0.038)	(0.041)
Hotline	coefficient	0.104**	0.088***	0.049*	0.128***	0.112***	0.092**
	standard error	(0.045)	(0.032)	(0.027)	(0.042)	(0.036)	(0.045)
Civic education	coefficient	0.069*	0.047	0.031	0.106**	0.079*	0.078*
	standard error	(0.039)	(0.034)	(0.025)	(0.047)	(0.042)	(0.042)
Network	coefficient	1.930	2.435	0.285**	1.949	1.759*	-0.078
	standard error	(2.073)	(2.071)	(0.140)	(1.301)	(1.038)	(0.404)
Endogenous network effect	coefficient	**-2.208**	**-2.917**	**-0.333***	**-2.240**	**-1.985**	**0.176**
	standard error	(2.558)	(2.527)	(0.184)	(1.563)	(1.251)	(0.565)
constant	coefficient	0.596***	0.642***	0.706***	0.579***	0.634***	0.665***
	standard error	(0.064)	(0.053)	(0.046)	(0.081)	(0.065)	(0.077)
No. of observations		946	946	946	434	434	434
Controls		Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions (1) to (3) include observations for targeted (in treated locations) and control respondents; regressions (4) to (6) include observations for untargeted (in treated locations) and control respondents. All regressions are IV and use only second-round data. We control for demographic characteristics (sex, age, married or in an union), enumeration area characteristics (has a post office, has a health center) and province dummies. Standard errors are clustered at the enumeration area level.

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