Interpersonal influence and network effects on voting behavior: Experimental evidence from Mozambique

Ana Vaz
St. Antony's College
University of Oxford

A thesis submitted in fulfillment of
the requirements for the degree of
Doctor of Philosophy in Economics

Hilary Term, 2014

Approximate Word Count: 63.000

Abstract

Interpersonal influence and network effects on voting behavior:

Experimental evidence from Mozambique

Ana Vaz, St. Antony's College

D.Phil. Thesis, Michaelmas Term 2013

During the 2009 Mozambican elections a randomized controlled trial implemented a voter education campaign. My thesis analyses the peer effects triggered by that campaign. The first essay looks at peer effects within the household. The second essay (co-authored with Marcel Fafchamps and Pedro Vicente) focuses on peer effects at the village level. The campaign targeted heads of household or their spouse (i.e., the primary target). In the first essay (Chapter 2) I test the effect of the campaign on other household members. I interpret this effect as evidence of interpersonal influence exerted by the primary target over other household members. The main finding of this analysis is that the transmission of voter education campaigns' effects tends to occur through sharing of opinions and social pressure, instead of information sharing. In this essay I also explore the determinants of interpersonal influence within the household. I test whether the effect of the campaign on other household members (i.e., secondary targets) varies with age, gender and education. I find a stronger effect on younger secondary targets, consistent with the idea that they are more susceptible to social pressure. The second essay (Chapter 3) examines whether the campaign's effect is transmitted within the village through social networks (kinship and chatting) and geographical proximity. We test whether the impact of the campaign on targeted and untargeted individuals depends on proximity to other targeted individuals in the village. Our main finding is that the campaign increases voter participation on average, but much less so among individuals who are socially or geographically closer to other targeted individuals. This result is interpreted as evidence of free-riding on voting as a civic duty.

Acknowledgements

I am very grateful to my supervisors. I wish to thank Marcel Fafchamps for all his suggestions and for always giving me very detailed feedback, and I wish to thank Sabina Alkire for her support and for the opportunity to work on interesting projects as her research assistant.

I am forever indebted to Pedro Vicente. This thesis was only possible due to his generosity and encouragement. I would also like to extend a word of thanks to the team with whom Pedro and I worked in Mozambique, especially to Jenny Aker, Miguel Ferreira, Alberto da Cruz, Egídio Chaimite, Aquílcia Samuel and Egídio Guambe.

I wish also to thank James Foster, Stephen Smith, Chrysanthi Hatzimasoura and Kyle Renner, who made me feel at home during my stay at George Washington University.

I acknowledge financial support from the Fundação para a Ciência e a Tecnologia (doctoral grant SFRH/BD/35943/2007).

Contents

Chapter 1

Introduction

Voting behavior has been the object of vast research in the past decades. However, most of this research was focused on developed countries where there is a long history of elections. This type of research is in its early stages in developing countries. In some of these developing countries elections are a relatively new phenomenon. In these new democracies the understanding of voting behavior is especially important, in order to inform policies that may ensure the healthy development of democracy and accountability mechanisms. This thesis contributes to this new strand of literature by documenting the impact of social influence on voting behavior in Mozambique.

The two empirical essays of this thesis examine the peer effects on voting behavior triggered by a voter education campaign. The first essay (Chapter 2) focus on the peer effects within the household; the second essay (Chapter 3) examines the peer effects at the village level.

During the 2009 Mozambican general elections a randomized controlled trial implemented a voter education campaign aimed at increasing voter participation. This campaign assumed three distinct forms. The first was the distribution of a free newspaper, which provided neutral information about the elections. The second was the creation and advertising of a text messaging hotline to report electoral problems. This intervention aimed at involving individuals in the electoral process, allowing them to contribute to the process. The third was a civic education campaign based on a leaflet about the voting procedure and text messages calling out the vote and providing information about the elections. This intervention combined an information component

with a nudge towards voting.

The campaign targeted the household head or his/her spouse. The post-treatment survey collected data regarding the experimental subject and a second adult living in the same household (hereafter called second interviewee). This was done to study the campaign spillover effects within the household. The experiment was also designed in order to identify targeted and untargeted individuals living in treated villages, as in each village the treatment was randomized across respondents/households. The purpose of this was to allow the studying of the diffusion effects to untargeted respondents within the village. During the surveys, we collected data to measure electoral behavior, interest and information about politics. We also gathered detailed information about social and geographic connectedness between individuals, including measures of chatting, kinship and geographical distance between respondents' houses.

In the first essay I focus on the role of interpersonal influence in the diffusion of the effects of the voter education campaign within the household. *Aker, Collier and Vicente (2013)*[1] have shown that all interventions have a positive impact on main respondents' turnout. Here I examine the campaign's effect on untreated individuals who cohabit with the main respondents (the second interviewees). A significant diffusion effect is interpreted as evidence of interpersonal influence within the household. I find different diffusion effects associated with different interventions and interpret that as evidence that different interventions trigger influence at different levels. The delivery of the newspaper seems to have, on average, no effect on other people in the household. The hotline intervention affects the interest and turnout of other individuals in the household. The civic education campaign only affects the turnout of other people in the household. This essay argues that the transmission of the voter education campaign's effect does not occur through information sharing, but through sharing of opinions and pressure.

Having established that the campaign's effect is transmitted within the household through interpersonal influence, I explore the determinants of such influence. I test whether the effect of the campaign on other household members varies with age, gender and education. Significant results are interpreted as evidence that interpersonal influence within the household depends on the characteristic being tested. I look at the characteristics of both the influencer (the experimental subject) and the target of influence (the second interviewee). I find that the indirect

[1] Article listed in the references section of Chapter 2.

effect of the campaign on individual turnout is stronger for younger second interviewees than for older ones. This is interpreted as evidence that younger individuals are more likely to respond to social pressure towards voting than older ones.

The second essay (co-authored with Marcel Fafchamps and Pedro Vicente) examines the peer effects at the village level. After studying the indirect effects of the voter education campaign within the household, this paper moves on to the analysis of the indirect effects within the treated villages. We start by estimating the average effects of the campaign on targeted and untargeted respondents (these results are explored in detail in *Aker, Collier and Vicente, 2013*). Then, we investigate whether the impact of the campaign on targeted and untargeted individuals depends on social and geographical proximity to (other) targeted individuals in the village. In simple terms, we estimate the heterogeneous effect of the campaign associated with the individual's network centrality. We estimate this network effect separately for targeted and untargeted respondents. The impact on targeted individuals is called a reinforcement effect; and the impact on untargeted subjects is called a diffusion effect.

We find that the campaign raises average voter participation and level of information about elections of targeted and untargeted respondents. This means that the reach of the voter education campaign goes beyond the treated households. The voting behavior of individuals living in households that did not receive the treatment is also impacted by the campaign.

The reinforcement and diffusion effects on political participation are negative across treatments and network measures. This means that the impact of the campaign on individual turnout is smaller among more connected individuals. On the other hand, we observe positive network effects on information and interest.

We argue that the campaign raises civic-mindedness and, hence, leads to an increase in participation. However, as more central individuals realize that others are more likely to vote and that electoral competition will diminish (given the large advantage of the incumbent), they feel less compelled to go vote. More central individuals free ride on others' participation. This effect does not occur in the case of information and interest probably because they do not involve any costly action.

Chapter 2

Interpersonal influence towards voting within Mozambican households

Voter education promotes voters' participation. One question that remains is how voter education campaigns can reach a significant part of the population. During the 2009 Mozambican elections, a field experiment implemented three voter education interventions: the distribution of a free newspaper, the creation of a SMS hotline to report electoral problems, and a civic education campaign. Based on a sample of untreated individuals living with experimental subjects, this paper examines the diffusion of the interventions' effects within the household. I find different spillover effects associated with different interventions and interpret that as evidence that different interventions trigger influence at different levels. I find that, on average, the delivery of the newspaper has no effect on other people in the household. The hotline intervention affects the preferences and behavior of other individuals, but not their information. The civic education campaign affects only the behavior of other people in the household. This chapter argues that the transmission of voter education

campaigns' effects does not occur through information sharing, but through sharing of opinions and pressure. Furthermore, this study provides evidence that the effect of the campaign on other household members is stronger the younger they are, consistent with the idea that they are more susceptible to social pressure within the household.

2.1 Introduction

Voter education contributes to make elections more democratic.[1] One question that still needs to be answered is how voter education campaigns can reach a significant part of the population in African countries where the access to mass media is still very limited.[2] Fortunately, there are factors that mitigate the effects of the limited access to mass media. First, the mobile phone is becoming a crucial tool for sharing information in the African context, with important repercussions in the political sphere (*Heacock, 2009; Ekine et al., 2010; Aker, Collier and Vicente, 2013*). Second, one of the most efficient means of spreading ideas and opinions is through interpersonal influence. *Lazarsfeld, Berelson and Gaudet (1944)*, when studying the voting behavior during the United States Presidential elections in 1940, found that people would identify personal and informal contacts as sources of influence on the decision to vote much more often than exposure to media.[3] In these circumstances, it is still possible that the information and opinions transmitted by the media reach the majority, even when only a minority has direct access to them.

This paper focuses on the role of interpersonal influence in the diffusion of the effects of voter education campaigns within households. During the 2009 elections in Mozambique, a field experiment implemented three media interventions aimed at increasing political participation: the distribution of a free newspaper, the creation of a text messaging hotline to report electoral

[1] Some studies have shown the benefits of voter education interventions in African countries. *Vicente and Wantchekon (2009)*, based on field experiments in Benin and São Tomé and Principe, argue that civic education campaigns reduce the impact of clientelism and vote buying on voting behavior. *Collier and Vicente (2009)* show that a campaign encouraging the Nigerian voters to oppose electoral violence reduced the impact of intimidation on turnout. *Aker, Collier and Vicente (2011)* show that different types of voter education interventions increase individuals' political participation and their information about politics.

[2] *Temin and Smith (2002)* examine the role of the media in the Ghana's 2002 elections. They find that although the media are very important in urban areas, they are almost irrelevant in rural areas. This happens because the access to media is very, very limited in rural and remote areas.

[3] Based on these evidence, *Katz and Lazarsfeld (1955)* developed the two step flow theory. According to this theory, in a first stage the information from the media reaches a group of individuals who pay more attention to the mass media and, in a second stage, these individuals transmit the media content, as well as their own interpretations, to the people with whom they interact.

problems and an electoral education campaign based on a leaflet and text messages. In the post-elections survey, a second person, aside from the experimental subject, was interviewed in each household. Based on this sample of untreated individuals, this paper provides evidence that the effects of voter education campaigns are transmitted to the people who cohabit with experimental subjects. Furthermore, this paper shows that different interventions trigger influence at different levels. The delivery of the newspaper, although it increases the information level of the individuals who directly received it, it has almost no effect on other people in the household. The hotline intervention, which affects the information, interest and turnout of treated individuals, only affects the interest and turnout of other individuals in the household. The civic education campaign only affects the turnout of treated individuals and other people in the household.

This chapter also explores the determinants of interpersonal influence regarding the decision to vote, within Mozambican households. It finds no robust evidence that gender or age of the experimental subject change the effect of the campaign on second interviewees. In other words, it finds no evidence that those characteristics are related to one's ability to exercise influence over other household members. This paper shows, however, that the indirect effect of the campaign is stronger on younger second interviewees than on older ones. This finding is interpreted as evidence that younger individuals are more likely to respond to social pressure towards voting than older ones.

This paper shows that the transmission of these voter education campaigns' effects does not occur through information sharing, but through sharing of opinions and pressure. This evidence is a bit worrying. One would like to think that this type of interventions would raise citizens' awareness and contribute to enabling them to make an informed voting decision; however, that is not always the case. Civic education intervention seems to work through persuasion: it mainly pressures treated individuals to vote, who in their turn pressure the people around them to vote as well.

The analysis focuses on the relationships within the household, therefore the findings discussed here do not immediately extrapolate to diffusion/influence between individuals in other contexts. Nevertheless, given that a large proportion of the individual's interactions with other people occurs within the household, and that the household is often the target unit of policy interventions, it is very useful to learn about diffusion/influence within the household.

This paper contributes to the literature on voter mobilization, which has been focused almost exclusively on the United States, by extending the analysis to the context of African countries. It relates closely to studies that analyze peer effects of voter interventions like *Nickerson (2008)*, *Gine and Mansuri (2011)* and *Fafchamps and Vicente (2013)*. The first shows that subjecting one individual to a call-out the vote message also increased the propensity to vote of his or her house mate in households with two registered voters. The second estimates the peer effects associated with geographical distance and friendship on female turnout triggered by a randomized voter awareness campaign. The third identifies the reinforcement and diffusion network effects triggered by a randomized campaign against political violence in Nigeria. This paper also links to the experimental research on social pressure as a determinant of electoral participation, namely to the work of *Gerber, Green and Larimer (2008)*, *Davenport (2010)* and *Panagopoulos (2010)*. *Gerber, Green and Larimer (2008)* were the first authors to provide experimental evidence of the impact of social pressure on turnout. *Davenport (2010)* extended their analysis exploring the potential of social pressure to increase the turnout of parts of society that tend to vote less. *Panagopoulos (2010)* provided evidence that appeals to shame are more effective in increasing turnout than appeals to pride.

The analysis of the determinants of influence within the household relates to the literature on bases of interpersonal influence (*Katz, 1957*; *French and Raven, 1959*; *Patchen, 1974*) and to the literature on social networks and diffusion of innovation that investigates what are the common characteristics of opinion leaders (*Kingdon, 1970*; *Booth and Babchuk, 1972*; *Chan and Misra, 1990*) and earlier adopters (*Adcock, Hirschman, and Goldstucker, 1977*; *Dickerson and Gentry, 1983*, *Brancheau and Wetherbe, 1990*). In the sense that this analysis also focuses on the characteristics of the receiver and not exclusively on those of the influencer, it acknowledges the argument put forward by *Watts and Dodds (2007)* that "it is generally the case that most social change is driven not by influentials but by easily influenced individuals influencing other easily influenced individuals" (*p. 442*).

The paper is organized as follows. Section 2.2 begins by presenting a conceptual framework to analyze the indirect effects of the campaign. Section 2.3 describes the testing strategy. The experimental design, the data and the descriptive statistics, including balance tests, are discussed in Section 2.4. Subsequently, Section 2.5 presents the econometric analysis of the average effects

and Section 2.6 presents the study of the heterogeneous effects. Section 2.7 concludes.

2.2 Conceptual framework

2.2.1 Interpersonal influence

The starting point of this analysis is that a voter education campaign, even when targeted at one person per household, tends to affect more people than those who were directly treated (e.g. *Nickerson, 2008*). This happens because the effect of the campaign is likely to spread to the people with whom the treated individual interacts, as she influences them. This section proposes a conceptual framework to examine these spillovers effects.

Human behavior is modeled as the result of a sequential process. First, the individual gathers information; then, based on that information, she develops her preferences; finally, she adopts a behavior according to her preferences. It is assumed that the other people may potentially influence the individual's behavior at those three levels, information, preferences and behavior. First, as the individual has to make decisions in an uncertain environment, she will gather information from as many sources as she can, including people with whom she interacts (*Montgomery and Casterline, 1996*). Second, as the gathering of information generally has costs, in certain circumstances the examples and opinions of others may become enough input for an individual to form her own preferences (*Arthur and Lane, 1993*); in other words, the individual may simply adopt or imitate other people's preferences. Third, other people may 'impose' a determined behavior to the individual through social pressure. Both intentional and unintentional pressure may lead the individual to adopt a behavior that otherwise she would not. [4][5] Under this model it is possible that influence at the level of information may lead to a change of preferences, which may trigger a change in behavior.

This idea is formalized as follows. I assume that individual i takes action x_i (casting a vote)

[4] A similar categorization of social influence was proposed by *Deutsh and Gerard (1955)*, who distinguished two types of social influence: informational and normative. Informational influence is defined "as an influence to accept information obtained from another as evidence about reality" (p. 629); and normative influence is "an influence to conform with the positive expectations of another" (p. 629).

[5] In the context of the interventions discussed in this paper, an example of influence at information level would be the treated individual teaching other household members the names of the presidential candidates. An example of influence at the level of preferences would be the treated individual convincing other household members that it is important to participate in the electoral process. And an example of influence at the behavior level would be the treated individual pressuring other household members to vote.

to maximize the following payoff function:

$$\max_{x_i \geq 0} E\left[U(G(x_i, x_{-i}), x_i; N) | \Omega_i\left(M, F\right)\right] - c\left(x_i\right) - d(\overline{x} - x_i)^2 \tag{2.1}$$

The decision to vote depends, first, on the expected utility of voting, conditional on the individual's information set. The individual derives utility from the outcome of the electoral process ($G(x_i, x_{-i})$) and from the act of voting itself, for instance due to civic mindedness (*Fiorina, 1976; Aldrich, 1993; Engelen, 2006*). Each argument of the utility function captures a different type of motivation to vote: instrumental and intrinsic. The election's outcome depends on the individual's decision to vote (x_i), as well as on the combined action of other individuals (x_{-i}). Motivation to vote may be affected by the opinions of other people. This is reflected in the shift parameter N. The individual's information set (Ω_i) depends on her access to media (M) and on the information she obtains from others (F).[6]

Second, the decision to vote also depends on the material cost of the action (e.g., transport cost), which is captured by $c\left(x_i\right)$; and on the social cost of deviating from behavioral norm $\overline{x}$ (e.g., *Akerlof, 1997*), captured by the expression $d\left(\overline{x} - x\right)^2$. This last element captures the extrinsic motivation to vote.[7]

All the elections in Mozambique have been won by the incumbent party or its candidates, FRELIMO (Frente de Libertação de Moçambique), and the gap between the two main parties has widened (in the parliamentary elections of 2004 the gap ascended to 32.3 percentage points).[8] As the final result of the elections is fairly known, some may argue that Mozambicans hardly vote for instrumental reasons. However, the instrumental value of the vote may not be associated with the victory of the preferred candidate, but with the turnout and the electoral results in the specific community. For example, if individuals believe that the geographical allocation of public

[6] *Montgomery and Casterline (1996)* refer to the information that the individual receives from other people as "social learning" and argue that it is a distinct concept from "social influence", which they relate mostly with the impact of "social conformity" in shaping individual's preferences. Here, I consider the information received from others as a result of influence.

[7] Extrinsic and instrumental motivations are often seen as synonyms. Here I use these terms to refer to different concepts. Extrinsic motivation occurs when the individual votes due to coercion, to please others, or to avoid blame. Instrumental motivation occurs when the individual votes because he hopes to achieve some benefit from the outcome of the elections. On the other hand, intrinsic motivation occurs when the individual votes because he values the act of voting itself. The categorization of motivation that I employ here is inspired in the categorization of extrinsic motivation proposed by the Self-Determination Theory (*Deci and Ryan, 1985*; *Ryan and Deci, 2000*).

[8] Based on results available on www.eisa.org.za.

funds is related with the electoral results at the level of the ballot station, they may go to the pools for instrumental reasons. Nevertheless, even if the instrumental motivation does not apply in the case of Mozambique, the present model still allows us to 'explain' voter participation based on the intrinsic (civic mindedness) and extrinsic (social norms) motivations to vote.

In the context of this model, the voter education campaign may affect the individual's behavior through three channels. First, it may change her information set Ω_i. Second, it may affect her instrumental or intrinsic motivation to vote by, for instance, increasing trust in the electoral process or raising her civic-mindedness (ΔN). And, third, it may create or reinforce social norms about voting, changing the social norm $\overline{x}$ or raising d. The voter education campaign may also trigger peer effects within the household. According to this model, the primary target of the campaign may affect the behavior of other household members in the same way as the campaign affected his own behavior.

According to this model, different types of voter education campaigns are likely to affect the behavior of its targets differently and, therefore, trigger different types of influence. The delivery of the newspaper consists in a basic information treatment. As a result, it should primarily affect the information set of treated individuals. Then, this increase in information may lead to a change in preferences and maybe ultimately affect behavior. In these circumstances, the spillover effect of this treatment is more likely to be related to information outcomes.

The hotline intervention aims at engaging treated individuals in the electoral process, allowing them to contribute to it. The main idea is that treated individuals will feel more empowered and more willing to participate. This is a campaign mostly targeted at preferences. In this case, the treated individuals will be more likely to influence the preferences and maybe the behavior of other household members, than their information set.

The civic education intervention, despite including an information component (providing some elementary electoral information via cell phone messages), consists basically in 'nudging'/persuading treated individuals to vote. Therefore, its effect should be mainly at the level of behavior. Consequently, it is more likely that individuals that received this treatment will influence the behavior preferences and behavior than the information set of other people.

I characterize the expected average effects of each type of intervention based on its main characteristics. It is possible, however, that some specific features of the intervention may trigger

different effects.[9]

The model presented here is built on the calculus of voting model developed by *Downs (1957)* and extended by *Riker and Ordeshook (1968)* and *Fiorina (1976)*, among others. Although there are alternative models to analyze the decision to vote (like for instance the minimax regret model of *Ferejohn and Fiorina, 1975*), the calculus of voting model remains the most widely cited model of decision to vote. Most importantly, this model provides a general and flexible decision-theoretic framework to which I can incorporate new features based on ideas from the social psychology literature in interpersonal influence (*Arthur and Lane, 1993*; *Deutsch and Gerard, 1955*; *French, 1956*; *Fishbein and Ajzen, 1975*; *Montgomery and Casterline, 1996*). In particular, I adapted the model to consider additional motivations to vote and to highlight the potential channels of interpersonal influence.

2.2.2 Determinants of influence

Interpersonal influence can be seen as a directed flow between two people. It can be a flow of information, opinions or instructions. Similarly to other directed flows, individual i's influence over individual j is likely to depend on: (1) the characteristics of i, the influencer; (2) the characteristics of j, the receiver or target of influence; and (3) the characteristics of their connection or link.

It is generally accepted that certain characteristics make some individuals more able to exert influence and that these characteristics may vary across settings (*French and Raven, 1959*; *Patchen, 1974, Kingdon, 1970*; *Booth and Babchuk, 1972*; *Chan and Misra, 1990*). As interpersonal influence may occur at three different levels - information, preferences and behavior - it is possible that the characteristics that make one individual more influential at one level are not the same that make another person more influential at other level. Taking this logic to the extreme, there may be three different types of person: one who is considered a reliable source of information; other who is considered a role model; and, finally, another to who is recognized some sort of authority or hierarchical power. Of course in real life there is not such a clear separation

[9] It is also possible that the effect of the campaign on main respondent's is not transmitted to other household members (for instance, if the effect on main respondents is very small, or if people in the household do not interact).

between these three types of people.

Influence may also depend on the characteristics of the receiver (*Patchen 1974*), at least in two ways. The characteristics of the receiver may affect the likelihood that the individual becomes target of influence; and may determine into what extent the individual assimilates the message sent by the influencer.

Finally, the context and the connection between the influencer and the receiver may impact on the effectiveness of interpersonal influence. For instance, "the communication networks or patterns of interaction in the group" (*French, 1956, p. 181*) are among the factors that mediate influence. Therefore, the specific characteristics of the household may determine the pattern of interaction in the family and hence impact on the influence flow.

2.3 Testing strategy

2.3.1 Average effects

Reduced form specification

I assume that the impact of each intervention can be divided into a direct and an indirect or spillover effect. The direct effect results from the visit by door-to-door interviewers and the reception of leaflets, cell phone messages or newspapers; this effect only affects the individuals treated by the program. The indirect or spillover effect is generated by the interaction with other people who did receive those visits and received the leaflets, cell phone messages or newspapers. In other words, I assume that part of the interventions' effects is transmitted by the treated individuals to the people with whom they interact through interpersonal influence.

The combined (i.e., direct and indirect) average effects of the campaign can be estimated as follows. Let y_i^P, z_i^P, and w_i^P be measures of electoral information, preferences, and behavior, respectively, of the main respondent living in household i. Let $T_i = 1$ if the household was targeted by the campaign and 0 otherwise. The campaign took three distinct forms that I test separately. For the sake of the presentation, here I focus on a single treatment.

Since treatment is randomly assigned, the homogeneous (average) effects of the campaign can

be estimated using regressions of the form:

$$\begin{aligned} y_i^P &= \alpha_1^P + \beta_1^P T_i + u_i^P \\ z_i^P &= \alpha_2^P + \beta_2^P T_i + v_i^P \\ w_i^P &= \alpha_3^P + \beta_3^P T_i + e_i^P \end{aligned} \tag{2.2}$$

These regressions can also be estimated with village or individual fixed effects to net out possible time effects and individual unobservables. Coefficients β_1^P, β_2^P and β_3^P are average treatment effects on information, preferences, and behavior, respectively.

The estimation of average treatment effects is not the focus of this paper. This issue is covered in detail in *Aker, Collier and Vicente (2011)*. Here I focus on estimating the indirect effects of the campaign. With the purpose of constructing a sample of individuals who were only subject to the indirect effects of the interventions, a second person, besides the experimental subject, was interviewed in each household (this person will be referred to as second interviewee). I estimate the average indirect effect of the campaign on these non-treated individuals living in households in treated areas. Let y_i^S, z_i^S, and w_i^S be measures of electoral information, preferences, and behavior, respectively, of the second interviewee living in household i. Given the randomized nature of the data,[10] the spillover (average) effects of the campaign are estimated using the following regressions:

$$\begin{aligned} y_i^S &= \alpha_1^S + \beta_1^S T_i + u_i^S \\ z_i^S &= \alpha_2^S + \beta_2^S T_i + v_i^S \\ w_i^S &= \alpha_3^S + \beta_3^S T_i + e_i^S \end{aligned} \tag{2.3}$$

I also estimate these effects controlling for individual characteristics and geographical location (province dummies).[11] All regressions are estimated using ordinary least squares (OLS) and the

[10] In the sample of second interviewees, the assumption of random selection is two-folded. First, as in the sample of main respondents, the selection of second interviewees into the treatments cannot be correlated with individual's characteristics that affect the outcomes of interest. Second, the selection cannot be correlated with household and respondents' (both main and second) characteristics that affect the spillover effect. Formally, $E\left(y^S\right) = E\left(y^S|T, h^S\right)$, where h^S represents the household and respondents' characteristics that affect the spillover effects.

[11] Following *Deaton's (2009)* advice, I keep the number of covariates as low as possible.

standard errors reported are adjusted by clustering at local level (enumeration area). Coefficients β_1^S, β_2^S and β_3^S are average indirect effects on information, preferences, and behavior, respectively.

In the context of my conceptual framework, a significant effect on an information outcome ($\beta_1^S > 0$) will be interpreted as evidence of influence in the form of sharing information by treated individuals.

The interpretation of a significant effect on preferences' outcomes ($\beta_2^S > 0$) will depend on whether the same treatment had also a significant impact on information or not. If a given treatment has a significant effect on preferences but not on information ($\beta_2^S > 0$ and $\beta_1^S = 0$), it can be inferred that this change resulted from influence directly on preferences. But if a given treatment affects both the information set and the preferences ($\beta_2^S > 0$ and $\beta_1^S > 0$), it is not possible to distinguish which part of the effect on preferences, if any, was due to direct influence on preferences, and which resulted from the change in the information set.

A similar situation occurs in the interpretation of a significant effect on behavior outcomes. If an intervention has a significant impact on behavior but not on preferences ($\beta_3 > 0$ and $\beta_2 = 0$) I will conclude that the spillover effect is the result of social control exerted by the treated individual on other household members, leading the last ones to change their conduct, even if their preferences remain unchanged. On the other hand, if an intervention affects both, preferences and behavior ($\beta_3^S > 0$ and $\beta_2^S > 0$), I cannot infer what part of the change in behavior resulted from the change in preferences and what was due to social control.

Although, under the conceptual framework described above, influence may be directly exerted at more than one level, this specification only allows the identification of the first level at which influence is exerted. For the subsequent levels one cannot distinguish what is the result from direct influence from what is triggered by changes in the previous levels.

I will test the indirect effect of the three types of treatment separately because I believe that different treatments have different effects on treated individuals and, therefore, may trigger different types of influence.

Citizen's participation in elections involves two distinct decisions: first, whether to vote and, second, how to vote. According to the literature, the decision to vote depends less on the expected gain of one's favorite candidate winning and more on a subjective benefit that individuals draw from the act of voting (*Fiorina, 1976; Aldrich, 1993; Engelen, 2006*), and

social pressure (*Knack, 1992; Nickerson, 2008* and *Gerber, Green and Larimer, 2008*). On the other hand, the voting choice depends on policy preferences, partisanship or party identification (*Stokes, Campbel and Miller, 1958*; *Miller, 1991* and *Bartels, 2000*) and social influence (*Beck, Dalton, Green and Huckfeldt, 2002*). Therefore, the set of information and preferences that determine those behaviors are not identical. For instance, information regarding the elections and the voting procedure are more relevant to the decision of whether to vote; while information regarding the campaign and policy proposals of each candidate or party is more likely to affect the decision of how to vote. Similarly, a change in the preference/interest on elections is more likely to affect the decision to vote, while a change in the preferences towards the different parties is more relevant to the voting choice.

Here, the analysis will be focused on the decision to vote. The main reason for this is that all three interventions were aimed at promoting political participation, but not at changing voting patterns. All the leaflets and civic education cell phone messages included only general information about the elections, without making any judgment about candidates or parties. The same can be said about the newspaper that through all electoral campaign maintained an independent position. The only possible exceptions are the messages about electoral problems disseminated to hotline treated subjects. These messages normally reported the problem and the party who caused it.

The main caveat of this analysis is that the presence of significant spillover effects of the interventions on non-treated people who cohabit with treated individuals is not unequivocal evidence of interpersonal influence. I cannot completely rule out the possibility that some other members of the household also received the treatment directly (read the newspaper or the messages on the cell phone). However, in some of these cases the access to the treatment - the newspaper for instance - may be conditioned by the treated subject, who by allowing other member to access to it is also exerting influence.

Structural form specification

Using the reduced form specification, I test whether the effects of the interventions spillover to other people in the household. However, I would like to argue that these effects are transferred by treated individuals to other household members. This argument implies two steps: first,

the interventions affect treated individuals; and, then, these treated individuals influence other people in the household. The analysis based on the reduced form specification may provide evidence that the effects of the interventions reached other people in the household, but does not necessarily show that treated individuals influenced these other people.

In order to directly test if treated individuals influence other members of the household, I regress second interviewee's outcomes on each outcome of the main respondent living in the same house. There are several reasons why the information, preferences and behavior of two individuals living in the same household should be correlated. As I am interested in isolating the influence triggered by the voter education campaign, I instrument main respondent's outcomes with the three interventions. The purposes of this exercise are: first, to identify what type of effects on treated individuals – either on information set, preferences, or behavior – lead to an effect on second interviewees; and, second, to analyze into what extent treated individuals 'simplify' the interventions' effects when they transmit it to other people in the household. For instance, the interventions may expand treated individuals' information set, and those individuals, based on that new information, may choose to influence the preferences or behavior of other people in the household, instead of sharing the information.

I estimate a two-step model for each combination of second interviewee's outcome and main respondent's outcome. This means that I regress each second interviewee's outcome on each predicted outcome of the main respondent.

An alternative to this estimation strategy would be the adoption of an instrumental variable approach. However, that approach would restrict the estimation sample to the 320 targeted households[12] for which I have information for both main and second interviewees, and would imply the use of the same covariates in the first and second steps. Nevertheless, I report the estimation results of the instrumental variable model as a robustness check.

From now on, let z_i^P and y_i^S be measures of electoral information, preferences, or behavior for the main respondent and the second interviewee, respectively. Consider three treatment dummy variables that are equal to 1 when the household i is located in a village that was selected for the corresponding treatment, newspaper (N_i), hotline (H_i) and civic education (C_i). In the first step, I regress main respondent's post-treatment outcome (z_i^P) on the interventions (N_i, H_i

[12] If I also consider the untargeted households, this sample would be of 365 households.

and C_i), controlling for main respondent's demographic characteristics (X_i^P) and geographical location (G_i):

$$z_i^P = \delta_1 + \beta_n N_i + \beta_h H_i + \beta_c C_i + X_i^{P\prime}\gamma_1 + G_i'\tau_1 + \epsilon_{i,l} \tag{2.4}$$

I estimate this first step equation using the panel observations (≈975).[13] Using the estimation results of this step, I predict the post-treatment outcomes of interest for all main respondents.[14]

As I intend to analyze the transmission of the interventions' effects, the second step regression is only estimated for the main respondent's outcomes that were affected by the interventions (for which at least one of the interventions' coefficient is statistically significant in the first step regression).

In the second step, I regress second interviewee's outcome (y_i^S) on the predicted outcome of the main respondent living in the same household ($\widehat{z_i^P}$). In this step I control for the set of variables that were included in the first step equation (X_i^P and G_i) and also for second interviewee's characteristics (X_i^S). The equation takes the following form:

$$y_i^S = \delta + \beta\widehat{z_i^P} + X_i^{P\prime}\gamma + G_i'\tau + X_i^{S\prime}\rho + \varepsilon_i \tag{2.5}$$

Both equations are estimated using OLS. I use Hardin's sandwich cluster-robust covariance estimator to obtain the covariance matrix on which I base my inference (*Hardin, 2002*).[15]

The coefficient β measures the effect of the main respondent's outcome z_i^P on the second interviewee's outcome y_i^S.

The specification above assumes that the second interviewee's outcome is not directly affected by the interventions. Otherwise, the coefficient β will be biased upwards, as it will also capture the direct effect of the intervention. With the purpose of overcoming this limitation, I try to estimate a new two-step model in which I control for the direct effect of the interventions on the second

[13] I excluded the households in treated areas that did not received the treatment.

[14] The set of covariates $\mathbf{X}_i^P$ only includes data collected during the pre-elections survey in order to enable the estimation of the predicted outcomes for all treated subjects living with second interviewees.

[15] Hardin's sandwich estimator of variance is very similar to the two-step Murphy-Topel estimator, but it allows to take into account the clustered nature of the data. The formula used here assumes that the two estimating equations are derived from models with valid log likelihoods and that it is possible to compute the second derivatives of the second model's log likelihood.

interviewee's outcome. In these circumstances, I instrument main respondents' outcomes with interactions between the interventions and characteristics that trigger heterogeneous responses to the treatments.[16] This identification strategy relies on two main assumptions. First, the interaction terms must be correlated with main respondents' outcomes, conditional on the other covariates (including the interventions). Second, the direct effect of the interventions on second interviewees must not depend on the characteristics of the main respondents. If the interaction terms are associated with heterogeneous effects of the interventions, the first assumption should be verified. However, as I have three different interventions and several outcome variables, for each intervention I need at least one interaction term that is significant in the regression of the main respondent's outcome (in which the intervention has a significant impact). As the selection into the treatment was random, the main threat to the second assumption is related with the selection of second interviewees. If the selection into the sample of second interviewees is not conditional on unbservables that affect the second interviewees response to the treatments and that are correlated with the characteristics of the main respondents, the assumption should be verified.

Aker, Collier and Vicente (2013) identified a set of characteristics that generate heterogeneous responses to the treatments. I used the interactions that they found to be significant as instruments to estimate the main respondent's outcomes.

In the first step of this new model I regress the post-treatment outcome of the main respondent (z_i^P): on the three interventions (N_i, H_i and C_i); on the set of characteristics that generate heterogeneous effects (Q_i^P); on interactions between the demeaned equivalent of those characteristics[17] and the corresponding intervention for which they generated heterogeneous effects; on main respondent's demographic characteristics (X_i^P); and on province dummies (G_i). Adding the interactions guarantees that the main respondent's predicted outcome varies with the treatment, but also with the individual's characteristics. This should allow me to distinguish, in the second step regression, the direct effect of the campaign (captured by the interventions) and the indirect effect that is diffused by the main respondent (captured by the main respondent's

[16] A similar approach is presented in *Card (1993)*. The use of interaction terms as instruments is also explored in *Miguel, Satyanath and Sergenti (2004)*.

[17] The demeaned equivalent of characteristic q_i^P is $\widetilde{q_i^P} = q_i^P - \frac{1}{n}\sum_{i=1}^{n} q_i^P$, where n is the total sample size.

predicted outcome).

The main respondents' characteristics identified by *Aker, Collier and Vicente (2013)* are age (age_i^P), whether the individual completed 12 years of education (edu_i^P), whether the individual is a farmer ($agric_i^P$), an artisan (art_i^P) or works at home (dom_i^P), and whether the household owns cattle (cat_i^P). Based on the conclusions of *Aker, Collier and Vicente (2013)*, I include interaction terms between the newspaper intervention and all characteristics with the exception of whether the individual is a farmer; I include an interaction between the hotline and variable whether the individual is a farmer; and I also include interaction between civic education and education, whether the individual is an artisan and whether the household owns cattle. The equation takes the following form:

$$\begin{aligned} z_i^P &= \delta_1 + \beta_{1n} N_i + \beta_{1h} H_i + \beta_{1c} C_i + X_i^{P\prime}\gamma_1 + Q_i^{P\prime}\boldsymbol{\theta}_1 + \\ &+\kappa_g \cdot N_i \cdot \widetilde{age_i^P} + \kappa_{edu} \cdot N_i \cdot \widetilde{edu_i^P} + \kappa_{art} \cdot N_i \cdot \widetilde{art_i^P} + \\ &+\kappa_{dom} \cdot N_i \cdot \widetilde{dom_i^P} + \kappa_{cat} \cdot N_i \cdot \widetilde{cat_i^P} + \lambda_{agr} \cdot Hi \cdot \widetilde{agric_i^P} + \\ &+\pi_{edu} \cdot C_i \cdot \widetilde{edu_i^P} + \pi_{art} \cdot C_i \cdot \widetilde{art_i^P} + \pi_{cat} \cdot C_i \cdot \widetilde{cat_i^P} + G_i'\tau_1 + \epsilon_i \end{aligned} \tag{2.6}$$

I estimate this regression using all main respondents for which I have panel data.

I only estimate the second step regression if in the first step the coefficients of the seven interaction terms are jointly significant (F-test) at a level of at least 10 percent.

In the second step, I regress the second interviewee's outcome (y_i^S) on the predicted outcome of the main respondent living in the same household ($\widehat{z_i^P}$). In this regression I control for second interviewee's demographics characteristics ($\mathbf{X}_i^S$), and all covariates included in the first step regression, with the exception of the interaction terms:

$$y_i^S = \delta + \mu \widehat{z_i^P} + \beta_n N_i + \beta_h H_i + \beta_c C_i + X_i^{P\prime}\gamma + Q_i^{P\prime}\theta + G_i'\tau + X_i^{S\prime}\rho + \varepsilon_i \tag{2.7}$$

I estimate both equations using OLS. As before, I use Hardin's sandwich cluster-robust covariance estimator to obtain the covariance matrix on which I base my inference.

The coefficient μ measures the effect of the main respondent's outcome z_i^P on the second

interviewee's outcome y_i^S after controlling for the direct effect of the interventions on y_i^S.

2.3.2 Heterogeneous effects

The setting under analysis provides an excellent opportunity to explore what are the diffusion dynamics within households. To explore the determinants of the diffusion within the household, I will test the heterogeneous indirect effects of the interventions on the second interviewees.

Reduced form specification

Let f_i represent a characteristic of household i (for instance, age of main respondent or gender of second interviewee); and $\widetilde{f_i}$ denote the demeaned equivalent of that characteristic ($\widetilde{f_i} = f_i - \frac{1}{n}\sum_{i=1}^{n} f_i$, where n is the total sample size). Again, for the sake of the presentation, I focus on a single treatment. Given the randomized nature of the data, the heterogeneous indirect effect associated with characteristic f is estimated using the following regression:

$$y_i = \alpha + \beta T_i + \gamma f_i + \delta \left(T_i \times \widetilde{f_i} \right) + \varepsilon_i \tag{2.8}$$

Coefficient δ is the heterogeneous indirect effect associated with characteristic f. A significant coefficient ($\delta \neq 0$) will be interpreted as evidence that characteristic f is a determinant of influence, i.e. that f enhances (or reduces if the coefficient is negative) main respondent's influence on second interviewee's outcome. If f is a characteristic of the second interviewee, I cannot distinguish whether individuals with such characteristic are more vulnerable to influence or whether they are more likely to be target of influence.

Equation (2.8) estimates the heterogeneous effect associated with one characteristic (e.g. age) of one of the interviewees in the household, *ceteris paribus*. It is possible, however, that the impacts of primary and second interviewee's characteristic (f_i^P and f_i^S, respectively) are correlated. To account for that possibility, I estimate the heterogeneous effects associated with main and second interviewees' characteristics jointly:

$$y_i = \alpha + \beta T_i + \gamma_1 f_i^P + \gamma_2 f_i^S + \delta_1 \left(T_i \times \widetilde{f_i^P} \right) + \delta_2 \left(T_i \times \widetilde{f_i^S} \right) + w_i \tag{2.9}$$

It is also possible that different characteristics (e.g., age and education) may be correlated. Unfortunately, my dataset is too small to robustly test the heterogeneous effects associated with all characteristics jointly.[18]

All these regressions are also estimated controlling for individual demographic characteristics and geographical location (province dummies). Finally, all regressions are estimated using ordinary least squares (OLS) and the standard errors reported are adjusted by clustering at local level (enumeration area).

Structural form specification

I also estimate a two-step model for each combination of second interviewee's outcome, main respondent's outcome and characteristic that I intend to test. Due to the size of the sample, I only estimated the model that uses the interventions as instruments.

The first step of the model was estimated as presented above. In the second step, I regress the outcome of the second interviewee (y_i^S), on the predicted outcome of the main respondent living in the same household ($\widehat{z_i^P}$), on the characteristic I intend to test (f_i), and on an interaction term between the predicted outcome and the demeaned equivalent of the characteristic ($\widehat{z_i^P} \times \widetilde{f}_i$).[19] In this regression I control for the same set of covariates that were included in the first step ($\mathbf{X}_i^P$ and $\mathbf{G}_i$), plus the second interviewee's characteristics ($\mathbf{X}_i^S$). So, the second step regression takes the form:

$$y_i^S = \alpha + \beta\widehat{z_i^P} + \lambda f_i + \delta\left(\widehat{z_i^P} \times \widetilde{f}_i\right) + X_i^{P\prime}\gamma + X_i^{S\prime}\rho + G_i'\tau + \varepsilon_i \tag{2.10}$$

The coefficient β measures the average effect of main respondent's outcome z_i^P on second interviewee's outcome y_i^S. The coefficient δ measures the heterogeneous effect, i.e., it measures into what extent the effect of the main respondent's outcome on the second interviewee's outcome varies with characteristic f. A significant heterogeneous effect ($\delta \neq 0$) will be interpreted as

[18] Initially, I thought of testing for heterogeneous effects associated with the differences between the first and second respondents' characteristics. I abandoned that idea because the functional form including those variables is less meaningful for individuals in control areas, than the model including the characteristics directly. The interpretation of the coefficients of variables like 'same gender' and 'age difference' is problematic, as they may capture correlations between second interviewee's characteristic and the outcome of interest.

[19] For each second interviewee's outcome, the number of second step regressions that I will estimate will be the number of main respondent's outcomes that pass the significance test in the first step, multiplied by the number of characteristics that I want to test.

evidence that factor f impacts on main respondent's influence over second interviewee.

An alternative to the two-step model is an instrumental variable model (IV). So, I use the instrumental variable methodology proposed by *Wooldridge (2003)* to estimate the following equation:

$$y_i^S = \alpha + \beta z_i^P + \lambda f_i + \delta \left(z_i^P \times \widetilde{f}_i \right) + X_i^{P\prime}\gamma + X_i^{S\prime}\rho + G_i'\tau + \varepsilon_i \tag{2.11}$$

The two endogenous variables, the main respondent's outcome (z_i^P) and the interaction between that outcome and the demeaned equivalent of the characteristic under analysis ($z_i^P \times \widetilde{f}_i$), are instrumented with main respondent's predicted outcome ($\widehat{z_i^P}$) and the interaction between that prediction and the demeaned equivalent of the characteristic ($\widehat{z_i^P} \times \widetilde{f}_i$). The predicted outcome is computed based on the estimates of the following regression:

$$z_i^P = \alpha_1 + \beta_1 T_i + X_i^{P\prime}\gamma_1 + X_i^{S\prime}\rho_1 + G_i'\tau_1 + \epsilon_i \tag{2.12}$$

I estimate this equation using OLS and based on the IV estimation sample, i.e., on the sample of households for which I have post-treatment information for main and second interviewees.

Coefficient β in equation (2.11) measures the average impact of main respondent's outcome on second interviewee's outcome that is triggered by the voter education campaign. Coefficient δ measures how much that impact varies with characteristic f.

2.4 Data and descriptive statistics

The empirical analysis is based on experimental data collected around the October 2009 general elections in Mozambique.[20] During the experiment, three different forms of voter education interventions were implemented: the distribution of a newspaper, the creation of a text messaging hotline to report electoral problems and an electoral education campaign based on a leaflet and text messages.

In this section, firstly, I present a brief description of the interventions; secondly, I give a concise account of the experiment design; thirdly, and finally, I present the descriptive statistics

[20] This project was coordinated by Pedro Vicente, Jenny Aker and Paul Collier.

of the sample under analysis.[21]

2.4.1 Treatments

The newspaper intervention consisted in the distribution of the free newspaper @Verdade, once a week, since the baseline survey until the week after the elections.[22] It was the first time this newspaper was distributed in the treated locations (until then it was only distributed in the city of Maputo). The newspaper maintained a neutral position during the electoral period and focused on providing electoral information. During the electoral period all issues of the newspaper included a version of the civic education leaflet and the newspaper also advertised a national hotline to report electoral problems (see *Figure 2.1*), very similar to the one that the hotline intervention used.

The hotline treatment consisted in the distribution of a leaflet (see *Figure 2.2*) inviting people to send messages reporting electoral problems that occurred in their locations during the electoral campaign and the elections-day. During the elections-day and the 14 days before, every message received was confirmed with a correspondent in the respective location, and the information was disseminated to the cell phones of the experimental subjects. During that period the treated individuals were also sent cell phone text messages reminding them to use the hotline to report problems. The civic education treatment consisted in providing the individuals with basic information regarding the general elections. It began with the distribution of a leaflet (see *Figure 2.3*) during the baseline survey, which explained thoroughly each step of the voting procedure.[23] Then, experimental subjects were sent cell phone text messages (5 a day) during the 14 days before the elections and on the elections-day. These messages emphasized the importance of voting (get out the vote) and provided information about: the date of the elections, the different elections taking place (for president, parliament, and provincial assemblies), the presidential candidates and the parties running for parliament, the absolute secrecy of the vote and the need to mark only one X on each ballot paper.

[21] More details on the treatments and experimental design can be found in *Aker, Collier and Vicente (2013)*.
[22] All treated individuals received at least one issue, but it was not possible to ensure that all individuals received all issues.
[23] This leaflet was made available by the National Electoral Commission.

Figure 2.1: Newspaper @Verdade (front page – edition before election; civic education page; hotline page)

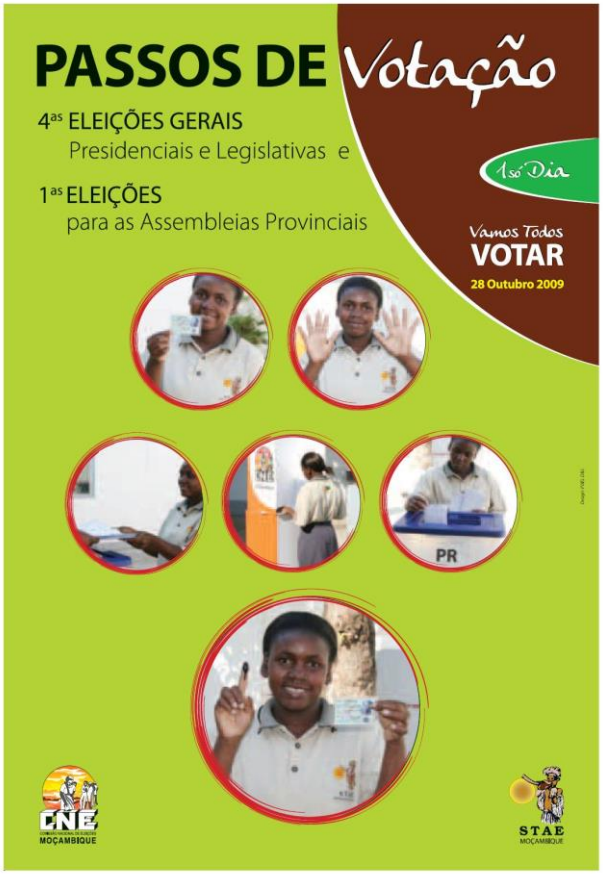


Figure 2.2: Hotline leaflet

Figure 2.3: Civic education leaflet by CNE/STAE

2.4.2 Measurement

The experiment covered 161 polling locations distributed in four provinces, Cabo Delgado, Zambezia, Gaza, and Maputo-Province. These locations are nationally representative of the population of Mozambique that has cell phone coverage. These locations resulted from a two-stage clustered representative sampling, first on provinces, then on enumeration areas; using as weights the number of registered voters per pooling location.[24] Given the importance of the access to cell phones in the implementation of the interventions, the pooling locations not covered by a cell phone network were removed from the sampling framework. In practice, this consisted in disregarding all the pooling locations that were more than 5 km away from a cell phone antenna.[25] Despite this, during the baseline survey seven locations had to be replaced by the closest pooling station, because when the team of interviewers arrived there they found that there was no cell phone coverage.

Of the 161 enumeration areas selected for the study, 40 were randomly assigned to each of the three treatments[26], and 41 locations serve as the control group, with no treatment administered.[27] The treatments were allocated following a stratified randomized procedure (*Bruhn and McKenzie, 2009*): first, in each province were formed clusters of four closest enumeration areas; then, each treatment was randomly allocated to one location in each cluster (*Aker, Collier and Vicente, 2013*, p. 13).

Sampling the households in each location followed standard household representativeness, n^{th} calls (*Aker, Collier and Vicente, 2013*, p. 14). In each selected household, the primary interviewee was the household head or his/her partner. Once selected the potential interviewee, the interview, and hence the delivery of the treatment, was only conducted if the respondent had

[24] The sample was defined based on the list of pooling stations, as well as the number of registered voters, provided by the CNE/STAE in their 2004 elections (disaggregated) electoral data electronic publication. The information relative to 2009 elections was only available one month before the election-day, when the pre-elections survey was already being conducted (*Aker, Collier and Vicente, 2013*).

[25] In order to identify all the pooling stations that were more than 5 km away from a cell phone antenna, *Aker, Collier and Vicente (2013)* used data made available by the two cell phone operators (MCell and Vodacom) on the precise geographical location of each of their antennae.

[26] During the baseline survey one of the interviewers team mistakenly delivered the civic education treatment in an enumeration area that was supposed to receive the newspaper treatment. Therefore, our sample of treated locations includes 39 newspaper areas, 40 hotline areas and 41 civic education areas.

[27] It was supposed to be only 40 control areas instead of 41. While in the field, during the pre-elections survey, one of the teams thinking that would not be able to conduct the survey in one given location replaced it by the closest one; however, later the team was also able to conduct the survey in the original location. Therefore there is an extra control location.

access to a cell phone.[28]

For measuring purposes, pre- and post-elections surveys were run in the treated and control locations. The baseline survey took place between September 17 and October 14, and included 1766 households/primary respondents, approximately 11 per experimental location. Note that in each treatment location, on average 2 of the 11 respondents did not receive the treatment. I will refer to these respondents and their households as untargeted. The post-elections survey started after the election results were made public, and lasted from November 16 until December 9.

During the post-elections survey, the enumerators travelled in pairs; in each household one would interview the experimental subject and the other would interview a second adult (who did not receive the treatment). It was previously defined a list of types of household members: spouse of main respondent, son, daughter, other male member of the household and other female member of the household. The selection of the second interviewee alternated between these types of household members. In each household the interviewer requested to speak with an individual of a certain type, following this list and depending on the type that was requested in the previous household visited. Within the categories son and daughter, the interviewer would start by trying to interview the eldest, and if he or she were not available he would ask to talk with the second eldest, and so forth. Within the types other household members, the interviewer would begin by trying to interview the parent, if he or she was not available or did not exist, the interviewer would move on to siblings, nephews/nieces and grandchildren. Within each of these categories, the enumerator would start by asking for the eldest to the youngest. If the selected type of household member did not exist, or was not available to be interviewed before the team of interviewers left the enumeration area, the interviewer would select a second interviewee of the following type of household member.[29]

The post-elections survey included 1154 main interviewees[30] and 518 second interviewees, but

[28] Note that access to cell phone was not understood as ownership of a cell phone. So, people who did not own a cell phone but had an easy access to a neighbor or family member owning a cell phone, were also included in the sample.

[29] For instance, if in the last household visited the interviewer asked to talk with the spouse, in this new household he would ask to interview the oldest son. If the main respondent in this household did not have sons, the interviewer would ask to talk with the eldest daughter that lives in the household. If none of the daughters was available to be interviewed, the interviewer would ask to talk with the father of the main respondent. And so forth until the enumerator identified a second adult to be interviewed.

[30] There is no evidence that sample attrition is correlated with geographical location, gender, age and occupation.

only in 365 households it was possible to interview both, the main and a second, interviewees. This sample included 175 main interviewees that lived in treated areas but did not receive the treatment and 59 second interviewees that lived in these untargeted households.

The main respondents' questionnaires (in both pre- and post-elections surveys) were designed with the purpose of collecting data to measure electoral behavior, perceptions about electoral problems and authority, interest and information about politics (*Aker, Collier and Vicente, 2013*). For logistic reasons, the second interviewees' questionnaire (only conducted in the post-elections survey) was a shorter version of the main respondents' questionnaire.

Based on the conceptual framework, I consider three types of outcomes: basic electoral information, interest in elections and turnout. I constructed two proxies of the information set relevant for the individual's decision to vote or not. There is, however, the risk that these indicators do not capture the relevant information and, hence, I may be unable to identify influence on relevant information. In terms of preferences, although the dataset does not include a specific measure of the respondents' preferences towards voting, it has measures of their interest in the elections and public matters, which may be a reasonable approximation of their interest in voting. A summary of the outcome variables is presented in *Table 2.1.*

My first proxy of information, information about elections, comes from questions asking respondents to list the names of the presidential candidates and the main parties running for the parliamentary election. The second proxy of information, information about voting, is based on questions asking the respondents to explain the concepts of electoral abstention, null vote and blank vote. The proxy of interest is based on respondent's self-reported interest in the presidential election, parliamentary election, provincial assemblies' election, and generally in public matters. These questions are combined into indices. I construct these indices following the approach of *Kling, Liebman and Katz (2007).* The indices correspond to the mean of the z-scores of the survey-indicators. The z-scores are calculated taking as reference the control group mean and standard deviation.[31]

[31] Following *Kling, Liebman, and Katz (2007)*, when an individual has a valid response to at least one component measure of the index, I impute any missing values for other component measures at the individual's group (control, newspaper, hotline or civic education) mean.

Table 2.1: Dependent variables

Measure	Indicators	Description / Phrasing of the question	Range / Scale	
Index of information about elections	Knows names of presidential candidates	Do you know who were the candidates to President in the October 2009 elections?	0 - 3	
	Knows names of parties	Are you able to list five parties that ran for the Parliament in the October 2009 elections?	0 - 3	
Index of information about voting	Knows concept of electoral abstention	Do you know what "electoral abstention" means?	1 - 3	No - Yes
	Knows concept of null vote	Do you know what "null vote" means?	1 - 3	No-Yes
	Knows concept of blank vote	Do you know what "blank vote" means?	1 - 3	No-Yes
Index of interest in elections	Interest in presidential election	With how much interest did you follow the 2009 presidential election?	1 - 4	None-Very much
	Interest in parliamentary election	With how much interest did you follow the 2009 parliamentary election?	1 - 4	None-Very much
	Interest in provincial assemblies' election	With how much interest did you follow the 2009 provincial assemblies' election?	1 - 4	None-Very much
	Interest in public matters	With how much interest do you follow public matters?	1 - 4	None-Very much
Turnout	Self-reported	Which of the following sentences best describes your situation during the 2009 Elections: (1) I was not a registered voter and I was not interested in voting; (2) I was not a registered voter but I would have liked to have voted; (3) I was a registered voter but I chose not to vote; (4) I was a registered voter but I was not able to vote; and (5) I was a registered voter and I voted.	0 - 1	Missing if (1) or (2); 0 if (3) or (4); 1 if (5)
	Adjusted	Self-reported turnout adjusted by considering as non-voters those who did not answer correctly the basic questions about the elections day, namely about the number of ballot papers and ballot boxes.	0 - 1	
	Finger measure	What finger was inked after voting?	0 - 1	1 if showed inked finger without hesitation
	Index 1	Composite index measuring how well the respondent answered questions on circumstances and events during the election day	0 - 7	
	Index 2	Composite index based on the sub-group of these questions that focus on what occurred at the ballot station	0 - 7	
	Interviewer's assessment	Enumerator assessment of likelihood that respondent voted	0 - 7	

As all the interventions aimed at increasing turnout, one of the main concerns of the survey instruments was to provide a measure of individual turnout. I consider six measures of turnout. The first measure is self-reported turnout.[32] The second one is the self-reported voting behavior adjusted by considering as non-voters those who answered wrongly basic questions regarding the voting procedure.[33] The third one is an indicator of whether the respondent showed without hesitation the finger that had been dipped in ink during the voting procedure. The fourth and fifth measures are indices constructed by *Aker, Collier and Vicente (2013)*. Index 1 is a composite index that measures how well the respondent answered questions regarding what occurred during the election-day. Index 2 focuses on how well the respondent answered questions about what occurred at the ballot station. Finally, the sixth measure is the enumerator's assessment of the likelihood that the respondent actually voted. The three last measures take values between 0 and 7. To facilitate results' interpretation and comparison, I normalized them, dividing by 7.

2.4.3 Descriptive statistics

The descriptive statistics of the demographic characteristics of the samples of primary and second interviewees are displayed in *Table 2.2*.

Looking at the sample of targeted primary respondents, I find that most interviewees are women; and that they are, on average, between 36 and 39 years old. In all groups most of the interviewees are either married or in a union (from 69.1 percent in control areas to 76.7 percent in newspaper areas). The percentage of respondents that never attended school varies between 15.4 in newspaper areas and 22.1 in civic education areas. Approximately a third of the main respondents work in agriculture (the percentage varies between 31.3 in newspaper areas and

[32] The information collected in the baseline survey suggests that self-reported voting behavior is a biased measure. According to that measure 95 percent of the registered respondents voted in the general elections of 2004, while according to the STAE statistics the national turnout was around 36 percent. There may be some explanations why the turnout among household heads and their partners is above the national level; however a difference of 60 percentage points seems to hint that people tend to misreport their voting behavior. Therefore, the post-elections questionnaire was carefully designed in order to provide reliable measures of individual turnout. In addition to directly ask about the individual voting behavior, an entire module of the questionnaire was committed to obtain a detailed account of the respondent's elections-day and, specifically, about the actual voting procedure (e.g., number of ballots, number of ballot boxes, whether the ballots had pictures, etc.). Finally, in the end of this module, the enumerator was asked to assess the likelihood that the interviewee had voted based on her reactions and behavior.

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

36.3 in civic education areas). The average household size is around 6. In terms of balance tests, at standard significance levels, there are very few differences between treated and control groups. First, there is a larger percentage of respondents who are either married or in a union in newspaper areas. Second, the household size is slightly larger in hotline areas. And in civic education areas there is a lower percentage of respondents whose main occupation is artisan and a larger percentage of respondents who are public servants.[34]

Looking at the sample of second interviewees, I find that in all four groups most of the second interviewees are also women; and that these interviewees are, on average, between 30 and 32 years old. The percentage of interviewees who are either married or in a union varies between 42.2 in civic education areas and 62.8 in hotline areas. The percentage of respondents with no schooling varies between 16.8 percent in the hotline group and 24.4 in the control group. The most common occupation among second interviewees is also agriculture. In the four groups the average household size is between 6 and 7 people. The majority of second interviewees are either the spouse or child of the main interviewee. In terms of balance tests, at standard significance levels, there are few differences between treated and control groups. First, the hotline group has a larger percentage of respondents that are either married or in a union and, hence, a lower percentage of individuals who are single, than the control group. Second, the household monthly expenditure is larger in hotline areas. Third, in the newspaper group there are less second interviewees who are the spouse of the main respondent. And, fourth, in the civic education group there is a lower number of second interviewees who are a daughter of the main respondent.[35]

Overall, it seems that the observables are balanced across treatment and control groups, which leads us to believe that the unobservable characteristics should also be.

[34] For the sample of main respondents, I computed 69 differences across comparison groups and found 2 statistically significant differences at the 10 percent level and one statistically significant difference at the five percent level. This pattern is consistent with a randomized assignment of the treatment.

[35] For the sample of second interviewees, I computed 75 differences across comparison groups and found one statistically significant difference at the 10 percent level, and four statistically significant differences at the five percent level.

Table 2.2: Differences across treatment and control areas

		Sample of targeted main respondents (post-elections)				Sample of second respondents			
		Control	**Newspaper**	**Hotline**	**Civic Education**	**Control**	**Newspaper**	**Hotline**	**Civic Education**
Basic Demographics	**female**	0.563	-0.031	-0.024	0.060	0.623	-0.053	-0.035	-0.068
			(0.044)	(0.044)	(0.041)		(0.062)	(0.060)	(0.061)
	age	36.957	-0.811	1.713	0.280	31.520	-0.981	0.756	-0.004
			(1.394)	(1.461)	(1.329)		(1.756)	(1.712)	(1.732)
	single	0.186	-0.042	0.006	-0.039	0.443	0.020	-0.122**	0.089
			(0.037)	(0.041)	(0.036)		(0.063)	(0.059)	(0.062)
	married or in a union	0.692	0.092**	0.039	0.050	0.504	-0.033	0.124**	-0.082
			(0.040)	(0.047)	(0.044)		(0.063)	(0.060)	(0.062)
	no school	0.179	-0.025	0.006	0.042	0.244	-0.023	-0.076	-0.049
			(0.033)	(0.040)	(0.041)		(0.053)	(0.049)	(0.052)
	incomplete primary school	0.276	0.063	-0.053	0.031	0.221	0.033	0.049	0.021
			(0.047)	(0.041)	(0.043)		(0.054)	(0.053)	(0.053)
	incomplete secondary school	0.168	-0.019	0.003	-0.013	0.244	0.010	-0.033	-0.033
			(0.038)	(0.037)	(0.037)		(0.055)	(0.051)	(0.052)
	completed at least primary school	0.449	0.016	0.068	-0.042	0.496	-0.013	0.022	0.043
			(0.057)	(0.058)	(0.057)		(0.084)	(0.083)	(0.089)
	household size	5.789	0.227	0.483*	0.068	6.271	0.360	0.298	0.389
			(0.282)	(0.261)	(0.286)		(0.340)	(0.312)	(0.338)
	household head	0.742	0.037	-0.024	-0.017	-	-	-	-
			(0.049)	(0.044)	(0.044)				
Assets and expenditure	**house**	0.853	-0.018	0.016	0.013	0.893	-0.065	0.005	-0.034
			(0.032)	(0.030)	(0.030)		(0.043)	(0.038)	(0.041)
	land	0.652	-0.047	-0.045	-0.055	0.656	-0.058	0.000	-0.063
			(0.059)	(0.057)	(0.056)		(0.061)	(0.058)	(0.060)
	cattle	0.254	0.004	0.007	-0.006	0.260	0.003	0.040	-0.025
			(0.049)	(0.052)	(0.045)		(0.055)	(0.055)	(0.054)
	cell phone	0.709	0.022	0.056	0.022	0.679	0.042	0.072	0.024
			(0.064)	(0.065)	(0.065)		(0.058)	(0.055)	(0.058)

Note: Standard errors reported; these are corrected by clustering at the location (enumeration area) level.

* significant at 10%; ** significant at 5%; *** significant at 1%. These results come from OLS regressions.

Table 2.2: Differences across treatment and control areas (cont.)

		Sample of targeted main respondents (post-elections)				Sample of second respondents			
		Control	Newspaper	Hotline	Civic Education	Control	Newspaper	Hotline	Civic Education
Occupation	has a job	0.240	-0.033	0.029	-0.052	0.183	-0.011	0.052	-0.048
			(0.041)	(0.049)	(0.042)		(0.048)	(0.050)	(0.046)
	agriculture	0.353	-0.064	-0.028	-0.000	0.323	-0.009	-0.027	-0.017
			(0.066)	(0.070)	(0.065)		(0.059)	(0.057)	(0.059)
	retail informal sector	0.043	0.002	0.008	0.006	0.046	-0.005	0.006	-0.022
			(0.022)	(0.019)	(0.019)		(0.026)	(0.027)	(0.023)
	artisan	0.050	0.017	0.001	-0.030*	0.023	-0.015	-0.008	-0.015
			(0.022)	(0.019)	(0.016)		(0.016)	(0.017)	(0.016)
	unskilled worker	0.054	0.018	0.006	-0.013	0.015	0.018	0.007	-0.007
			(0.023)	(0.021)	(0.020)		(0.019)	(0.017)	(0.014)
	teacher	0.047	-0.002	0.013	-0.022	0.023	0.010	0.021	0.001
			(0.019)	(0.027)	(0.015)		(0.021)	(0.022)	(0.019)
	student	0.040	0.001	-0.014	0.018	0.131	0.051	0.010	-0.018
			(0.018)	(0.019)	(0.022)		(0.046)	(0.042)	(0.041)
	domestic	0.147	-0.017	-0.032	-0.008	0.123	0.001	-0.056	-0.002
			(0.032)	(0.032)	(0.031)		(0.042)	(0.036)	(0.041)
Relationship to primary interviewee	spouse	-	-	-	-	0.473	-0.131**	0.027	-0.060
							(0.063)	(0.062)	(0.062)
	children	-	-	-	-	0.240	-0.010	-0.005	-0.034
							(0.054)	(0.053)	(0.052)
	daughter	-	-	-	-	0.156	-0.037	-0.053	-0.093**
							(0.044)	(0.041)	(0.039)
	son	-	-	-	-	0.086	0.025	0.047	0.057
							(0.038)	(0.039)	(0.040)

Note: Standard errors reported; these are corrected by clustering at the location (enumeration area) level.

* significant at 10%; ** significant at 5%; *** significant at 1%. These results come from OLS regressions.

2.5 Results: Average effects

2.5.1 Reduced form specification

Results

I start by presenting the estimates of the average effects of the different intervention on the main respondents. These effects are analyzed in the next Chapter and are explored in full detail in *Aker, Collier and Vicente (2013)*. The purpose of including these results here is to identify what type of influence we would expect to occur in households in different treatment areas. The regression results are presented in *Table 2.3*. All the regressions are estimated using OLS and considering only post-election data. For each outcome of interest I present two sets of estimates: the first with no controls and the second including individual characteristics and provincial dummies and EA controls.[36]

The newspaper has a significant positive effect on electoral information and on the measure of political participation open letter. The coefficient of this intervention on the regression of interest and all turnout proxies is relatively close to zero, which leads me to believe that the newspaper has no effect on those outcomes. The hotline has a positive effect on electoral information, and a strong and consistent positive effect on turnout. Unlike the newspaper, the coefficient of the hotline in the regression of interest is relatively large, despite statistically insignificant. The magnitude of this coefficient suggests that the hotline might have a positive impact on interest, which I cannot precisely estimate. Civic education only has a significant effect on turnout, and only when measured by the proxies adjusted turnout, measure based on the finger, index 2 and the interviewer's assessment.[37] Most of these effects are in line with what we would expect based on the model described above and given the general nature of each intervention.

[36] By construction the average indices of information and interest living in control areas are zero (intercept term of the regressions with no controls). The coefficients in the regressions should be read in terms of deviation units in the regressions of indices, and in terms of percentage points (divided by 100) in the regressions of turnout measures and 'open letter'.

[37] I also computed the adjusted p-values using the Holm-Bonferroni method to take into account the family of hypothesis associated to each treatment (2 specifications x 11 outcomes). Based these p-values only some effects of the hotline would be significant at a level of at least 10%, namely the effects on the turnout proxies interviewer assessment (with and without controls), self-reported turnout, index 1 and index 2 (without controls); and the effect on information about the elections (with controls). If I adjust the p-values taking into account all the hypothesis tested in the table (3 treatments x 2 specifications x 11 outcomes), only the effect of the hotline on the proxy interviewer assessment (without controls) would be significant.

Table 2.3: Average effect of the three treatments on targeted respondents information, interest and political participation

	Information				Interest in elections		Turnout												Open letter	
	About elections		About voting				Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)		(15)	(16)	(17)	(18)	(19)
Newspaper	0.122	0.192**	0.082	0.128	0.008	-0.059	0.024	0.012	0.051	0.015	0.011	0.005	0.028	0.013	0.034	0.016	0.035	0.024	0.087*	0.109**
	(0.095)	(0.085)	(0.103)	(0.093)	(0.104)	(0.084)	(0.037)	(0.038)	(0.048)	(0.042)	(0.042)	(0.042)	(0.035)	(0.035)	(0.035)	(0.034)	(0.039)	(0.038)	(0.050)	(0.053)
Hotline	0.165*	0.233***	0.074	0.125	0.103	0.090	0.080***	0.070**	0.120***	0.094**	0.063**	0.053*	0.075***	0.063**	0.080***	0.065**	0.094***	0.085***	-0.037	-0.002
	(0.089)	(0.083)	(0.091)	(0.080)	(0.106)	(0.078)	(0.026)	(0.029)	(0.044)	(0.041)	(0.028)	(0.030)	(0.025)	(0.027)	(0.027)	(0.027)	(0.026)	(0.028)	(0.035)	(0.038)
Civic education	0.037	0.046	-0.047	0.019	0.078	-0.001	0.045	0.032	0.085*	0.073*	0.056*	0.043	0.043	0.031	0.052*	0.039	0.073***	0.069**	0.044	0.061
	(0.097)	(0.101)	(0.094)	(0.093)	(0.094)	(0.063)	(0.028)	(0.030)	(0.044)	(0.042)	(0.029)	(0.033)	(0.026)	(0.027)	(0.028)	(0.028)	(0.027)	(0.029)	(0.048)	(0.055)
constant	0.000	-0.226	-0.000	-0.141	0.000	0.026	0.859***	0.887***	0.706***	0.824***	0.807***	0.783***	0.788***	0.841***	0.757***	0.835***	0.753***	0.728***	0.153***	0.140*
	(0.066)	(0.145)	(0.066)	(0.141)	(0.075)	(0.118)	(0.021)	(0.060)	(0.031)	(0.074)	(0.020)	(0.070)	(0.019)	(0.056)	(0.020)	(0.056)	(0.020)	(0.063)	(0.026)	(0.083)
Adjusted R-squared	0.003	0.062	0.001	0.067	-0.000	0.110	0.006	0.023	0.008	0.061	0.002	0.016	0.006	0.026	0.007	0.041	0.010	0.024	0.011	0.013
No. of observations	976	865	979	868	976	865	953	845	953	845	953	845	953	845	953	845	953	845	973	817
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 2.4: Average indirect effect of the three treatments on second interviewees' information, interest and political participation

	Information				Interest in elections		Turnout												Open letter	
	About elections		About voting				Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Newspaper	-0.072	-0.114	0.037	0.031	-0.036	0.061	0.018	0.007	0.059	0.056	0.035	0.017	0.026	0.016	0.042	0.030	0.052	0.057	-0.008	-0.007
	(0.140)	(0.118)	(0.125)	(0.110)	(0.131)	(0.105)	(0.052)	(0.054)	(0.062)	(0.060)	(0.055)	(0.058)	(0.049)	(0.050)	(0.049)	(0.049)	(0.047)	(0.051)	(0.029)	(0.032)
Hotline	0.060	0.005	0.089	0.132	0.175	0.211**	0.061	0.051	0.043	0.032	0.100**	0.072	0.056	0.044	0.063	0.051	0.099**	0.088*	0.070*	0.057
	(0.134)	(0.112)	(0.117)	(0.102)	(0.127)	(0.104)	(0.044)	(0.044)	(0.054)	(0.048)	(0.049)	(0.050)	(0.041)	(0.041)	(0.042)	(0.041)	(0.048)	(0.047)	(0.036)	(0.037)
Civic education	-0.042	-0.037	0.028	0.061	0.099	0.170	0.071	0.081*	0.075	0.093*	0.093*	0.082	0.065	0.076*	0.074*	0.088**	0.110**	0.110***	0.030	0.022
	(0.129)	(0.117)	(0.106)	(0.099)	(0.141)	(0.108)	(0.044)	(0.044)	(0.064)	(0.055)	(0.054)	(0.055)	(0.042)	(0.041)	(0.044)	(0.042)	(0.045)	(0.042)	(0.031)	(0.032)
constant	-0.000	0.380***	0.000	-0.081	-0.000	-0.310**	0.856***	0.761***	0.737***	0.653***	0.771***	0.735***	0.797***	0.755***	0.764***	0.729***	0.722***	0.561***	0.054**	-0.014
	(0.086)	(0.134)	(0.073)	(0.147)	(0.098)	(0.153)	(0.037)	(0.073)	(0.040)	(0.085)	(0.038)	(0.082)	(0.034)	(0.067)	(0.034)	(0.068)	(0.034)	(0.070)	(0.021)	(0.037)
Adjusted R-squared	-0.004	0.076	-0.005	0.123	0.003	0.221	0.002	0.043	-0.002	0.075	0.006	0.032	0.001	0.047	0.003	0.056	0.011	0.062	0.007	0.021
No. of observations	459	427	459	427	459	427	426	397	426	397	426	397	426	397	426	397	426	397	459	427
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Now, I turn to the average indirect effects of the interventions on the second interviewees. The regression results are presented in *Table 2.4*. As before, I present two sets of estimates for each dependent variable, one without controls and another controlling for individual characteristics and geographic location.[38]

The coefficient of the newspaper intervention is not significant in any of the regressions of second interviewees' outcomes. This means that I cannot find evidence of spillovers in the households targeted by this intervention. Although the delivery of the newspaper increased the information set of people who received it, there is no evidence that this additional information reached the other people in the household.

In households that received the hotline intervention there is evidence of spillovers on interest in elections, and on political participation. As in newspaper areas, although the hotline has a positive impact on the level of electoral information of the targeted respondents, this effect does not seem to be transmitted to the other household members. The evidence suggests that influence was directly at the level of preferences (measured by the proxy interest). The indirect effect on turnout is significant when using the proxies based on the inked finger (10 percentage points, when using no controls) and on the interviewer's assessment (8.8 and 9.9 percentage points, with and without controls). In this scenario we cannot conclude that there was direct influence on behavior, because the increase in individual turnout may be a result, at least partially, of the significant increase in the interest in elections.

In civic education areas I find a positive indirect effect on second interviewees' turnout. Looking at the significant coefficients, this effect varies between 7.4 percentage points on index 2 (without controls) and 11 percentage points on the interviewer's assessment. I interpret this as evidence of influence directly at the level of behavior. This suggests that in these areas second interviewees vote more because they are persuaded to do it and not because they are more interested or more informed. The effects of civic education on both targeted and second interviewees suggest that this intervention works mainly through persuasion, it pressures targeted individuals to vote, who in their turn exert pressure (intentionally or not) on the people around them to also vote. This interpretation is based on the assumption that the index interest is a

[38] Unfortunately, none of the coefficients presented in the Table are significant when we adjust the p-values using the Holm-Bonferroni method.

good proxy for preferences and, as a results, it captures other positive motivations to vote like civic-mindedness or enthusiasm about the elections.

Contrary to what standard economic models would assume, the transmission of the interventions' effects does not seem to be driven by information sharing. Although two of the interventions (newspaper and hotline) have a positive impact on information of the targeted, I find no evidence of information sharing within the household. The transmission of the effects seems to occur through sharing of preferences and pressure. Therefore, some campaigns that are aimed at raising citizens' awareness and enable them to make a well-informed voting decision, like the civic education intervention, might be contributing to increase social pressure to vote.

In a broader context, the findings reported here constitute an important contribution to the literature on voter mobilization, mostly because they corroborate some conclusions that so far were based almost exclusively on US data.[39] Firstly, they show that, in Mozambique, the effects of campaigns aimed at increasing political participation, especially empowerment campaigns (like the hotline) and call out the vote campaigns (like the civic education intervention), are transmitted to people who live with targeted individuals. This corroborates the conclusion of *Nickerson (2008)*. This author, using data from a randomized experiment conducted in two North American cities, showed that the effect of call out the vote door-to-door campaigns is passed to other people in the household. Secondly, these results provide evidence of people voting due to social control in a developing country. This adds to the evidence provided by *Gerber, Green and Larimer (2008)*, *Davenport (2010)* and *Panagopoulos (2010)*, based on United States' data. *Gerber, Green and Larimer (2008)* demonstrated the "power of a single mailer disclosing the voting behavior of oneself and one's neighbors" (p. 42) in increasing individual turnout. *Davenport (2010)* showed that a similar strategy dramatically increased the turnout among people who rarely vote. *Panagopoulos (2010)*, provided evidence that appeals to shame (informing individuals that the names of all non-voters would be published in the newspaper) are more effective in increasing turnout than appeals to pride (informing individuals that the names of all voters would be published in the newspaper). In addition to all this, the evidence regarding the newspaper areas has shown the limited capacity of information campaigns based

[39] Another exception is the study of *Gine and Mansuri (2011)*, who provide evidence of a field experiment on female voter turnout in Pakistan.

on providing reading material in reaching other people besides the targeted individuals.

Robustness

One potential weakness of this econometric analysis is related with the attrition in the sample of main and second interviewees. This lack of information is largely due to the absence of the experimental subject or a second adult in the household at the moment of the enumerators visit, rather than the refusal to be interviewed. Most of the visits took place during the day. Due to logistic constraints and given the fact that the second interviewees were not the main targets of the experiment, it was given priority to locate and interview main respondents. As a result, even if the potential second interviewee was only momentarily absent, sometimes it was not feasible to hold the work of the enumerators team to wait for such individual. The most reported specific reason for sample attrition for both, main and second, respondents was agricultural season. The post-elections survey took place during the rainy season in Mozambique. Most of the agricultural work in the fields ("machambas") occurs during this period. As the agricultural plots tend to be located at a fair distance from home (*Sheldon, 1999*; *De Vletter, 2001*), often agricultural workers migrate during this season (*Aker, Collier and Vicente, 2013*). Probably because in some cases all the family migrated (for example, that may be the case of families where all adults are agricultural workers, and they take the young children with them), the team of enumerators found some of the households empty. In Maputo Province working and travelling were also often reported as reasons for attrition. These reasons are only relevant in Maputo Province because this is the province more urbanized – and thus with more non-farm employment opportunities – and more "connected" (*Cungara et al., 2011*).[40] If the selection into the sample of second interviewees is correlated with the outcomes of interest, my results may not be generalized to the rest of the population. For instance, if people who spend more time at home are more likely to be included in the sample and are also more likely to be influenced by other people in the household, my estimates of the indirect effects of the interventions on the other household members will be biased upwards.[41]

[40] Sample attrition at the level of main respondents is also discussed in the next Chapter and in *Aker, Collier and Vicente (2013)*.

[41] A similar argument can be done with respect to the attrition in the sample of main respondents. Next Chapter provides some evidence that the estimates presented in *Table 2.4* are robust to selection bias.

Table 2.5: Average indirect effect of the three treatments on second interviewees' outcomes, imputing missing information

	Information				Interest in elections		Turnout												Open letter	
	About elections		About voting				Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Newspaper	-0.063	-0.103	0.005	-0.010	0.035	0.055	0.025	0.028	0.087	0.088	0.054	0.051	0.037	0.039	0.056	0.056	0.063	0.070	0.007	0.000
	(0.130)	(0.117)	(0.108)	(0.096)	(0.132)	(0.106)	(0.048)	(0.045)	(0.069)	(0.063)	(0.058)	(0.054)	(0.044)	(0.040)	(0.045)	(0.041)	(0.049)	(0.045)	(0.040)	(0.040)
Hotline	0.066	0.035	0.141	0.130	0.194	0.184*	0.071	0.065	0.051	0.046	0.107*	0.104*	0.069	0.064	0.076*	0.070*	0.100**	0.101**	0.073*	0.068*
	(0.118)	(0.110)	(0.112)	(0.094)	(0.128)	(0.104)	(0.048)	(0.046)	(0.064)	(0.057)	(0.060)	(0.057)	(0.044)	(0.041)	(0.045)	(0.042)	(0.048)	(0.045)	(0.042)	(0.041)
Civic education	0.037	0.001	0.085	0.059	0.134	0.126	0.060	0.062	0.085	0.085	0.065	0.058	0.058	0.060	0.069	0.073*	0.078	0.072	0.032	0.028
	(0.124)	(0.117)	(0.114)	(0.102)	(0.126)	(0.100)	(0.047)	(0.046)	(0.062)	(0.057)	(0.057)	(0.055)	(0.043)	(0.041)	(0.045)	(0.042)	(0.048)	(0.045)	(0.038)	(0.037)
constant	-0.003	0.372**	-0.002	0.048	-0.007	-0.130	0.841***	0.774***	0.708***	0.688***	0.761***	0.687***	0.777***	0.757***	0.742***	0.732***	0.717***	0.560***	0.057**	-0.029
	(0.056)	(0.148)	(0.055)	(0.128)	(0.070)	(0.126)	(0.034)	(0.061)	(0.045)	(0.083)	(0.039)	(0.070)	(0.030)	(0.058)	(0.031)	(0.058)	(0.035)	(0.057)	(0.026)	(0.052)
No. of observations	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366	1,366
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for second respondents living with targeted (in treated locations) and control respondents. All regressions are OLS and use imputed data. In the second column for each outcome I control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table 2.6: Average indirect effect of the three treatments on second interviewees' outcomes, using the inverse probability weighting correction

	Information				Interest in elections		Turnout												Open letter	
	About elections		About voting				Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Newspaper	-0.127	-0.182	0.057	0.053	-0.059	0.039	0.013	0.004	0.047	0.045	0.027	0.006	0.025	0.015	0.043	0.030	0.053	0.055	-0.004	0.002
	(0.156)	(0.135)	(0.131)	(0.123)	(0.132)	(0.102)	(0.061)	(0.062)	(0.067)	(0.068)	(0.063)	(0.067)	(0.056)	(0.058)	(0.056)	(0.058)	(0.057)	(0.059)	(0.032)	(0.036)
Hotline	0.054	-0.034	0.172	0.216*	0.166	0.196*	0.062	0.049	0.020	0.002	0.101*	0.065	0.058	0.041	0.064	0.047	0.104*	0.087*	0.095**	0.086**
	(0.134)	(0.112)	(0.124)	(0.116)	(0.125)	(0.105)	(0.052)	(0.050)	(0.062)	(0.057)	(0.057)	(0.056)	(0.048)	(0.047)	(0.049)	(0.048)	(0.055)	(0.052)	(0.042)	(0.043)
Civic education	-0.020	-0.054	0.074	0.103	0.090	0.128	0.065	0.057	0.057	0.054	0.083	0.054	0.062	0.055	0.070	0.065	0.094*	0.082*	0.038	0.031
	(0.136)	(0.124)	(0.112)	(0.108)	(0.135)	(0.109)	(0.054)	(0.052)	(0.069)	(0.063)	(0.063)	(0.063)	(0.050)	(0.049)	(0.051)	(0.050)	(0.054)	(0.049)	(0.036)	(0.036)
constant	0.025	0.418***	-0.033	-0.092	0.023	-0.299*	0.856***	0.741***	0.744***	0.649***	0.772***	0.738***	0.796***	0.732***	0.763***	0.713***	0.725***	0.534***	0.052**	-0.040
	(0.094)	(0.162)	(0.076)	(0.160)	(0.091)	(0.158)	(0.046)	(0.076)	(0.046)	(0.098)	(0.047)	(0.095)	(0.042)	(0.073)	(0.042)	(0.075)	(0.043)	(0.074)	(0.021)	(0.039)
Adjusted R-squared	-0.002	0.077	-0.001	0.117	0.004	0.228	0.001	0.044	-0.004	0.058	0.004	0.031	0.001	0.041	0.002	0.042	0.009	0.068	0.013	0.030
No. of observations	442	410	442	410	442	410	411	382	411	382	411	382	411	382	411	382	411	382	442	410
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for second respondents living with targeted (in treated locations) and control respondents. All regressions are OLS and use as sampling weights the inverse of the probability of the individual being selected into the second interviewees sample. In the second column for each outcome I control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

I try to empirically correct for the selection into the sample of second interviewees using two methods: multiple imputation and a correction through inverse probability weighting.

I use multiple imputation method to replace the missing values of outcome and control variables, and re-estimate the average effects of the campaign. Here I will focus on the regressions of the second interviewees' outcomes, but the regressions of the main respondents outcomes are reported in the next Chapter.

Multiple imputation is only valid if the data is missing at random, which implies that the imputation model must include all the covariates that affect missingness. Otherwise the parameter estimates will be biased (*Carpenter, Kenward and Vansteelandt, 2006*). I included in the imputation model all the variables that I use in the empirical analysis, other characteristics of the household and of the respondents, characteristics of the enumeration area, and interactions between the interventions and characteristics of the household and respondents. I used multivariate normal regression.[42] Given the high proportion of missing information, I used 70 imputations.

The estimates of the second interviewees' regressions using the imputed data are presented in *Table 2.5*. These estimates seem to corroborate the findings presented above. In hotline areas there is a positive indirect effect on interest and turnout of second interviewees. In fact, using this data the coefficient of the hotline is significant in more regressions of turnout proxies and it also tends to be higher than when using the original complete data. In civic education areas I also find a positive effect on turnout, in average of 6.9 percentage points; but this effect is only significant in the regression of the measure index 2 with controls.

Like multiple imputation, the inverse probability weighting correction also assumes that I observe all variables that affect selection. Based on that assumption, I can consistently estimate the probability of selection of second interviewees; and use the inverse of those probabilities as sampling weights in the OLS estimation of the indirect effects of the interventions.

The probability of selection was estimated with base on a broad set of baseline characteristics of the main respondent and the household, on a set baseline outcome variables of the main respondent, and on the interaction between these covariates and the three interventions.[43] I find

[42] Given than most variables are categorical, I 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 often normal imputation models have a good performance for linear regressions, even when the variables are non-normal.

[43] As there is no baseline information for second interviewees, I have to rely on main respondent and household

that the probability of selecting a second interviewee is highly correlated with the household size, the geographical location and whether the household was selected to not receive the treatment.[44] The recalculated estimates of second interviewee's regressions are presented in *Table 2.6*. In terms of the hotline treatment, I find effects on interest and turnout similar to the ones I had identified in *Table 2.4*, but in addition I also find a positive effect on the index information about voting. This last piece of evidence suggests that there may be some information sharing in households targeted by the hotline. In terms of civic education, I find again a positive indirect effect on turnout, but this effect is only significant in the regressions of the measure based on interviewers' assessment.

Overall, these two exercises corroborate the findings presented above. However, they also suggest that the effects associated with the hotline are more robust than the ones associated with the civic education.

Both methods, multiple imputation and the inverse probability weighting, presume selection on observables. If the selection of second interviewees depends on unobservable variables, the estimates presented here are likely to be biased.[45]

information to estimate the probability of selecting a second interviewee.

[44] The estimates of different specifications of the model of selection are presented in *Table A.1* in Appendix A.

[45] The Altonji ratio for the significant coefficients presented in *Table 2.4* varies between 2.5 (for the coefficient of the hotline in the regression of the turnout measure based on the finger) and 743.7 (for the coefficient of civic education in the regression of interviewer's assessment), with a median ratio of 6.5. This means that to attribute the entire effects identified here to selection, selection on unobservables would have to be at least 2.5 times greater than selection on observables (*Nunn and Wantchekon. 2011*).

Table 2.7: Average indirect effects on second interviewee's outcomes considering only those who reported to have access to a different phone than the main respondent

	Information				Interest in elections		Turnout												Open letter	
	About elections		About voting				Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Hotline	0.045	0.032	0.089	0.190	0.253*	0.209	0.060	0.067	0.047	0.053	0.111*	0.107	0.063	0.072	0.070	0.079	0.119**	0.119*	0.110*	0.118*
	(0.180)	(0.149)	(0.155)	(0.152)	(0.143)	(0.139)	(0.053)	(0.060)	(0.067)	(0.063)	(0.063)	(0.066)	(0.048)	(0.053)	(0.048)	(0.052)	(0.061)	(0.061)	(0.060)	(0.068)
Civic education	-0.040	-0.138	0.126	0.060	0.073	0.118	0.060	0.052	0.065	0.067	0.093	0.074	0.070	0.062	0.082	0.072	0.130**	0.136**	0.034	0.041
	(0.175)	(0.158)	(0.147)	(0.142)	(0.166)	(0.124)	(0.054)	(0.055)	(0.076)	(0.067)	(0.074)	(0.080)	(0.050)	(0.050)	(0.051)	(0.050)	(0.056)	(0.055)	(0.052)	(0.059)
constant	0.032	0.458**	0.021	0.225	-0.009	-0.482**	0.868***	0.622***	0.735***	0.490***	0.779***	0.495***	0.793***	0.574***	0.757***	0.546***	0.725***	0.293***	0.097***	-0.061
	(0.126)	(0.228)	(0.098)	(0.236)	(0.103)	(0.231)	(0.042)	(0.118)	(0.044)	(0.128)	(0.050)	(0.133)	(0.037)	(0.107)	(0.036)	(0.105)	(0.045)	(0.108)	(0.036)	(0.082)
Adjusted R-squared	-0.009	-0.001	-0.006	0.122	0.006	0.309	-0.002	0.050	-0.007	0.082	0.008	0.062	0.002	0.056	0.007	0.070	0.023	0.144	0.007	0.005
No. of observations	191	179	191	179	191	179	178	168	178	168	178	168	178	168	178	168	178	168	191	179
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Other caveat of the empirical analysis is that the effect of the intervention on the second interviewees may result not exclusively from influence but also from the sharing of the mobile phone or the newspaper (common exposure). In hotline and civic education areas, 43 percent of the second interviewees living with targeted respondents (93 individuals) reported to have access to the same mobile number that the main respondent did. This means that when measuring the effects of the interventions on these individuals I may actually be measuring the direct effect of the text messages, instead of measuring the effect of interpersonal influence exerted by the treated individual. In order to investigate if my findings in the civic education and hotline areas are being driven by the common access to the cell phone, I estimate the effects of those interventions considering only the second interviewees who reported to have access to a different phone than the main respondent. The estimates are displayed in *Table 2.7*. These point estimates are not very far from the ones presented in *Table 2.4*. In addition, even using this limited sample, I find that the hotline has significant indirect effects in the same measures as when using the sample of all second interviewees; and that the indirect effect of the civic education on the turnout measure based on the interviewer assessment remains significant. This evidence corroborates the argument that the effects of the hotline and civic education on second interviewees result from influence and not direct exposure to the interventions messages.

2.5.2 Structural form specification

This section describes and discusses the empirical findings using the structural form specification.

I start by presenting and discussing the estimation results of the two-step model that uses the three interventions as instruments to estimate the main respondent's outcome variables.

The results of the first step regression are very similar to the ones presented in *Table 2.3*.[46] As I find no significant evidence that the interventions affect the interest of targeted respondents, no second step regressions were estimated for this outcome. *Table 2.8* presents the estimates of coefficient β in equation (2.5).[47] Each cell of this table corresponds to a different second step

[46] In order to maintain reasonable degrees of freedom, in these regressions I did not include the characteristics of the enumeration and I only included the characteristics of main respondents that are unbalanced across control and treatment areas. This avoids further reduction in the estimation sample and limits the number of parameters to estimate.

[47] The standard errors presented in this Table were obtained using Hardin's sandwich cluster-robust estimator. In *Table A.2* in Appendix I present the bootstrapped standard errors.

regression. The value presented in cell (a, b) - in row a and column b - is the estimated coefficient of main respondent's standardized predicted outcome A (head of row a) in the regression of second interviewee's outcome B (head of column b).[48] The table only includes second interviewees' outcomes in which at least one of main respondent's outcomes had a significant effect.

I find no significant evidence that information of main respondents affects second interviewees. In terms of behavior, main respondent's turnout affects second interviewee's interest in elections and turnout. This evidence is consistent with the finding that influence assumes, at least partially, the form of social control. However, because there is also evidence of main respondent's influence on second interviewee's preferences, I cannot exclude that the change in behavior was also, at least partially, triggered by that change in preferences. It is possible that main respondents make second interviewees feel more interested in the elections, and this increases their probability of voting.

This evidence corroborates the conclusions presented in the previous section. Using a two-step model with the interventions as instruments I also only find evidence of a significant effect on the second interviewee's preferences and turnout. There is a transmission of the interest in elections and of the decision whether to vote, within the household.

An alternative to the estimation of the two step model would be to adopt an instrumental variable approach. As a robustness check, I estimated effects of main respondent's outcomes on second interviewee's outcomes using the two stage least squares estimator.

Assuming that the second interviewees were not directly exposed to the campaign, I use the three interventions - newspaper, hotline and civic education - as instruments. I regress each second interviewee's outcome on each main respondent's outcome, controlling for the main respondent's and second interviewee's characteristics (that are not balanced across control and treatment groups) and geographical location. The standard errors are adjusted by clustering at local level.

[48] To facilitate the interpretation of the coefficients, I standardized main respondents' predicted outcome. The coefficients of these variables should be interpreted in terms of standard deviation units.

Table 2.8: Model with interventions as instruments - Second step estimation results

Main respondent's predicted outcome	Second interviewee's outcomes					
	Interest in elections	Turnout				
		Self-reported	Finger measure	Index 1	Index 2	Interviewer's assessment
	(1)	(2)	(3)	(4)	(5)	(6)
Self-reported turnout	0.191**	0.058*	0.072*	0.050	0.053	0.079**
	(0.091)	(0.035)	(0.041)	(0.032)	(0.033)	(0.040)
Adjusted turnout	0.267**	0.086	0.104	0.075	0.082	0.116*
	(0.134)	(0.056)	(0.064)	(0.051)	(0.052)	(0.062)
Finger measure	0.217	0.082*	0.092	0.071	0.075	0.099*
	(0.135)	(0.047)	(0.058)	(0.044)	(0.047)	(0.059)
Index 1	0.213**	0.063	0.080*	0.055	0.060	0.089**
	(0.101)	(0.040)	(0.046)	(0.037)	(0.038)	(0.045)
Index 2	0.236**	0.072	0.089*	0.064	0.070*	0.102**
	(0.112)	(0.045)	(0.051)	(0.041)	(0.042)	(0.049)
Interviewer's assessment	0.177**	0.059*	0.070*	0.052*	0.057*	0.080**
	(0.081)	(0.033)	(0.039)	(0.030)	(0.031)	(0.036)

Note: Each cell corresponds to a linear model regression of an outcome variable of the second interviewee on the standardized predicted main respondent's outcome variable. The values displayed are the coefficients of standardized predicted main respondent's outcome variables (displayed in rows) in the regression of second interviewee's outcome variables (displayed in columns). Standard errors reported between parenthesis; these were obtained using Hardin's sandwich cluster-robust covariance estimator.

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

Table 2.9: Instrumental variable approach - Second step estimation results

Main respondent's predicted outcome	Second interviewee's outcomes	
	Interest in elections	Turnout
		Interviewer's assessment
	(1)	(2)
Self-reported turnout	0.409*	0.189*
	(0.227)	(0.101)
Adjusted turnout	0.424	0.174
	(0.274)	(0.107)
Finger measure	0.552	0.287
	(0.356)	(0.193)
Index 1	0.792	0.371*
	(0.484)	(0.212)
Index 2	0.751	0.349*
	(0.475)	(0.200)
Interviewer's assessment	0.523*	0.234*
	(0.285)	(0.125)

Note: Each cell corresponds to the estimation of the effects of one of main respondent's outcomes on one of second interviewee's outcomes using an instrumental variable model. The values displayed are the coefficients of main respondent's outcome variables (displayed in rows) in the regressions of second interviewee's outcome variables (displayed in columns). Regressions include observations for second interviewees living in treated and control areas. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table 2.9 reports the estimates of the parameters of interest considering the complete sample of households for which I have information regarding both respondents.[49] Similarly to *Table 2.6*, the value presented in cell (a, b) - in row a and column b - is the estimated coefficient of main respondent's outcome A (head of row a) in the regression of second interviewee's outcome B (head of column b). The table only includes second interviewees' outcomes in which at least one of the main respondent's outcomes had a significant effect. Only the effects of main respondent's outcomes for which at least one of the interventions is an informative instrument (i.e. significant at least at 10 percent level) are considered. The conclusions presented in the previous section are corroborated.

Now, I look at the estimation results of the two-step model that controls for the direct effect of the interventions on the second interviewee's outcome and uses the interaction terms as instruments.

The estimated coefficients of the interventions and the interactions terms in the first step regression are displayed in *Table A.3* in Appendix. The hypotheses that the interaction terms are jointly equal to zero cannot be rejected in the regressions of main respondent's information and interest indices. This suggests that the interactions are not informative instruments for these outcomes.[50]

Table 2.10. presents the estimates of the second step regressions of second interviewee's turnout (measure based on the interviewer's assessment) on main respondent's turnout (all proxies). The regressions presented in columns (1) to (6) were estimated considering the full sample of second interviewees; in columns (7) to (12) the main and second interviewees who had access to the same mobile phone were excluded.[51]

[49] If I estimate this model excluding the households in treated areas that did not receive the treatment, the estimation sample of turnout regressions is reduced to 287 observations. With this very limited sample I still find significant effects of main respondent's turnout, measured by the interviewer's assessment, on second interviewee's interest and turnout (also measured with the interviewer assessment). Here I present the estimates of the regressions considering all the households for which I have information (the estimation sample of turnout regressions increases to 329 observations).

[50] In these circumstances, I cannot reliably estimate the impact of main respondent's information and interest on turnout using this specification.

[51] To obtain an estimation sample as big as possible, I included the observations from untargeted households (those that were located in treated areas, but where the treatment was not delivered). Nevertheless, similar estimates are obtained if I exclude those observations (*Table A.4* in Appendix).

Table 2.10: Model with interactions as instruments - Second step estimation results

Covariates	Dependent variable: Second interviewee's turnout, measured by interviewer assessment											
	Sample of all second interviewees						Excluding interviewees who share cell phone					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.021	0.023	0.022	0.021	0.017	0.015	-0.014	-0.014	-0.011	-0.017	-0.019	-0.025
	(0.049)	(0.044)	(0.051)	(0.049)	(0.050)	(0.050)	(0.063)	(0.064)	(0.065)	(0.064)	(0.065)	(0.065)
Hotline	0.022	0.017	0.037	0.027	0.023	0.029	0.005	-0.004	0.019	0.004	0.001	0.014
	(0.051)	(0.058)	(0.050)	(0.050)	(0.051)	(0.050)	(0.067)	(0.071)	(0.068)	(0.066)	(0.067)	(0.062)
Civic education	0.052	0.039	0.046	0.052	0.046	0.046	0.015	0.000	0.005	0.012	0.005	0.010
	(0.046)	(0.051)	(0.050)	(0.046)	(0.047)	(0.047)	(0.057)	(0.062)	(0.063)	(0.057)	(0.059)	(0.057)
Main respondent's turnout (standardized)												
Self-reported	0.053*						0.064*					
	(0.029)						(0.034)					
Adjusted		0.072						0.086*				
		(0.044)						(0.046)				
Finger measure			0.048*						0.064*			
			(0.029)						(0.037)			
Index 1				0.053*						0.065*		
				(0.028)						(0.034)		
Index 2					0.059*						0.072*	
					(0.031)						(0.036)	
Interviewer's assessment						0.046*						0.061**
						(0.024)						(0.030)
Adjusted R-squared	0.032	0.031	0.031	0.032	0.032	0.032	0.045	0.046	0.044	0.046	0.046	0.046
No. of observations	466	466	466	466	466	466	274	274	274	274	274	274
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include observations for second respondents living in treated and control areas. All regressions are OLS and in all regressions I control for demographic characteristics and province dummies. Standard errors reported between parenthesis; these were obtained using Hardin's sandwich cluster-robust covariance estimator.

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

The estimates of the coefficients of main respondent's turnout, coefficient μ in equation (2.7), are always positive, and in most cases they are significant at 10 percent level. On the other hand, the coefficients of the interventions are never significant and in most cases relatively low. When I exclude the interviewees that shared the phone, the coefficient of main respondent's turnout increases and the coefficients of the intervention decrease, becoming very close to zero in most regressions. This evidence corroborates the conclusion that the effect of the interventions on second interviewees results from main respondents' influence.

Furthermore, using this specification I only obtain evidence that main respondent's turnout affects second interviewee turnout. This corroborates the argument that the campaigns' effects are transmitted to second interviewees, at least partially, through pressure.[52]

Unfortunately, the structural specification does not allow to clearly explicit the channels of influence of the different interventions. As the different interventions trigger different effects on main respondent's outcomes, the different effects of main respondent's outcomes on second interviewee's outcomes result from the combination of influence effects in the different treatment areas.

2.6 Results: Heterogeneous effects

In this section I examine whether the effect of the campaign on other household members (i.e., secondary targets) varies with age, gender and education. The analysis is restricted to these three characteristics due to data limitations in terms of variables and number of observations.[53]

[52] This model, unlike the previous, does not identify a significant effect of main respondent's outcomes on second interviewee's interest. In fact, in the second step regressions of second interviewee's interest the coefficient of main respondent's outcome tends to be negative and insignificant, while the coefficient of the hotline tends to be high and significant (even in the regressions of second interviewees who do not share the phone). The most probable explanation for this is that my instruments (the interactions) are not capturing the effect of the hotline on the main respondent that triggers influence on second interviewee's interest. Another possible explanation would be that the hotline has a direct effect on second interviewees' interest in elections. However, given that there is no significant evidence that this type of intervention has even an impact on the interest of those who were targeted, this explanation is not very plausible.

[53] The estimates of the two-step and instrumental variable models discussed in this section were estimated including the observations from untargeted households (those that were located in treated areas, but where the treatment was not delivered).

2.6.1 Gender

Gender is captured by a dummy variable that assumes the value one when the respondent is male and zero otherwise.

I find no significant heterogeneous indirect effects associated with the gender of the main respondent (the influencer). With respect to second interviewee's (receiver) gender, using the reduced form specification, I find a positive effect on information about elections in newspaper areas; and a positive effect on turnout in civic education areas (but only significant for the measure based on the interviewer's assessment). These effects remain significant when I estimate the heterogeneous indirect effects associated with main and second interviewees' gender jointly (equation (2.9)).

Using the two step and the instrumental variables specification I cannot find evidence of significant heterogeneous effects associated with the gender of any of the interviewees. One reason why this specification does not capture any of those effects might be because each of them approximately concerns only one third of the treated individuals and, hence, gets 'diluted' in the structural model.

Although Mozambique is a society where the patriarchal culture remains strong (*Tvedten, Paulo and Montserrat, 2008*), I find no evidence that men are more likely than women to exercise influence over other household members. In addition, I also do not find evidence that women are more likely to be influenced within the household that men. If anything, it seems that in civic education areas men are more likely to be influenced towards voting than women (within the household context).

2.6.2 Age

I find practically no evidence of significant heterogeneous indirect effects associated with main respondent's age.[54]

[54] The only exception is a significant negative effect of the hotline on information about voting, when including controls.

Table 2.11: Regressions of turnout measures with interaction with second interviewee's age - Reduced form specification

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.023	0.019	0.065	0.066	0.036	0.034	0.031	0.027	0.046	0.043	0.064	0.072
	(0.052)	(0.051)	(0.062)	(0.058)	(0.057)	(0.056)	(0.048)	(0.047)	(0.048)	(0.046)	(0.047)	(0.047)
Hotline	0.067	0.052	0.043	0.034	0.100**	0.074	0.059	0.045	0.066	0.052	0.102**	0.089*
	(0.044)	(0.044)	(0.053)	(0.048)	(0.050)	(0.051)	(0.040)	(0.041)	(0.041)	(0.042)	(0.047)	(0.047)
Civic education	0.085*	0.083*	0.092	0.098*	0.109**	0.084	0.077*	0.079*	0.088**	0.091**	0.132***	0.113***
	(0.044)	(0.044)	(0.063)	(0.055)	(0.055)	(0.052)	(0.042)	(0.040)	(0.044)	(0.041)	(0.043)	(0.041)
Second interviewee's age	0.006***	0.007***	0.010***	0.011***	0.005**	0.007**	0.006***	0.006***	0.006***	0.007***	0.008***	0.009***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Newspaper x age	**-0.002**	**-0.002**	**-0.007***	**-0.007****	**-0.004**	**-0.004**	**-0.002**	**-0.002**	**-0.002**	**-0.003**	**-0.004**	**-0.004**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Hotline x age	**-0.003**	**-0.004**	**-0.006***	**-0.005**	**-0.002**	**-0.002**	**-0.003**	**-0.003**	**-0.004****	**-0.004***	**-0.004**	**-0.005**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Civic x age	**-0.004***	**-0.005***	**-0.010*****	**-0.010*****	**-0.003**	**-0.004**	**-0.005****	**-0.005****	**-0.006*****	**-0.006****	**-0.004**	**-0.005**
	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
constant	0.651***	0.593***	0.399***	0.389***	0.594***	0.520***	0.611***	0.591***	0.558***	0.532***	0.462***	0.343***
	(0.089)	(0.113)	(0.099)	(0.129)	(0.092)	(0.130)	(0.078)	(0.103)	(0.075)	(0.102)	(0.090)	(0.114)
Adjusted R-squared	0.029	0.056	0.030	0.090	0.017	0.055	0.023	0.061	0.029	0.079	0.053	0.085
No. of observations	408	397	408	397	408	397	408	397	408	397	408	397
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 2.11 displays the estimates of the campaign's heterogeneous indirect effects associated with the second interviewee's age. These results show that, on one hand, younger second interviewees tend to vote less than older ones: the age coefficient is positive and significant in all turnout regressions. On the other hand, they show that the indirect effect of civic education, on turnout is larger for younger second interviewees than for older ones: the interaction coefficient is negative in all turnout regressions and significant in the regressions of self-reported and adjusted turnout, index 1 and index 2. These results are corroborated when I estimate the heterogeneous effects associated with both respondents' age.[55]

Using the two step specification, the coefficient of the interaction between main respondent's turnout and the second interviewee's age is always negative in the regressions of second interviewees' turnout, but never statistically significant. In the IV model, the coefficient of this interaction term is also always negative, and in some regressions significant or borderline significant (*Table A.5* in Appendix). Therefore, these estimates seem to corroborate the finding that the indirect effect of the campaigns on second interviewees varies negatively with their age.

I find no evidence that age has an impact on one' ability to influence other household members towards voting. This result is particularly interesting as age tends to be one of the determinants of the family hierarchy, and since in African societies there is a traditional value of respect for elders.

Evidence suggests that, on average, younger people vote less than older people, but the indirect effects of the interventions seem to be contributing to close that gap. I believe this evidence may be the result of two related phenomena. On one hand, younger people may be more vulnerable to pressure within the household and, hence, more likely to be influenced towards voting. On the other hand, older people, besides being less vulnerable to pressure, may also feel less inclined to vote once they know other household members will vote – a free-riding effect.[56]

[55] According to Whipple's Index the reported age of main and second interviewees is relatively accurate. Nevertheless, I re-estimated these equations replacing the variable age with age group. The conclusions are the same.

[56] It could also be argued that the potential indirect effect of the interventions is smaller for older people because they already have a high propensity to vote. So, the negative heterogeneous effect on turnout associated with the receiver's age would not reflect that older people are less prone to be influenced, but simply that they already conform to the "enforced" behavior - a saturation effect. I believe this interpretation is not supported by the data. First, if such saturation effect existed it would manifest itself for both main and second interviewees. But I find no evidence that experimental subjects' response to the interventions depends on their age. Second, I find that the average turnout of second interviewees aged forty or older (approximately 25 percent of the sample) is higher in control areas than in treated areas (for all measures, but only close to significant in the case of self-reported).

2.6.3 Education

Although I find practically no significant heterogeneous effects associated with literacy of main or second interviewees, there is evidence of such effects when I raise the threshold of education to complete primary school (5 years of compulsory school plus 2 years). In the following regressions education is captured by a dummy variable that equals one if the respondent has completed at least primary school and zero otherwise.

Regarding the education level of the main respondent, I find that the interactions with the newspaper and the hotline are negative in the regressions of all turnout measures. However, only two of these interaction terms are significant: the interaction with the hotline in the regression of the adjusted turnout, and the interaction with the newspaper in the regression of the measure based on the interviewer's assessment (both only when I include controls).

The results presented in *Table 2.12* show that the newspaper effects on second interviewee's outcomes depend on his or her education. These estimates show that the impact of the newspaper on second interviewee's probability of voting is 18.5 to 23.7 percentage points higher (depending on the turnout measure considered) for those who completed primary school than for those who did not.

When I jointly estimate the heterogeneous effects associated with both interviewees' education, the findings listed above are corroborated and even strengthened. The estimates displayed in *Table 2.13* show more significant heterogeneous effects on turnout measures, especially associated with main respondent's education; and the absolute values of the point estimates the interaction terms are higher than the ones in the regressions where I only include the interactions with one of the interviewer's education.

This provides some corroboration to the idea of free-ridding by elder household members.

Table 2.12: Regressions of turnout measures with interaction with second interviewee's education level - Reduced form specification

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.023	0.022	0.064	0.073	0.044	0.041	0.031	0.030	0.047	0.045	0.061	0.077*
	(0.050)	(0.050)	(0.059)	(0.058)	(0.056)	(0.057)	(0.047)	(0.046)	(0.046)	(0.045)	(0.046)	(0.046)
Hotline	0.060	0.051	0.041	0.031	0.101**	0.073	0.055	0.043	0.062	0.051	0.098**	0.087*
	(0.043)	(0.044)	(0.050)	(0.048)	(0.050)	(0.051)	(0.039)	(0.040)	(0.040)	(0.041)	(0.048)	(0.047)
Civic education	0.069	0.081*	0.072	0.093*	0.091*	0.084	0.062	0.075*	0.071*	0.086**	0.107**	0.110***
	(0.043)	(0.042)	(0.060)	(0.053)	(0.054)	(0.052)	(0.040)	(0.039)	(0.041)	(0.040)	(0.045)	(0.041)
Second respondent has primary school	-0.104**	0.010	-0.160*	0.005	0.040	0.136*	-0.101**	0.007	-0.096*	0.020	-0.059	0.043
	(0.052)	(0.065)	(0.083)	(0.088)	(0.074)	(0.079)	(0.049)	(0.061)	(0.053)	(0.063)	(0.057)	(0.062)
Newspaper x education	**0.219*** **	**0.198** **	**0.300*** **	**0.237** **	**0.121**	**0.148**	**0.212*** **	**0.188** **	**0.211*** **	**0.185** **	**0.223*** **	**0.212*** **
	(0.077)	(0.081)	(0.112)	(0.112)	(0.101)	(0.098)	(0.073)	(0.077)	(0.077)	(0.080)	(0.077)	(0.077)
Hotline x education	**0.093**	**0.063**	**0.107**	**0.053**	**-0.065**	**-0.066**	**0.090**	**0.056**	**0.091**	**0.051**	**0.052**	**0.059**
	(0.072)	(0.076)	(0.109)	(0.108)	(0.096)	(0.096)	(0.069)	(0.073)	(0.073)	(0.076)	(0.087)	(0.086)
Civic x education	**0.137***	**0.084**	**0.192**	**0.097**	**0.045**	**-0.002**	**0.144** **	**0.092**	**0.163** **	**0.110**	**0.132**	**0.072**
	(0.076)	(0.083)	(0.120)	(0.121)	(0.093)	(0.089)	(0.072)	(0.078)	(0.077)	(0.081)	(0.084)	(0.077)
constant	0.911***	0.728***	0.821***	0.624***	0.750***	0.615***	0.850***	0.726***	0.815***	0.687***	0.753***	0.500***
	(0.042)	(0.082)	(0.051)	(0.096)	(0.060)	(0.098)	(0.038)	(0.072)	(0.041)	(0.073)	(0.047)	(0.084)
Adjusted R-squared	0.010	0.062	0.007	0.086	0.011	0.063	0.011	0.067	0.013	0.081	0.021	0.091
No. of observations	426	397	426	397	426	397	426	397	426	397	426	397
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 2.13: Regressions with interactions with both respondents education - Reduced form specification

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.018	0.016	0.061	0.070	0.040	0.036	0.028	0.026	0.045	0.042	0.055	0.070
	(0.048)	(0.049)	(0.059)	(0.059)	(0.056)	(0.056)	(0.045)	(0.045)	(0.046)	(0.045)	(0.044)	(0.044)
Hotline	0.059	0.048	0.045	0.032	0.099**	0.068	0.054	0.041	0.063	0.049	0.094**	0.081*
	(0.041)	(0.043)	(0.050)	(0.048)	(0.050)	(0.052)	(0.037)	(0.039)	(0.039)	(0.040)	(0.047)	(0.047)
Civic education	0.065	0.076*	0.070	0.090*	0.088	0.080	0.060	0.072*	0.069*	0.084**	0.103**	0.104***
	(0.041)	(0.041)	(0.061)	(0.053)	(0.054)	(0.052)	(0.039)	(0.038)	(0.041)	(0.039)	(0.043)	(0.039)
Main respondent has primary schooling	0.103	0.137*	0.070	0.119	0.101	0.142*	0.080	0.108	0.059	0.091	0.139**	0.183***
	(0.067)	(0.072)	(0.091)	(0.092)	(0.076)	(0.080)	(0.064)	(0.068)	(0.064)	(0.068)	(0.068)	(0.069)
Second respondent has primary schooling	-0.132**	-0.031	-0.179**	-0.029	0.014	0.098	-0.122**	-0.027	-0.112**	-0.009	-0.096	-0.010
	(0.053)	(0.068)	(0.080)	(0.091)	(0.075)	(0.081)	(0.049)	(0.063)	(0.052)	(0.066)	(0.061)	(0.065)
Newspaper x main education	**-0.201****	**-0.212****	**-0.161**	**-0.157**	**-0.130**	**-0.135**	**-0.180***	**-0.182****	**-0.150**	**-0.151***	**-0.220****	**-0.228****
	(0.098)	(0.095)	(0.127)	(0.123)	(0.110)	(0.109)	(0.093)	(0.091)	(0.094)	(0.091)	(0.093)	(0.090)
Hotline x main education	**-0.167***	**-0.188***	**-0.230***	**-0.250***	**-0.141**	**-0.167**	**-0.138**	**-0.150**	**-0.127**	**-0.142**	**-0.169**	**-0.196**
	(0.095)	(0.107)	(0.126)	(0.134)	(0.111)	(0.125)	(0.090)	(0.102)	(0.091)	(0.103)	(0.108)	(0.123)
Civic x main education	**-0.012**	**-0.071**	**-0.032**	**-0.102**	**-0.051**	**-0.140**	**-0.010**	**-0.061**	**0.011**	**-0.045**	**-0.026**	**-0.118**
	(0.083)	(0.093)	(0.127)	(0.131)	(0.086)	(0.091)	(0.081)	(0.089)	(0.083)	(0.091)	(0.081)	(0.085)
Newspaper x second education	**0.271*****	**0.253*****	**0.342*****	**0.283****	**0.155**	**0.185***	**0.258*****	**0.235*****	**0.249*****	**0.225*****	**0.280*****	**0.271*****
	(0.084)	(0.088)	(0.112)	(0.115)	(0.103)	(0.098)	(0.078)	(0.082)	(0.080)	(0.084)	(0.083)	(0.080)
Hotline x second education	**0.142***	**0.122**	**0.180**	**0.138**	**-0.025**	**-0.017**	**0.131***	**0.104**	**0.129**	**0.096**	**0.099**	**0.114**
	(0.082)	(0.094)	(0.110)	(0.117)	(0.108)	(0.115)	(0.078)	(0.089)	(0.081)	(0.093)	(0.104)	(0.111)
Civic x second education	**0.127**	**0.101**	**0.195**	**0.129**	**0.052**	**0.038**	**0.137***	**0.108**	**0.150***	**0.121**	**0.123**	**0.100**
	(0.080)	(0.090)	(0.133)	(0.138)	(0.094)	(0.091)	(0.076)	(0.085)	(0.081)	(0.090)	(0.082)	(0.077)
constant	0.883***	0.682***	0.803***	0.582***	0.723***	0.565***	0.829***	0.689***	0.799***	0.657***	0.715***	0.438***
	(0.055)	(0.096)	(0.063)	(0.105)	(0.071)	(0.112)	(0.050)	(0.087)	(0.052)	(0.087)	(0.058)	(0.093)
Adjusted R-squared	0.019	0.070	0.011	0.088	0.008	0.062	0.018	0.072	0.018	0.082	0.032	0.104
No. of observations	426	397	426	397	426	397	426	397	426	397	426	397
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Using the two step model specification, I find no consistent evidence of heterogeneous effects associated with education.[57] On the other hand, the IV estimates corroborate the findings listed above. First, the IV estimates of the interaction with main respondent's education are always negative in the regressions of turnout measures, although not significant. Second, the coefficient of the interaction with second interviewee's education is always positive and in a few cases it is significant (*Table A.6* in Appendix).

Overall, the indirect effects of the campaigns seem smaller in households where the main respondent has completed primary school. This evidence suggest that individuals with higher education tend to exert less influence within the household.

Although one might expect that individuals with lower education would be more prone to be influenced than individuals with higher education, data suggests otherwise. It seems that second interviewee's education enhances main respondent's influence on second interviewee's information and turnout. A potential explanation for this is that individuals who received primary education pay more attention to political matters and, thus, have a higher probability of being the interlocutors of conversations about that topic.

Given these two effects together, a possible scenario is that primary education raises individuals' electoral information and awareness. So, in households where the experimental subject did not complete primary education but the second interviewee did, the second interviewee is in a better position to understand the purpose of the campaign and explain it to the main respondent. This type of interaction between the two individuals contributes to the second interviewee assimilating new information, feeling more empowered and, thus, more likely to vote. As these heterogeneous effects are significant mainly in newspaper areas, an alternative explanation would be that main respondents who have not completed primary education may ask other people in the household who have higher schooling to read them the newspaper. In these circumstances, the indirect effect of the newspaper on second interviewees would result from the direct exposure to the newspaper and not only from influence of the main respondent. Nevertheless, even in this situation the main respondent may be exerting influence by conditioning the access of other household members to the newspaper.

[57] In fact, the coefficient of the interaction between the main respondent's outcome and second interviewee's education is even negative in the regressions of some turnout measures.

2.6.4 Robustness checks

To examine the sensitivity of these findings to attrition I re-estimated the reduced form regressions using the inverse probability weighting correction. All the findings listed above are corroborated (*Tables A.7* and *A.8* in Appendix).

The specification of a model testing for heterogeneous effects associated with the difference between the respondents' characteristics may be questionable.[58] Nevertheless, I also estimated the heterogeneous effect associated with the variables 'different gender' (dummy variable that assumes value one if the interviewers are of different gender), 'age difference' and 'main respondent has higher education' (dummy variable that assumes value one if main respondent has higher level of education that the second interviewee). With respect to the variable 'different gender', I find a negative heterogeneous effect of newspaper in the regressions of interest in elections and a negative heterogeneous effect of civic education in adjusted turnout.[59] The interaction between civic education and the age difference between the two interviewees is positive and significant in all regressions of second interviewee's turnout measures. This corroborate the conclusion that younger interviewees are more vulnerable to pressure towards voting (*Table A.9* in Appendix). The interaction between newspaper and the dummy that equals one if the main respondent has higher education that the second interviewee is negative and significant in the regressions of all turnout measures with the exception of the one based on the finger.

In order to examine into what extent the gender and age pattern of heterogeneous effects could be capturing an effect associated with the relationship between the two respondents, I also tested for heterogeneous effects associated with the type of relationship between the main and second interviewees.[60] I find no significant effect associated with the relationships: spouses, parents - children, men - wife and women - husband.[61]

[58] The functional form including difference in characteristics is less meaningful for individuals in control areas, than the model including the characteristics directly. The interpretation of the coefficients of variables like 'different gender' and 'age difference' is problematic, as they may capture correlations between second interviewee's characteristic and the outcome of interest.

[59] This may suggest that the difference between the genders of the interviewees has an impact separately from the gender of the interviewees.

[60] As in the case of differences in characteristics, the functional form including the relationship between the respondents may not be meaningful for individuals in control areas.

[61] It is hard to reliably test the heterogeneous effects associated with the relationships father - children and mother - children because there are few observation of such pairs.

2.7 Conclusion

Interpersonal influence plays an important role in the diffusion of information, opinions and behaviors. The main hypothesis of this study is that voter education campaigns affect not only the targeted subjects but also the people with whom they interact. During the 2009 elections in Mozambique, a field experiment implemented three interventions aimed at increasing political participation: the distribution of a newspaper, the creation of a SMS hotline to report electoral problems and an electoral education campaign. This paper investigates the spillover of these interventions' effects to people who lived with treated subjects. To do this I use an innovative dataset that includes post-treatment outcomes for experimental subjects and for a second person living in the same household. Although I cannot completely rule out the risk of common exposure, I believe that the main reason for such spillovers is interpersonal influence. In order to better understand that process, this study proposes a simple model of human behavior subject to interpersonal influence that distinguishes influence at three levels: information, preferences, and behavior.

Findings confirm that the interventions affect people living with treated subjects. This implies that most impact evaluations of voter mobilization interventions tend to underestimate the total effect and, hence, the cost-effectiveness of the interventions.

The results also show that different interventions trigger influence at different levels. In households that received the newspaper treatment, basically an information intervention, there is almost no evidence of influence. Although there was an increase of the information outcomes of the treated individuals, this additional information did not reach the other household members. In households that received the hotline treatment, similar to an empowerment intervention, there is evidence of influence directly on interest in elections, as well as of an increase in turnout.

In households that received the civic education treatment, similar to a call out to vote campaign, there is only robust evidence of influence directly on individual turnout. This suggests that these people were more likely to vote, not because they were more interested or better informed, but due to social pressure. Since the most significant effect of civic education intervention on treated individuals was also on turnout, I conclude that this type of intervention works mainly through persuasion: it pressures treated people to vote, who in their turn also pressure people

with whom they interact.

In this chapter, I also explored the determinants of interpersonal influence regarding the decision to vote, within Mozambican households. Contrary to common wisdom, I found no evidence that gender or age determine one's ability to exercise influence over other household members regarding the decision to vote. With respect to the characteristics of the target of influence, the results suggest that younger individuals are more likely to respond to social pressure towards voting than older ones.

Contrary to what most economic models would assume, these results show that the transmission of these interventions' effects does not occur through information sharing, but mostly through social pressure. So, although interpersonal influence extends the reach of voter education interventions, one should be aware that a significant part of the message is lost along the way. Treated individuals seem to transmit to others their own interpretation of the content of the intervention's message.

This chapter focused on interpersonal influence within the household. However, a significant part of interpersonal influence occurs outside the household (for example among friends and neighbors). So, the next chapter focuses on the peer effects at the village level, triggered by the voter education campaign.

Bibliography

[1] Adcock, W. Jr., Hirschman, E. and Goldstucker, J. (1977). "Bank credit card users: An updated profile," in *Advances in Consumer Research*, Vol. 4, ed. William D. Perreault, Jr., Atlanta, GA: Association for Consumer Research.

[2] Aldrich, J. (1993). "Rational choice and turnout," *American Journal of Political Science*, 37(1), 246-278.

[3] Aker, J., Collier, P. and Vicente, P. C. (2013). "Is information power? Using cell phones during an election in Mozambique," Universidade Nova de Lisboa, mimeograph.

[4] Akerlof, G. A. (1997). "Social distance and social decisions," *Econometrica*, 65(5), 1005-1027.

[5] Arthur, W. and Lane, D. (1993). "Information contagion," *Structural Change and Economic Dynamics*, 4(1), 81-104.

[6] Bartels, L. (2000). "Partisanship and voting behavior, 1952-1996," *American Journal of Political Science*, 44(1), 35-50.

[7] Beck, P., Dalton, R., Green, S. and Huckfeldt, R. (2002). "The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices," *American Political Science Review*, 96(1), 57-73.

[8] Booth, A. and Babchuk, N. (1972). "Informal medical opinion leadership among the middle aged and elderly," *The Public Opinion Quarterly*, 36(1), 87-94.

[9] Brancheau, J. and Wetherbe, J. (1990). "The adoption of spreadsheet software: Testing innovation diffusion theory in the context of end-user computing," *Information Systems Research*, 1(2), 115-143.

[10] Bruhn, M. and McKenzie, D. (2009). "In pursuit of balance: Randomization in practice in development field experiments," *American Economic Journal: Applied Economics*, 1(4), 200-32.

[11] Card, D. (1993). "Using geographic variation in college proximity to estimate the return to schooling," NBER WP No. 4483.

[12] Carpenter, J., Kenward, M. and Vansteelandt, S. (2006). "A comparison of multiple imputation and doubly robust estimation for analyses with missing data," *Journal of Royal Statistics Society: Series A (Statistics in Society)*, 169(3), 571–584.

[13] Chan, K. and Misra, S. (1990). "Characteristics of the opinion leader: A new dimension," *Journal of Advertising*, 19(3), 53-60.

[14] Collier, P., and Vicente, P. C. (2009), "Votes and violence: Evidence from a field experiment in Nigeria," BREAD, Working Paper 232.

[15] Cungara, B., Fagilde, G., Garrwtt, J., Uaiene, R. N., and Headey, D. (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/

[16] Davenport, T. (2010). "Public accountability and participation: The effects of a feedback intervention on voter turnout in a low salience election," *Political Behavior*, 32(3), 337-368.

[17] Deaton, A. (2009). "Instruments of development: randomization in the tropics, and the search for the elusive keys to economic development," NBER Working Paper Series, Vol. w14690. Available at SSRN: http://ssrn.com/abstract=1335715, Version of 8th November 2010.

[18] Deci, E., and Ryan, R. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Publishing Co.

[19] Deutsch, M. and Gerard, H. B. (1955). "A study of normative and informational social influences upon individual judgment," *Journal of Abnormal & Social Psychology*, 51(3), 629-636.

[20] De Vletter, F. (2001). "Coping with extreme poverty through traditional skills: The case of the xirundzu basket makers of Mozambique", ILO/SAMAT Discussion Paper no. 18, International Labour Organization, Southern African Multidisciplinary advisory team (ILO/SAMAT), Harare, Zimbabwe. Available at: http://docsfiles.com/pdf_ilo_discussion_paper.html

[21] Dhillon, A. and Peralta, S. (2002). "Economic theories of voter turnout," *Economic Journal*, 112, F332–F352.

[22] Dickerson, M. and Gentry, J. (1983). "Characteristics of adopters and non-adopters of home computers," *The Journal of Consumer Research*, 10(2), 225-235.

[23] Downs, A. (1957). *An Economic Theory of Democracy*, New York: Harper and Row.

[24] Feddersen, T. (2004). "Rational choice theory and the paradox of not voting," *Journal of Economic Perspectives*, 18(1), 99-112.

[25] Fishbein, M., and Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, MA: Addison-Wesley. Retrieved November 2013 from http://people.umass.edu/aizen/f&a1975.html

[26] French, J. (1956). "A formal theory of social power," *Psychological Review*, 63, 181-194.

[27] French, J. and Raven, B. (1959). "The bases of social power," in D. Cartwright (Ed.), *Studies in Social Power*, pp. 150–67. Ann Arbor: Institute for Social Research.

[28] Ekine, S. (Ed.) (2010). *SMS Uprising: Mobile Activism in Africa*, Oxford: Fahamu Books.

[29] Engelen, B. (2006). "Solving the paradox – The expressive rationality of the decision to vote," *Rationality and Society*, 18(4), 419-441.

[30] Fafchamps, M. and Vicente, P. C. (2013). "Political violence and social networks: Experimental evidence from a Nigerian election," *Journal of Development Economics*, 101, 27-48.

[31] Feddersen, T. J. (2004). “Rational choice theory and the paradox of not voting,” *The Journal of Economic Perspectives*, 18(1), 99-112.

[32] Ferejohn, J., and Fiorina. M. (1974). "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review,* 68, 525-36.

[33] Fiorina, M. P. (1976). "The voting decision: Instrumental and expressive Aspects," *The Journal of Politics*, 38(2), 390-413.

[34] Gerber, A. S., Green, D. P. and Larimer, C. W. (2008). "Social pressure and voter turnout: Evidence from a large-scale field experiment," *American Political Science Review*, 102(1), 33-48.

[35] Gine, X., and Mansuri, G. (2011). “Together we will: Experimental evidence on female voting behavior in Pakistan,” World Bank, Policy Research Working Paper 5692.

[36] Hardin, J. W. (2002). "The robust variance estimator for two-stage models," *Stata Journal*, 2(3), 253-266.

[37] Heacock, R. (2009). "Mobile activism in African elections a comparative case study," DigiActive Research Series. Retrieved February 2011 from http://www.scribd.com/doc/17352255/RD-2-Mobile-Activism-in-African-Elections-a-Comparative-Case-Study.

[38] Hole, A. R. (2006). "Calculating Murphy-Topel variance estimates in Stata: A simplified procedure," *Stata Journal*, 6(4), 521-529.

[39] Huckfeldt, R. and Sprague, J. (1995). *Citizens, Politics and Social Communication: Information and Influence in an Election Campaign*, New York: Cambridge Univ. Press.

[40] Katz, E. (1957). "The two-step flow communication: An up-to-date report on an hypothesis," *Public Opinion Quarterly*, 21, 61-78.

[41] Katz, E., and Lazarsfeld, P. (1955). *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.

[42] Kingdon, J. (1970). "Opinion leaders in the electorate," *The Public Opinion Quarterly*, 34(2), 256-261.

[43] Kling, J., Liebman, J. and Katz, L. (2007): "Experimental analysis of neighborhood effects," *Econometrica*, 75(1), 83–119.

[44] Knack, S. (1992). "Civic norms, social sanctions, and voter turnout," *Rationality and Society*, 4(2), 133-156.

[45] Lazarsfeld, P., Berleson, B. and Gaudet, H. (1944). *The People's Choice: How The Voter Makes Up His Mind in a Presidential Campaign*, New York: Duell, Sloan and Pearce.

[46] Miguel, E., Satyanath, S. and Sergenti, E. (2004). "Economic shocks and civil conflict: An instrumental variables approach," *Journal of Political Economy*, 112(4), 725-753.

[47] Miller, W. E. (1991). "Party identification, realignment, and party voting: Back to the basics," *The American Political Science Review*, 85(2), 557-568.

[48] Montgomery, M. R. and Casterline, J. B. (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.

[49] MOZAMBIQUE Food Security Outlook, July to December 2009. Retrieved February 2011 from http://pdf.usaid.gov/pdf_docs/PNADS377.pdf.

[50] Murphy, K. M. and Topel, R. H. (1985). "Estimation and inference in two-step econometric models," *Journal of Business & Economics Statistics*, 3(4), 370-379.

[51] Nickerson, D. W. (2008). "Is voting contagious? Evidence from two field experiments," *American Political Science Review*, 102(1), 49-57.

[52] Nunn, N. and Wantchekon, L. (2011). "The slave trade and the origins of mistrust in Africa," American Economic Review, 101(7), 3221-3252.

[53] Panagopoulos, C. (2010). "Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior," *Political Behavior*, 32(3), 369-386.

[54] Patchen, M. (1974). "The locus and basis of influence on organizational decisions," *Organizational Behavior and Human Performance*, 11(2), 195-221.

[55] Riker, W. and Ordeshook, P. (1968). "A Theory of the Calculus of Voting," *American Political Science Review,* 62(1), 25–43.

[56] Ryan, R., and Deci, E. (2000). "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being," *American Psychologist*, 55(1), 68-78.

[57] Sheldon, K. (1999). "Machambas in the City: Urban Women and Agricultural Work in Mozambique," *Lusotopie*, pp. 121-140. Available at: http://www.uwcc.wisc.edu/info/women/sheldon.pdf

[58] Stokes, D. E., Campbell, A. and Miller, W. E. (1958). "Components of electoral decision," *The American Political Science Review*, 52(2), 367-387.

[59] Temin, J. and Smith, D. A. (2002). "Media matters: Evaluating the role of the media in Ghana's 2000 Elections," *African Affairs*, 101, 585-605.

[60] Tvedten, I., Paulo, M. and Montserrat, G. (2008). "Gender policies and feminisation of poverty in Mozambique," Report R:2008:13. Bergen, Norway: Chr. Michelsen Institute.

[61] Vicente, P. C. and Wantchekon, L. (2009). "Clientelism and vote buying: lessons from field experiments in African elections," *Oxford Review of Economic Policy*, 25(2), 292–305.

[62] Watts, D. J. and Dodds, P. S. (2007). "Influentials, networks, and public opinion formation," *Journal of Consumer Research*, 34(4), 441–458.

[63] Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

[64] Wooldridge, J. (2003). "Further results on instrumental variables estimation of average treatment effects in the correlated random coefficient model," *Economics Letters*, 79(2), 185-191.

[65] Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

Chapter 3

Voting and peer effects: Experimental evidence from Mozambique

Jointly with Marcel Fafchamps and Pedro Vicente

Voter education campaigns often aim to increase voter participation and political accountability. We follow randomized interventions implemented nationwide during the 2009 Mozambican elections using a free newspaper, leaflets, and text messaging. We investigate whether treatment effects were transmitted through social networks (kinship and chatting) and geographical proximity. For individuals personally targeted by the campaign, we estimate the reinforcement effect of proximity to other targeted individuals. For untargeted individuals, we estimate the diffusion of the campaign depending on proximity to targeted individuals. We find evidence for both effects, similar across the different treatments and across the different connectedness measures. We observe that the treatments worked through networks by raising the levels of information and interest about the election, in line with the average treatment effects. However, differently from those average effects, we find negative

network effects of voter education on voter participation. We interpret this result as a free-riding effect, likely to occur for costly actions.

3.1 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., *Fedderson, 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) make voting a legal obligation, most do not. The level of electoral participation therefore depends on the probability voters attribute to being pivotal and 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 vote-buying (*Vicente, 2007*), violence (*Collier and Vicente, 2009*), 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.

We study a voter education intervention aimed at increasing electoral participation. The voter education was implemented in collaboration with a free newspaper and a consortium of local NGOs. Three different treatments were administered nationwide across four provinces of the country. The first was the distribution of the free newspaper that focused on neutral information about the elections. The second was a text messaging hotline to which citizens could report electoral problems. The third was civic education delivered via a leaflet and text messages providing information about the elections. All treatments included an appeal to voter

participation in the elections. Treatments were allocated randomly across locations. Within locations a number of randomly selected individuals was 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 were not directly targeted by the campaign. We refer to them as 'untargeted'.

Our focus is on estimating the peer effects of the different treatments. Following *Fafchamps and Vicente (2013)*, we divide peer effects into diffusion and reinforcement effects. Diffusion effects refer to the effect of the campaign on untargeted individuals in treated locations who are geographically or socially proximate to targeted individuals. Reinforcement effects refer to whether the effect of treatment is stronger for targeted individuals who are close to other targeted individuals.

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 estimate average treatment effects using official voting records at the ballot station level. To estimate peer 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 increased voter turnout, among targeted and untargeted individuals. Based on ballot station records, both the newspaper and the civic education treatments achieved significant increases in voter turnout. Based on individual data, the hotline was particularly effective. We also document a clear increase in information about the election.

Peer effects on voter participation are, however, quite different as they are all negative. This holds for both reinforcement and diffusion effects, and for different measures of network centrality. This is true for both voter turnout and for our behavioral measure of political participation. Negative peer effects on turnout are particularly strong for the hotline. In contrast, 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 by the probability of affecting the electoral process and by non-instrumental motivations like civic-mindedness. By giving information about

the credibility of the elections, the campaign reassured voters about the integrity of the process. It may also have raised civic-mindedness. Both effects should be conducive to increased turnout, in line with the average effects that we find. However, peer effects can induce free-riding in turnout as more central voters realize that, because of the campaign, turnout will increase and their votes become less essential.

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, *Macours and Vakis (2008)* introduce explicit interaction among households while focusing on reinforcement effects only, and *Angelucci, De Giorgi, Rangel, and Rasul (2010)* extend the analysis to diffusion 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.

The studies by *Nickerson (2008)* 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 randomized get-out-the-vote campaigning in the U.S. to identify peer-effects in two-member households. The second assesses the impact of a voter awareness campaign on female turnout in Pakistan in which peer effects are estimated using geographical distance and friendship. This study uses a testing strategy very similar to *Fafchamps and Vicente (2013)*, which follows a campaign against political violence in Nigeria.[1] This experimental literature on voter mobilization was initiated by *Gerber and Green (2000)* for the US, who studied the impact of a leaflet

[1] The present study, however, focuses in a different context; considers a different type of intervention (directly aimed at increasing turnout); and it was designed *a priori* to measure the diffusion effects (the untargeted respondents were randomly selected during the baseline).

get-out-the-vote campaign.[2] *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. However, none of these studies investigates the peer effects triggered by mobilization.

The paper is organized as follows. In Section 3.2 we describe our testing strategy. We then offer the context of our experiment in Section 3.3. The treatments are introduced in detail in Section 3.4. Subsequently, in Section 3.5 we describe the data including outcome and network variables. In Section 3.6 we present our empirical results, including balance tests, average effects, network effects, and robustness. In Section 3.7 we provide interpretation, including conceptual framework and discussion of alternative explanations for our results. Section 3.8 concludes.

3.2 Testing strategy

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

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

$$y_i = \alpha + \beta T_i + \varepsilon_i \tag{3.1}$$

Coefficient β is the average treatment effect on electoral behavior, information, or interest. This regression can also be estimated with village and individual controls. Whenever comparable information about y_i is available at different points in time, a difference-in-differences version can also be estimated in which individual fixed effects net out possible individual unobservables. Estimation of (3.1) is covered in detail in *Aker, Collier, and Vicente (2011)*.

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

[2] *Wantchekon (2003)* initiated the study of voter interventions in developing countries through the analysis of campaign platforms in Benin.

using untargeted and control observations in a regression of the form:

$$y_i = \alpha^u + \beta^u T_v + \varepsilon_i$$

where $T_v = 1$ if the village was treated. The individual treatment variable T_i drops out since, by design, it is 0 for untargeted individuals. Coefficient β^u is an estimate of the average diffusion effect of the campaign on the electoral behavior, information, or interest of untargeted individuals.

Evidence of possible diagonal effects can be investigated by comparing estimates of β^u with estimates of β. For instance, assume we find that $\beta = 0$ but that $\beta^u > 0$. This indicates that the campaign affected the outcome for the untargeted individuals ($\beta^u > 0$), but not the outcome for those directly affected by the campaign ($\beta = 0$). Such evidence would suggest that treatment effects only operate indirectly.

A possible configuration of interest is when $\beta^u = 0$ for electoral information and interest but $\beta^u > 0$ for electoral behavior. Chapter 2 reports such effects on dependents of treated individuals and interpret these findings as suggesting that treated individuals – who typically are the head of household or his spouse – put pressure on dependents to vote, without necessarily providing them with information or convincing them that voting is a civic duty. Whether similar pressure is applied across households is unclear, and something we examine in this paper.

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

$$y_i = \alpha + \beta T_i + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} + \gamma \frac{1}{N} \sum_{j \neq i} g_{ij} T_i + \varepsilon_i \tag{3.2}$$

where N is the number of observations in i's village, and $\frac{1}{N} \sum_{j \neq i} g_{ij}$ is the proportion of neighbors

of i to whom i is connected, i.e., i's degree of connectedness. The coefficient of interest is γ.[3] Network reinforcement effects are tested by examining whether the effect of treatment is larger among individuals with more direct links to targeted individuals.[4] Regression (3.2) is estimated in levels and first-difference, using control and targeted individuals only, i.e., excluding untargeted individuals living in treated areas.

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

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

$$y_i = \alpha^u + \beta^u T_v + \delta^u \frac{1}{N} \sum_{j \neq i} g_{ij} + \gamma^u \frac{1}{N} \sum_{j \neq i} g_{ij} T_v + \varepsilon_i \tag{3.3}$$

Interpretation is similar to the one of heterogeneous reinforcement effects.

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

3.3 Context

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

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

[4] 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.

official aid assistance accounting for 22 percent of GNI in 2008.[5]

Mozambique became independent from Portugal in 1975, after which FRELIMO (Frente de Libertação de Moçambique), the independence movement, led a single-party, socialist regime. Beginning in 1977, Mozambique suffered a devastating civil war, fought between FRELIMO and RENAMO (Resistência Nacional Moçambicana). RENAMO was supported by Apartheid South Africa and, in the context of the cold war, by the USA. 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 assemblies' elections of October 28, 2009. The 2009 elections were relatively calm, with FRELIMO and Guebuza expected to win. The elections were relatively unproblematic and national and international observers generally considered 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.

3.4 Treatments

The data used in this paper come from a randomized control trial implemented in Mozambique around the time of the 2009 presidential, parliamentary, and provincial assemblies' elections.

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

Three treatments are investigated, all geared towards encouraging people to vote. The first treatment is the distribution of an independent newspaper providing electoral information; the second is a campaign to encourage voters to use an SMS-based hotline set up to report electoral problems; and the third is a civic education campaign, which gave information about the election and focused on participation in the election. The three interventions were designed and conducted with the institutional support and active collaboration of newspaper @Verdade (http://www.verdade.co.mz/) and a consortium of eight Mozambican NGOs, named Observatorio Eleitoral. For more details on these organizations, see *Aker, Collier, and Vicente (2013).*

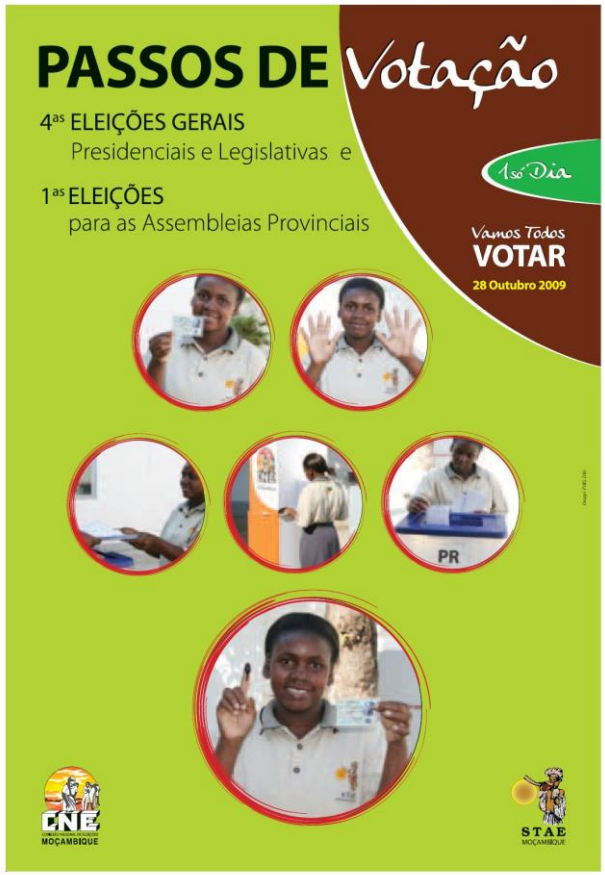
Voter education campaigns generally combine one or more of three elements: (i) information – providing information about the electoral process; (ii) nudging – repeatedly reminding people to vote;[6] and (iii) participation – offering voters the opportunity to circulate their observations about the electoral process.

The newspaper treatment combines information and participation elements. This treatment was organized around the distribution of a free newspaper, *@Verdade*, to experimental subjects in selected locations. None of the treated locations had received the newspaper before.[7] 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 covered the contents of the civic education treatment by including a version of the CNE/STAE leaflet on the steps for voting (see middle panel of *Figure 3.1*). The newspaper also advertised a national hotline for reporting electoral problems, but it was branded with a different slogan and different short-codes from the second treatment (see right panel of *Figure 3.1*). 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 at each location.

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

[7] 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.

Figure 3.1: Newspaper @Verdade (front page – edition before election; civic education page; hotline page)



The hotline treatment emphasizes the participation element. This treatment was organized around the setting-up of two short-code phone numbers contracted with the cell phone operators in Mozambique (Mcel and Vodacom). These short-codes constituted an SMS hotline as they were prepared to receive text messages reporting electoral problems. During the baseline survey, we conducted a door-to-door campaign providing information on the hotline: we distributed 10,000 leaflets (250 per location) primarily directed at targeted respondents, providing basic information about the hotline, i.e., short-codes, examples, format of the reports to be sent,[8] and the name of the sponsors. The leaflet is depicted in *Figure 3.2*. 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 a leaflet designed and made available by the electoral commission (CNE/STAE). A copy of the leaflet is displayed in *Figure 3.3*. 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.

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

Figure 3.2: Hotline leaflet

Figure 3.3: Civic education leaflet by CNE/STAE

3.5 Data

The project took place in four provinces, Cabo Delgado, Zambezia, Gaza, and Maputo-Province. The sampling base was the 2004 electoral map of the country (the 2009 map became available only one month before the election), and the unit of enumeration is the area covered by the polling station. Because the use of cell phones was central to all our treatments, we eliminated 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 were 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, based on information provided by the CNE/STAE in their electronic publication of (disaggregated) electoral data for the 2004 elections. Since all registered voters in the sampling frame have the same probability of being sampled, the 161 locations are nationally representative of the voting population of Mozambique that has access to mobile phone coverage. Of the 161 polling locations selected for our study, 40 were randomly assigned to each of the three treatments, and 41 locations serve as the control group, with no treatment administered. The allocation of locations to treatments and control followed a stratified randomization procedure (*Bruhn and McKenzie, 2009*). First, clusters of four similar locations were formed in each province, with similarity based on geography. Within each cluster, locations were then randomly assigned to one of the three treatments or to control. During the baseline survey, in the event that we found no cell phone coverage in a selected location, we replaced it by the closest polling location with cell phone coverage. That happened in seven locations.

In each of the EAs we conducted two face-to-face household surveys, one before the election and treatment, and one after. Sampling in each EA followed standard procedures for household representativeness (e.g., n'th house call by enumerators starting from the center of the EA, the polling location, typically a school). Interviews targeted the household head or his/her

spouse. Interview and subsequent treatment were 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 included 1,766 households/respondents, approximately 11 per EA. It took place from mid-September to mid-October 2009. The post-election survey started after the election results were announced in early November. It lasted for a similar period of time and sought the same respondents, reaching 1,154 of them. To check for selective attrition, in the next section, we verify whether observable characteristics vary systematically across treatments for the post-election sample. We also run our main results using a multiple imputation technique to account for missing observations.

Treatment was also randomized across respondents/households within each treated EA. Of the 11 households interviewed at baseline, two were, on average, randomly selected not to receive the treatment themselves. The other nine were directly targeted for treatment as described in the previous section. This was done specifically to study diffusion effects on individuals in treated locations not directly reached by the treatment.

3.5.1 Outcome and network variables

The outcomes of interest in this paper come from survey and behavioral data collected at the individual level, and from official voting results at the level of the ballot station. *Table 3.1* presents a summary of the survey outcome variables. These variables have been grouped into three sets: participation, information, and interest, in line with our conceptual framework.

Table 3.1: Survey outcomes

Measure	Indicators	Description / Phrasing of the question	Range / Scale	
Turnout	Self-reported turnout	Which of the following sentences best describes your situation during the 2009 Elections: (1) I was not a registered voter and I was not interested in voting; (2) I was not a registered voter but I would have liked to have voted; (3) I was a registered voter but I chose not to vote; (4) I was a registered voter but I was not able to vote; and (5) I was a registered voter and I voted.	0 - 1	Missing if (1) or (2); 0 if (3) or (4); 1 if (5)
	Adjusted turnout	Self-reported turnout adjusted by considering as non-voters those who did not answer correctly the basic questions about the elections day, namely about the number of ballot papers and ballot boxes.	0 - 1	
	Based on ink finger	What finger was inked after voting?	0 - 1	1 if showed inked finger without hesitation
	Index 1	Composite index measuring how well the respondent answered questions on circumstances and events during the election day	0 - 7	
	Index 2	Composite index based on the sub-group of these questions that focus on knowledge about the ballot station	0 - 7	
	Interviewer's assessment	Enumerator assessment of likelihood that respondent voted	0 - 7	
Index of basic electoral information	Knows candidates' names	Do you know who were the candidates to President in the October 2009 elections?	0 - 3	
	Knows parties	Are you able to list five parties that run for the Parliament in the October 2009 elections?	0 - 3	
Index of interest in elections	Interest in presidential election	With how much interest did you follow the 2009 presidential election?	1 - 4	None-Very much
	Interest in parliamentary election	With how much interest did you follow the 2009 parliamentary election?	1 - 4	None-Very much
	Interest in provincial assemblies' election	With how much interest did you follow the 2009 provincial assemblies' election?	1 - 4	None-Very much
	Interest in public matters	With how much interest do you follow public matters?	1 - 4	None-Very much

Our proxies of information and interest come, respectively, from questions asking respondents to list presidential candidates and parties running for the 2009 elections, and from questions asking about the interest respondents had in the presidential election, parliamentary election, provincial assemblies' election, and generally in public matters. The latter questions employed a subjective scale. To facilitate analysis and interpretation, we combine the questions described above into two indices, one for basic information about the elections and the other for political interest. The indices are constructed following the approach of *Kling, Liebman and Katz (2007)*. We normalize the survey-indicators using z-scores and 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.[9]

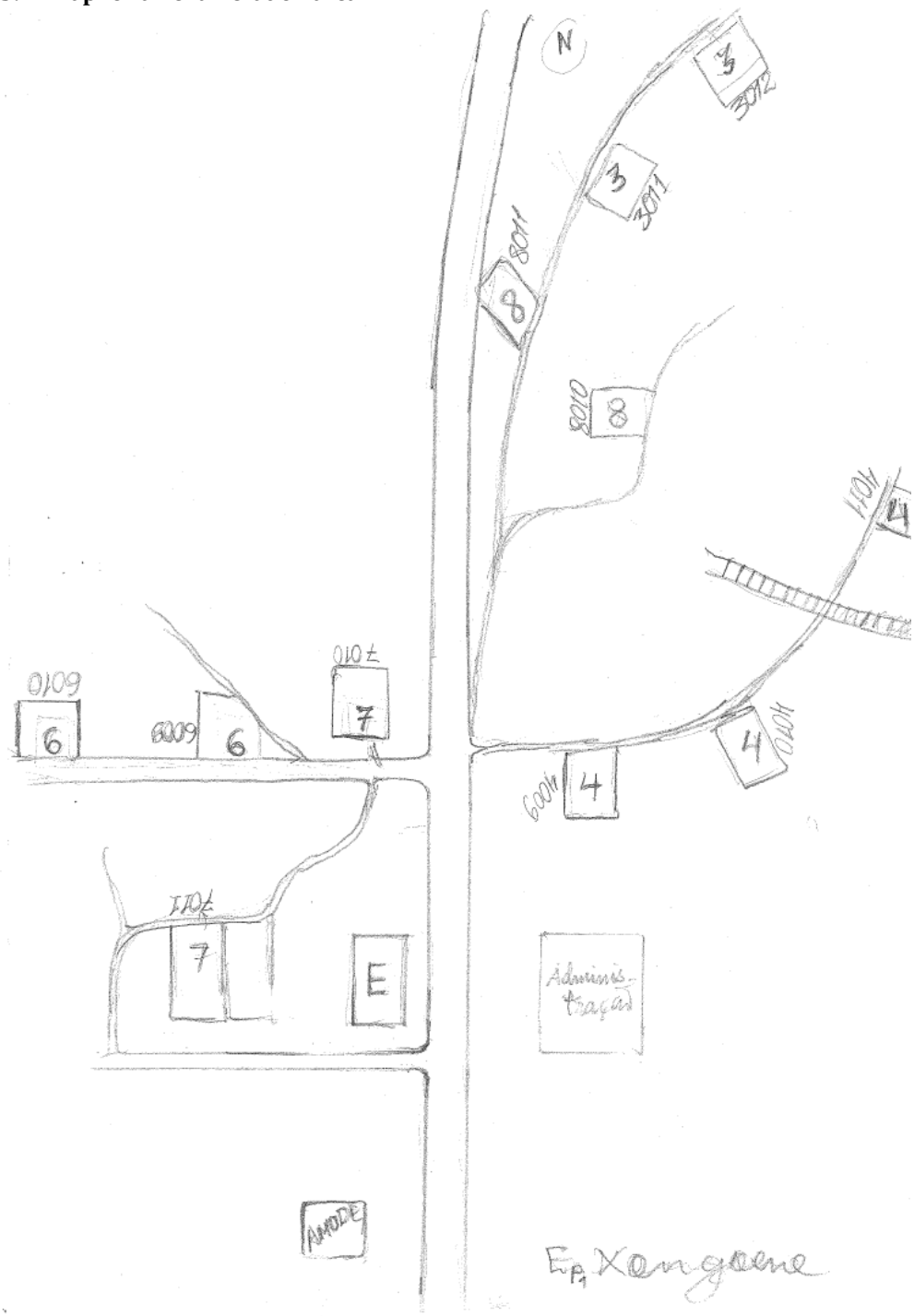
We were particularly careful with our measurement of voter turnout. We propose six turnout measures. The first one is self-reported turnout. The second is self-reported turnout adjusted by considering as non-voters those who did not answer correctly questions regarding ballot papers and boxes.[10] The third one is an indicator of whether the respondent showed without hesitation his/her inked finger to the enumerator – dipping one finger in indelible ink was the method used to prevent people from voting multiple times. Turnout index 1 is a composite index measuring how well the respondent answered questions on the sequence of events during the election day. The answer to each question is coded according to how convincing the response is. Turnout index 2 is based on the sub-group of these questions that focus on knowledge about the ballot station (e.g., the number of ballot papers, whether there were photos of the candidates, the number of ballot boxes, whether they were transparent, and whether they were colored). The last measure of turnout is an 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] 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.

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

Figure 3.4: Open Letter Leaflet

Figure 3.5: A map for an enumeration area



A behavioral measure of demand for political accountability, which we refer to as the 'open letter', is constructed 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 clear 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 3.4*.

Official voting results at the level of the ballot station were made available by the electoral commission of Mozambique. We employ results for the presidential and parliamentary elections of 2009. Ballot stations were matched with the enumeration areas in our experiment, which as mentioned were based on ballot stations themselves.

We collected three measures of social and geographical centrality. The first two variables are centrality measures based on social networks.[11] For the first one, which we call 'chatting', a link from i to j exists if i can identify the name of j when prompted, and i stated that he/she talks to j on a regular basis.[12] For the second, which we call 'kinship', a link from i to j exists if i can identify j by name and reports being related to j.[13] The third variable is a measure of geographical centrality calculated as the average distance to other sampled individuals in the same EA. Each enumerator was asked to locate each respondent on an approximate EA map, and to calculate the distance between interviews. See *Figure 3.5* 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

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

[12] The question asked was 'How frequently do you calmly chat about the day events with the following individuals or members of their households? Not at all-Frequently'.

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

maps differ in scale, distances are re-scaled to make them comparable across all locations.[14]

3.6 Empirical results

3.6.1 Balance

Table 3.2 presents 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, only three demographic characteristics are significantly different at the 10% level. In the sample of untargeted individuals the number of significant differences is reduced to two. The comparison between control and treated EAs in the follow-up survey is also presented in *Table 3.2*. We see a similar pattern: in both samples of targeted and untargeted respondents, most household demographics and EA characteristics are not significantly different. Panel attrition seems to have maintained comparability of the different experimental groups, as far as observables are concerned.

Social and geographical centrality variables are summarized in *Table 3.3*. 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 display average connectedness $\frac{1}{N}\sum_{j\neq i} g_{ij}$, as defined above, with the exception of geographical distance, which is the symmetric of connectedness. We do not observe any statistically significant differences across comparison groups.[16]

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

[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.

[16] If we pool the three interventions, we still find no statistically significant difference between control, targeted and untargeted groups.

Table 3.2: Differences across treatment and control areas - Demographics

		Post-Elections Sample						
			Targeted in treated locations			Untargeted in treated locations		
		Control	Newspaper	Hotline	Civic Education	Newspaper	Hotline	Civic Education
Basic demographics	female	0.563	-0.031	-0.024	0.060	-0.071	-0.016	0.067
			(0.044)	(0.044)	(0.041)	(0.061)	(0.057)	(0.059)
	age	36.957	-0.811	1.713	0.280	1.587	2.411	-1.145
			(1.394)	(1.461)	(1.329)	(2.077)	(2.027)	(1.945)
	single	0.186	-0.042	0.006	-0.039	0.024	0.001	0.036
			(0.037)	(0.041)	(0.036)	(0.065)	(0.055)	(0.080)
	married or in a union	0.692	0.092**	0.039	0.050	0.010	-0.004	-0.044
			(0.040)	(0.047)	(0.044)	(0.073)	(0.061)	(0.085)
	literate	0.821	0.034	-0.010	-0.038	-0.014	0.007	-0.062
			(0.034)	(0.044)	(0.042)	(0.056)	(0.060)	(0.065)
	no school	0.179	-0.034	0.010	0.038	0.014	-0.007	0.062
			(0.034)	(0.044)	(0.042)	(0.056)	(0.060)	(0.065)
	incomplete primary school	0.276	0.063	-0.053	0.031	-0.013	0.115*	-0.091
			(0.047)	(0.041)	(0.043)	(0.061)	(0.067)	(0.061)
	incomplete secondary school	0.168	-0.019	0.003	-0.013	-0.046	-0.012	-0.002
			(0.038)	(0.037)	(0.037)	(0.054)	(0.051)	(0.056)
	household head	0.742	0.037	-0.024	-0.017	0.065	0.055	-0.020
			(0.049)	(0.044)	(0.044)	(0.063)	(0.059)	(0.077)
	household size	5.789	0.227	0.483*	0.068	-0.060	0.094	0.665
			(0.282)	(0.261)	(0.286)	(0.367)	(0.439)	(0.467)
Ethnicity	changana	0.355	0.012	0.013	0.018	0.031	0.005	-0.034
			(0.095)	(0.092)	(0.092)	(0.107)	(0.106)	(0.099)
	macua	0.244	-0.054	-0.013	-0.035	0.019	0.006	-0.055
			(0.081)	(0.085)	(0.085)	(0.096)	(0.098)	(0.093)
	lomue	0.118	0.008	-0.046	-0.016	-0.031	-0.056	0.033
			(0.072)	(0.058)	(0.066)	(0.068)	(0.061)	(0.075)
	chuabo	0.100	-0.014	-0.023	0.010	0.005	-0.007	-0.025
			(0.058)	(0.051)	(0.060)	(0.062)	(0.062)	(0.053)
	chironga	0.061	-0.011	-0.027	-0.028	-0.043*	0.002	0.015
			(0.031)	(0.022)	(0.024)	(0.026)	(0.042)	(0.042)
	maconde	0.018	0.023	0.021	0.023	-0.000	0.013	0.058
			(0.020)	(0.032)	(0.026)	(0.021)	(0.033)	(0.052)

Note: Standard errors reported; these are corrected by clustering at the location (enumeration area) level.

* significant at 10%; ** significant at 5%; *** significant at 1%. These results come from OLS regressions.

Table 3.2: Differences across treatment and control areas - Demographics (cont.)

		Post-Elections Sample						
			Targeted in treated locations			Untargeted in treated locations		
		Control	Newspaper	Hotline	Civic Education	Newspaper	Hotline	Civic Education
Religion	catholic	0.401	-0.077	-0.060	-0.024	-0.051	-0.136*	0.043
			(0.060)	(0.062)	(0.056)	(0.081)	(0.077)	(0.093)
	protestant	0.319	0.050	0.036	0.029	-0.021	0.072	0.014
			(0.067)	(0.071)	(0.063)	(0.072)	(0.086)	(0.088)
	muslim	0.215	0.037	0.020	-0.018	0.031	0.066	-0.011
			(0.073)	(0.071)	(0.064)	(0.083)	(0.088)	(0.079)
Occupation	has a job	0.240	-0.033	0.029	-0.052	0.058	-0.053	0.038
			(0.041)	(0.049)	(0.042)	(0.067)	(0.059)	(0.068)
	agriculture	0.353	-0.064	-0.028	-0.000	0.058	0.022	0.055
			(0.066)	(0.070)	(0.065)	(0.090)	(0.091)	(0.089)
	retail informal sector	0.043	0.002	0.008	0.006	-0.025	0.004	-0.025
			(0.022)	(0.019)	(0.019)	(0.022)	(0.030)	(0.023)
	artisan	0.050	0.017	0.001	-0.030*	-0.015	-0.019	-0.050***
			(0.022)	(0.019)	(0.016)	(0.028)	(0.025)	(0.013)
	unskilled worker	0.054	0.018	0.006	-0.013	-0.018	-0.007	-0.017
			(0.023)	(0.021)	(0.020)	(0.029)	(0.029)	(0.030)
	teacher	0.047	-0.002	0.013	-0.022	0.025	0.000	0.064
			(0.019)	(0.027)	(0.015)	(0.037)	(0.029)	(0.042)
	student	0.040	0.001	-0.014	0.018	-0.022	0.007	-0.040***
			(0.018)	(0.019)	(0.022)	(0.022)	(0.030)	(0.013)
	domestic	0.147	-0.017	-0.032	-0.008	-0.022	-0.038	0.019
			(0.032)	(0.032)	(0.031)	(0.047)	(0.048)	(0.059)
Assets and expenditure	house	0.853	-0.015	0.027	0.008	-0.028	-0.025	0.036
			(0.034)	(0.031)	(0.033)	(0.066)	(0.056)	(0.054)
	land	0.652	-0.044	-0.028	-0.062	-0.056	-0.105	-0.023
			(0.059)	(0.060)	(0.059)	(0.090)	(0.084)	(0.088)
	cattle	0.254	0.016	-0.011	-0.000	-0.044	0.074	-0.032
			(0.052)	(0.052)	(0.047)	(0.073)	(0.077)	(0.072)
	cell phone	0.706	0.006	0.046	0.007	0.083	0.091	0.090
			(0.065)	(0.068)	(0.068)	(0.080)	(0.083)	(0.081)
	expenditure	122.452	4.816	3.641	-15.275	1.572	28.799	77.759
			(17.013)	(15.622)	(14.056)	(22.193)	(35.099)	(60.203)

Note: Standard errors reported; these are corrected by clustering at the location (enumeration area) level.

* significant at 10%; ** significant at 5%; *** significant at 1%. These results come from OLS regressions.

Finally, we display averages for our baseline voting variables at the ballot station level. These are voting records from the presidential and parliamentary elections of 2004. Results are presented in *Table 3.4*. We do not observe any statistically significant differences across comparison groups. Voting variables from our baseline survey are explored in full detail in *Aker, Collier, and Vicente (2013)*: since treated respondents were asked questions on politics after receiving leaflets (for civic education and the hotline) and the newspaper, there may be differences between comparison groups for the targeted due to first reactions or conformity bias. However, no clear evidence of such effects is apparent in the data.

3.6.2 Average effects

We start by presenting regression results of the average effects of the campaign. These results are explored in full detail in *Aker, Collier, and Vicente (2013)*. We start with measures of political participation, which is the main intended effect of the campaign. *Table 3.5a* presents the average effects of the three treatments on voter turnout and the sending of the open letter among targeted individuals. Since this information is only available in the follow-up survey, all regressions are estimated using post-elections data only. For each measure we present two sets of estimates: the first has no controls; the second including provincial dummies, EA controls and individual characteristics.[17]

We first look at reported values of outcome variables for control individuals. These are given by the intercept terms for all regressions with no control variables. Self-reported turnout is 86 percent. Since this 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, i.e., self-reported turnout lower than more objectively measured turnout. This being said, the lowest turnout measure still puts average turnout above 70 percent among respondents. The inked finger measure, which is arguably the most objective although not necessarily the most accurate, suggests a turnout of 81 percent on average among control individuals. Since our sample only contains household heads and their spouses, it is perhaps not surprising that

[17] We employ only unbalanced individual characteristics. Similar results are obtained if we include a longer list of controls but, following *Deaton's (2009)* advice, we report results with a more parsimonious list of controls.

average turnout in the sample is above the 2009 national average of 44 percent. 15 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 12 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 all cases when we use no controls. By the inked finger measure and when we use no controls, the civic education treatment caused a 5.6 percentage point increase in turnout, which compares well to the 6.3 percentage point associated with the hotline treatment. However, when we use controls the civic education treatment is significant only with the adjusted turnout and interviewer assessment measures. For the newspaper treatment we find a small positive effect when we use no controls (i.e., +3.1 percentage points on average across the six measures), but this effect is never statistically significant. When we use controls the effect diminishes (i.e., +1.4 percentage points on average). Regarding the open letter, we find a significant positive effect of the newspaper treatment. On average, this treatment increases the probability of sending a SMS proposing policy priorities in 10.9 percentage points (when using controls).

The average diffusion effect of the campaign on untargeted individuals is reported in *Table 3.5b*. We find statistically significant diffusion effects for two treatments, newspaper and hotline.[18] If anything, point estimates are larger for untargeted than targeted respondents: averaged over the six measures (no controls), we find 10.5 and 9.2 percentage point increases in turnout among untargeted households for the newspaper and hotline treatments, respectively. The treatment effect is also large for the civic education treatment, but it is only statistically significant when using the adjusted turnout and no controls. The result on the newspaper is surprising given that the newspaper treatment has no statistically significant effect on turnout among the targeted. This suggests that the diffusion effect of the newspaper treatment is stronger than its direct effect, possibly because of magnification of the importance of the newspaper among individuals without direct access to it. In contrast, the civic education treatment increases turnout

[18] The newspaper effect is not statistically significant when using adjusted turnout, and the hotline treatment is not significant when using the finger measure and including controls.

among the targeted but not among the untargeted, suggesting that it has a smaller and possibly negligible diffusion effect. We find positive average diffusion effects on the open letter for all treatments. However, none of these effects is significant.[19]

We now look at the effects of the treatments on information and interest in politics. These are displayed in *Table 3.6*. In our conceptual framework, we hypothesized several likely mechanisms of change produced by the voter education campaign: first, distributing information about the electoral process may convince voters of the integrity of the process, thereby raising $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right]$; second, the campaigns may increase non-instrumental motivation $\partial U/\partial x_i$ by raising civic-mindedness. We find significant positive effects of the hotline on information about the elections. This effect ranges between 16.5 and 23.3 standard deviation units for the targeted; for the untargeted it is equal to 22.6 standard deviation units (in the regression with controls). The newspaper also has a positive impact on information, which is significant for the targeted (19.2 standard deviation units). Neither the hotline nor the newspaper had a clear impact on interest in politics. The civic education treatment does not have any significant effects on any of these outcomes, even though the sign of the effect on the information index is consistently positive.

[19] We have run comparable regressions using individual fixed effects (available upon request). The comparison is between intended turnout at baseline and the three dichotomous turnout measures (self-reported, adjusted, and inked finger). We find similar results regarding the hotline treatment among targeted and untargeted respondents: the effect on turnout is positive and statistically significant in all regressions. For the average diffusion effect, the newspaper treatment remains significant in two of the three regressions, and civic education is significant in one (at the 10 percent level only). Virtually identical results are obtained if we omit individual fixed effects but include individual controls.

Table 3.5a: Average effect of the three treatments on targeted respondents' political participation

	Turnout												Open letter	
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Newspaper	0.024	0.012	0.051	0.015	0.011	0.005	0.028	0.013	0.034	0.016	0.035	0.024	0.087*	0.109**
	(0.037)	(0.038)	(0.048)	(0.042)	(0.042)	(0.042)	(0.035)	(0.035)	(0.035)	(0.034)	(0.039)	(0.038)	(0.050)	(0.053)
Hotline	0.080***	0.070**	0.120***	0.094**	0.063**	0.053*	0.075***	0.063**	0.080***	0.065**	0.094***	0.085***	-0.037	-0.002
	(0.026)	(0.029)	(0.044)	(0.041)	(0.028)	(0.030)	(0.025)	(0.027)	(0.027)	(0.027)	(0.026)	(0.028)	(0.035)	(0.038)
Civic education	0.045	0.032	0.085*	0.073*	0.056*	0.043	0.043	0.031	0.052*	0.039	0.073***	0.069**	0.044	0.061
	(0.028)	(0.030)	(0.044)	(0.042)	(0.029)	(0.033)	(0.026)	(0.027)	(0.028)	(0.028)	(0.027)	(0.029)	(0.048)	(0.055)
constant	0.859***	0.887***	0.706***	0.824***	0.807***	0.783***	0.788***	0.841***	0.757***	0.835***	0.753***	0.728***	0.153***	0.140*
	(0.021)	(0.060)	(0.031)	(0.074)	(0.020)	(0.070)	(0.019)	(0.056)	(0.020)	(0.056)	(0.020)	(0.063)	(0.026)	(0.083)
Adjusted R-squared	0.006	0.023	0.008	0.061	0.002	0.016	0.006	0.026	0.007	0.041	0.010	0.024	0.011	0.013
No. of observations	953	845	953	845	953	845	953	845	953	845	953	845	973	817
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 3.5b: Average effect of the three treatments on untargeted respondents' political participation

	Turnout												Open letter	
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Newspaper	0.088**	0.089**	0.079	0.047	0.140***	0.149***	0.088**	0.087**	0.094***	0.091**	0.142***	0.138***	0.058	0.072
	(0.037)	(0.040)	(0.064)	(0.058)	(0.037)	(0.040)	(0.034)	(0.036)	(0.036)	(0.036)	(0.038)	(0.041)	(0.066)	(0.066)
Hotline	0.090**	0.080*	0.158***	0.092*	0.092**	0.069	0.087***	0.074*	0.095***	0.068*	0.099**	0.081*	0.006	0.036
	(0.035)	(0.043)	(0.054)	(0.056)	(0.043)	(0.050)	(0.034)	(0.041)	(0.035)	(0.040)	(0.040)	(0.042)	(0.049)	(0.052)
Civic education	0.047	0.022	0.105*	0.071	0.061	0.052	0.053	0.033	0.065	0.045	0.042	0.023	0.107	0.109
	(0.045)	(0.050)	(0.061)	(0.058)	(0.049)	(0.055)	(0.040)	(0.045)	(0.040)	(0.044)	(0.046)	(0.048)	(0.070)	(0.075)
constant	0.859***	0.838***	0.706***	0.773***	0.807***	0.668***	0.788***	0.781***	0.757***	0.765***	0.753***	0.628***	0.153***	0.167*
	(0.021)	(0.095)	(0.031)	(0.114)	(0.020)	(0.095)	(0.019)	(-0.085)	(0.020)	(0.085)	(0.020)	(0.084)	(0.027)	(0.091)
Adjusted R-squared	0.008	-0.008	0.012	0.056	0.014	0.004	0.011	-0.002	0.014	0.017	0.019	0.023	0.003	0.011
No. of observations	437	379	437	379	437	379	437	379	437	379	437	379	449	386
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 3.6: Average effect of the three treatments on respondents' information and interest

	Homogeneous effects on targeted respondents				Homogeneous effects on untargeted respondents			
	Basic electoral information		Interest in elections		Basic electoral information		Interest in elections	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Newspaper	0.122	0.192**	0.008	-0.059	0.124	0.182	0.015	-0.038
	(0.095)	(0.085)	(0.104)	(0.084)	(0.135)	(0.115)	(0.181)	(0.147)
Hotline	0.165*	0.233***	0.103	0.090	0.181	0.226*	0.102	0.063
	(0.089)	(0.083)	(0.106)	(0.078)	(0.132)	(0.122)	(0.131)	(0.102)
Civic education	0.037	0.046	0.078	-0.001	0.152	0.179	-0.003	-0.080
	(0.097)	(0.101)	(0.094)	(0.063)	(0.119)	(0.128)	(0.154)	(0.133)
constant	0.000	-0.226	0.000	0.026	0.000	0.164	0.000	0.202
	(0.066)	(0.145)	(0.075)	(0.118)	(0.066)	(0.188)	(0.075)	(0.183)
Adjusted R-squared	0.003	0.062	-0.000	0.110	0.001	0.179	-0.005	0.220
No. of observations	976	865	976	865	453	395	454	396
Controls	Yes	Yes	Yes	Yes	No	Yes	No	Yes

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

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

Table 3.7: Average effect of the three treatments on the official electoral results (ballot-station level)

	Presidential elections				Parliamentary elections			
	Turnout	% null votes	% votes in Guebuza	% votes in Dhlakama	Turnout	% null votes	% votes FRELIMO	% votes RENAMO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Newspaper	0.055**	-0.007*	0.040*	-0.015	0.057**	-0.002	0.034	-0.020
	(0.025)	(0.004)	(0.020)	(0.016)	(0.025)	(0.003)	(0.021)	(0.016)
Hotline	0.051**	-0.005	0.025	-0.015	0.053**	-0.004	0.023	-0.017
	(0.025)	(0.004)	(0.020)	(0.016)	(0.025)	(0.003)	(0.021)	(0.015)
Civic education	0.053**	-0.003	0.046**	-0.032**	0.052**	-0.006*	0.039*	-0.038**
	(0.025)	(0.004)	(0.020)	(0.016)	(0.025)	(0.003)	(0.021)	(0.015)
constant	0.405***	0.044***	0.634***	0.164***	0.409***	0.036***	0.597***	0.168***
	(0.033)	(0.005)	(0.027)	(0.021)	(0.033)	(0.004)	(0.028)	(0.021)
Adjusted R-squared	0.389	0.171	0.673	0.582	0.377	0.145	0.666	0.637
No. of observations	161	161	161	161	161	161	161	161
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

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

Finally we report in *Table 3.7* the average effect of the treatments on actual electoral outcomes from official ballot station records. We find that all treatments had a strong and significant positive effect on voter turnout. This effect ranged between 5.1 and 5.5 percentage points for the presidential elections, and between 5.2 and 5.7 percentage points for the parliamentary elections. On voting patterns, we find a positive point estimate of all treatments on voting for the incumbent president and party (FRELIMO) and a negative effect of all treatments on voting for the main challenger candidate and party (RENAMO). However, only civic education is statistically significant in every case.[20] This treatment leads to an increase in the score of FRELIMO and the incumbent president by 3.9 and 4.6 percentage points, respectively. It also reduces the share of votes of RENAMO and its presidential candidate by 3.8 and 3.2 percentage points, respectively. We therefore conclude that our treatments increased voter turnout, benefited the incumbent, and hurt the chances of the challenger.[21]

3.6.3 Network effects on political participation

Next we turn to network effects on political participation, i.e., voter turnout and open letter. In *Table 3.8* we show our regressions relating to the inked finger measure of voter turnout. We employ the three centrality variables presented earlier: chatting, kinship, and geographic proximity. We start by measuring network reinforcement effects with regression (3.2). Results are displayed in columns (1)-(3). We then estimate network diffusion effects with regression (3.3). Results are displayed in columns (4)-(6). All regressions are estimated using only data from the follow-up survey. We control for provincial dummies, EA characteristics and individual characteristics. The main focus is on the coefficient of $\frac{1}{N}\sum_{j\neq i} g_{ij}$ and of the interaction terms $\frac{1}{N}\sum_{j\neq i} g_{ij}T_i$ (reinforcement) and $\frac{1}{N}\sum_{j\neq i} g_{ij}T_v$ (diffusion).

We obtain strong positive coefficients for $\frac{1}{N}\sum_{j\neq i} g_{ij}$ particularly when using chatting and kinship as centrality measure, but also when using geographical proximity.[22] This implies that,

[20] The estimate of the effect of the newspaper treatment on voting for the incumbent president is also significant at 10 percent level.

[21] In 2004 elections, RENAMO obtained more votes than FRELIMO in 22 of the 161 enumeration areas. Twenty of those EAs are located in Zambezia. In the 2009 elections there were only 4 EAs where RENAMO obtained more votes than FRELIMO. These EAs are all located in Zambezia.

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

without treatment, individuals who are more central socially in their community are more likely to vote. A likely reason is that these individuals feel more pressure to perform their voting duty. From this data 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 term, we find a negative coefficient for most treatments and centrality measures. Results indicate that reinforcement through geographical proximity is consistently negative and statistically significant for all three treatments: the average respondent in the control group (in terms of geographical distance) sees a decrease in the probability of voting by 12.8, 13, and 11.1 percentage points through network reinforcement for the newspaper, hotline, and civic education, respectively. A negative and statistically significant effect is also present for diffusion through chatting for the hotline treatment. The average respondent in the control group (in terms of chatting) sees a decrease in the probability of voting by 9.4 percentage points through network diffusion for the hotline.

We display full results for turnout measure index 1 in *Table 3.9*. The same pattern of negative network effects emerges. Negative network effects are particularly prominent for the hotline: all proximity variables yield negative reinforcement and diffusion, except for reinforcement through geographical proximity, which is not significant. For the civic education treatment, reinforcement through kinship is also negative and statistically significant. *Table 3.10* shows the interaction coefficients for the remaining voter turnout measures. Significant effects are all negative. The hotline stands out as inducing most network reinforcement and diffusion effects. For self-reported turnout and for interviewer assessed turnout, the newspaper also induces negative network reinforcement and diffusion through geographical proximity, and civic education also induces negative reinforcement effects through kinship and geographical proximity. [23]

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

Table 3.8: Network effects on turnout measure based on inked finger

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.016	0.004	0.005	0.167***	0.163***	0.161***
	(0.042)	(0.043)	(0.050)	(0.040)	(0.040)	(0.044)
Hotline	0.059*	0.052*	0.074**	0.076	0.074	0.070
	(0.033)	(0.031)	(0.035)	(0.052)	(0.052)	(0.056)
Civic education	0.048	0.041	0.055	0.081	0.074	0.022
	(0.035)	(0.034)	(0.041)	(0.052)	(0.051)	(0.064)
Network	0.268**	0.233**	0.082*	0.341***	0.302***	0.093*
	(0.128)	(0.112)	(0.043)	(0.131)	(0.112)	(0.049)
Network x Newspaper	**0.033**	**-0.207**	**-0.105***	**-0.106**	**-0.088**	**-0.033**
	(0.152)	(0.186)	(0.058)	(0.166)	(0.189)	(0.046)
Network x Hotline	**-0.125**	**-0.215**	**-0.107****	**-0.438****	**-0.231**	**-0.103**
	(0.155)	(0.162)	(0.053)	(0.216)	(0.170)	(0.082)
Network x Civic education	**-0.116**	**-0.171**	**-0.091****	**-0.075**	**0.145**	**-0.107**
	(0.173)	(0.154)	(0.043)	(0.162)	(0.231)	(0.071)
constant	0.772***	0.782***	0.783***	0.640***	0.654***	0.610***
	(0.070)	(0.070)	(0.076)	(0.098)	(0.099)	(0.100)
Adjusted R-squared	0.032	0.015	0.017	0.027	0.008	0.015
No. of observations	845	845	721	379	379	324
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

Table 3.9: Network effects on turnout measure index 1

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.020	0.013	-0.004	0.100***	0.098***	0.096**
	(0.034)	(0.035)	(0.040)	(0.036)	(0.036)	(0.039)
Hotline	0.066**	0.062**	0.068**	0.080*	0.081**	0.087**
	(0.028)	(0.027)	(0.029)	(0.043)	(0.041)	(0.041)
Civic education	0.035	0.030	0.037	0.053	0.051	0.016
	(0.028)	(0.028)	(0.033)	(0.045)	(0.043)	(0.050)
Network	0.247**	0.219**	0.064	0.270***	0.253**	0.073*
	(0.098)	(0.099)	(0.041)	(0.101)	(0.102)	(0.044)
Network x Newspaper	**-0.053**	**-0.119**	**-0.060**	**-0.116**	**-0.071**	**-0.051**
	(0.118)	(0.129)	(0.051)	(0.135)	(0.175)	(0.040)
Network x Hotline	**-0.200***	**-0.255****	**-0.069**	**-0.307****	**-0.305***	**-0.110***
	(0.113)	(0.110)	(0.046)	(0.137)	(0.163)	(0.063)
Network x Civic education	**-0.181**	**-0.254***	**-0.061**	**-0.112**	**0.116**	**-0.040**
	(0.136)	(0.131)	(0.041)	(0.118)	(0.198)	(0.060)
constant	0.834***	0.838***	0.839***	0.761***	0.769***	0.752***
	(0.056)	(0.056)	(0.062)	(0.088)	(0.090)	(0.088)
Adjusted R-squared	0.039	0.028	0.031	0.017	0.004	0.015
No. of observations	845	845	721	379	379	324
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

Table 3.10: Estimates of the interaction coefficients in turnout regressions using the remaining turnout proxies

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

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

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

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

Results for the open letter are displayed in *Table 3.11*. 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. The average respondent in the control group (in terms of kinship) sees a decrease in the probability of sending a text message of 4 to 11.4 percentage points through networks for the civic education. The interpretation may be similar to one for voter turnout: sending an open letter is a costly action that potentially suffers from free-riding.

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

Table 3.11: Network effects on behavioral measure open letter

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

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

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

3.6.4 Network effects on information and interest in politics

We now seek to identify the channels through which the treatments affect political participation. We have already noted that the treatments have a direct positive effect on information about the election among targeted and untargeted individuals. But we could not find statistically significant effects for interest in politics. We now examine network diffusion and reinforcement effects on information and interest in politics. We want to know whether information and interest are transmitted across networks, and, if yes, whether the effects are negative like for political participation.

As in the previous section, we employ the three centrality variables presented – chatting, kinship, and geographic proximity – and we estimate network reinforcement effects with regression (3.2) and network diffusion effects with regression (3.3). All regressions are estimated using post-election data only. We control for provincial dummies, EA characteristics and individual characteristics. The main focus is on the coefficient of $\frac{1}{N}\sum_{j\neq i} g_{ij}$ and of the interaction terms $\frac{1}{N}\sum_{j\neq i} g_{ij}T_i$ (reinforcement) and $\frac{1}{N}\sum_{j\neq i} g_{ij}T_v$ (diffusion).

Table 3.12 shows the results for our index of information about the elections. We only find a statistically significant effect: network reinforcement through kinship when analyzing the impact of the newspaper treatment. The estimated coefficient implies that the average respondent in the control group (in terms of kinship) sees an increase in the index of information of 0.07 standard deviation units through network reinforcement for the newspaper treatment. This is a relatively small effect, but unlike for political participation, it is positive.

Table 3.12: Network effects on the index of basic electoral information

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.194**	0.194**	0.192**	0.185	0.159	0.196
	(0.088)	(0.084)	(0.095)	(0.118)	(0.136)	(0.122)
Hotline	0.233***	0.236***	0.231**	0.232**	0.269**	0.232*
	(0.082)	(0.082)	(0.096)	(0.118)	(0.109)	(0.135)
Civic education	0.047	0.046	0.044	0.187	0.181	0.299**
	(0.099)	(0.099)	(0.118)	(0.128)	(0.128)	(0.123)
Network	-0.074	-0.394	0.125	-0.051	-0.357*	0.108
	(0.246)	(0.275)	(0.129)	(0.244)	(0.215)	(0.112)
Network x Newspaper	**0.274**	**0.641***	**-0.206**	**0.091**	**-0.122**	**-0.270**
	(0.368)	(0.376)	(0.145)	(0.566)	(1.171)	(0.181)
Network x Hotline	**-0.126**	**0.073**	**-0.110**	**-0.645**	**-0.698**	**-0.119**
	(0.314)	(0.382)	(0.127)	(0.590)	(0.727)	(0.182)
Network x Civic education	**-0.270**	**-0.163**	**-0.177**	**0.273**	**0.612**	**-0.207**
	(0.360)	(0.353)	(0.141)	(0.481)	(0.658)	(0.132)
constant	-0.227	-0.232	-0.211	0.153	0.183	0.126
	(0.144)	(0.145)	(0.174)	(0.181)	(0.180)	(0.213)
Adjusted R-squared	0.062	0.066	0.074	0.176	0.183	0.182
No. of observations	865	865	741	395	395	340
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

Table 3.13: Network effects on the index of interest in elections

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	-0.034	-0.051	-0.076	-0.012	0.039	-0.056
	(0.079)	(0.079)	(0.091)	(0.141)	(0.142)	(0.148)
Hotline	0.102	0.094	0.118	0.064	0.085	0.076
	(0.071)	(0.073)	(0.087)	(0.098)	(0.095)	(0.116)
Civic education	0.006	0.006	0.005	-0.040	-0.084	0.021
	(0.060)	(0.061)	(0.065)	(0.129)	(0.152)	(0.142)
Network	-0.262	-0.344	0.010	-0.317*	-0.209	-0.054
	(0.178)	(0.270)	(0.064)	(0.167)	(0.315)	(0.087)
Network x Newspaper	**1.158*** **	**1.110*** **	**0.021**	**1.372*** **	**1.925*** **	**0.444** **
	(0.266)	(0.399)	(0.091)	(0.484)	(0.675)	(0.181)
Network x Hotline	**0.839*** **	**0.809** **	**-0.118**	**-0.077**	**-0.063**	**0.077**
	(0.234)	(0.380)	(0.110)	(0.428)	(0.634)	(0.225)
Network x Civic education	**0.445***	**0.135**	**-0.059**	**1.226*** **	**-0.043**	**0.041**
	(0.255)	(0.432)	(0.072)	(0.453)	(1.029)	(0.141)
constant	-0.010	0.004	0.009	0.176	0.147	0.247
	(0.111)	(0.115)	(0.130)	(0.176)	(0.177)	(0.204)
Adjusted R-squared	0.130	0.120	0.091	0.237	0.225	0.229
No. of observations	865	865	741	396	396	341
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

Results for our index of interest in politics are presented in *Table 3.13*. We find a larger number of significant interaction effects, all of which are positive. The most robust network effects are found for the newspaper treatment. Results suggest that both reinforcement and diffusion are channeled through kinship and chatting, and that only diffusion is channeled through geographical proximity. Almost all network effects of the newspaper are significant at the 1 percent level. They imply that the newspaper increases interest in politics for the average respondent in the control group: in terms of chatting, by between 0.25 (reinforcement) and 0.29 (diffusion) standard deviation units; in terms of kinship, by between 0.12 (reinforcement) and 0.21 (diffusion) standard deviation units; in terms of geographical proximity, by 0.54 standard deviation units (diffusion).[24] We find slightly lower network reinforcement effects for the hotline, through chatting and kinship. Slightly lower reinforcement and diffusion network effects are also present for the civic education treatment, but only through chatting.

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

3.6.5 Robustness check: multiple imputation

Although balance tests do not indicate that panel attrition significantly affected the comparability of treatment and control groups, we nevertheless test how sensitive our results are to missing post-elections observations. We use the multiple imputation method to replace the missing values of outcome and control variables; and re-estimate the average and network effects on political participation using the full sample of baseline respondents. We employ multivariate normal regressions.[25] 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

[24] The figures listed here are obtained by multiplying the coefficient of the respective interaction term by the average of the network variable among control respondents (displayed in *Table 3.3*).

[25] 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 often normal imputation models have a good performance for linear regressions, even when the variables are non-normal.

of the enumeration area, and interactions between the interventions and characteristics of the household and respondents.

Recalculated estimates of the average treatment effects on political participation are similar to the ones obtained earlier. In *Table 3.14a* 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 3.5a*. We find a very similar pattern of significant effects. When looking at the untargeted respondents (*Table 3.14b*) the differences to the main results (*Table 3.5b*) are even smaller.

In terms of network effects, the coefficients of the network interaction terms remain negative for most treatments and network measures. Again, we find a very similar pattern of significant results, although with a much smaller magnitude. *Table 3.15* displays the estimates of peer effects on turnout measure index 1 using imputed data. Comparing these results with the ones displayed in *Table 3.9* we see that most of the significant network effects remain so, although they have smaller magnitudes. Overall, the estimates using imputed data corroborate the network effects identified earlier.

Table 3.14a: Average effect of the three treatments on targeted respondents' political participation when imputing missing information

	Turnout												Open letter	
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Newspaper	0.013	0.008	0.028	0.019	0.003	-0.004	0.017	0.012	0.020	0.014	0.020	0.014	0.082*	0.075
	(0.026)	(0.025)	(0.037)	(0.032)	(0.030)	(0.029)	(0.025)	(0.024)	(0.026)	(0.023)	(0.028)	(0.026)	(0.050)	(0.050)
Hotline	0.060***	0.050**	0.090***	0.078***	0.045*	0.030	0.054***	0.045**	0.058***	0.046**	0.067***	0.055***	-0.032	-0.027
	(0.022)	(0.022)	(0.034)	(0.030)	(0.024)	(0.024)	(0.019)	(0.019)	(0.021)	(0.019)	(0.022)	(0.021)	(0.037)	(0.035)
Civic education	0.033	0.025	0.061*	0.061*	0.041	0.035	0.030	0.026	0.036	0.034*	0.051*	0.050*	0.042	0.024
	(0.023)	(0.022)	(0.036)	(0.032)	(0.025)	(0.025)	(0.021)	(0.020)	(0.022)	(0.020)	(0.022)	(0.022)	(0.044)	(0.044)
constant	0.870***	0.872***	0.726***	0.794***	0.819***	0.786***	0.801***	0.823***	0.771***	0.815***	0.770***	0.734***	0.150***	0.192**
	(0.016)	(0.042)	(0.025)	(0.058)	(0.017)	(0.049)	(0.014)	(0.038)	(0.016)	(0.038)	(0.015)	(0.041)	(0.028)	(0.076)
No. of observations	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514	1,514
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 3.14b: Average effect of the three treatments on untargeted respondents' political participation when imputing missing information

	Turnout												Open letter	
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Newspaper	0.063*	0.062*	0.058	0.051	0.103***	0.102***	0.064**	0.067**	0.069**	0.069**	0.105***	0.099***	0.080	0.077
	(0.033)	(0.035)	(0.053)	(0.051)	(0.038)	(0.039)	(0.030)	(0.031)	(0.031)	(0.031)	(0.034)	(0.035)	(0.068)	(0.068)
Hotline	0.069**	0.065*	0.111**	0.103**	0.068*	0.061	0.062**	0.063*	0.067**	0.065**	0.073**	0.065*	-0.002	-0.000
	(0.033)	(0.036)	(0.046)	(0.046)	(0.038)	(0.040)	(0.031)	(0.033)	(0.032)	(0.033)	(0.034)	(0.035)	(0.054)	(0.056)
Civic education	0.036	0.029	0.077	0.082*	0.049	0.052	0.039	0.037	0.046	0.047	0.030	0.030	0.109	0.085
	(0.038)	(0.038)	(0.048)	(0.046)	(0.043)	(0.043)	(0.032)	(0.032)	(0.032)	(0.032)	(0.039)	(0.038)	(0.070)	(0.070)
constant	0.870***	0.886***	0.726***	0.804***	0.819***	0.766***	0.801***	0.825***	0.771***	0.808***	0.770***	0.706***	0.150***	0.202**
	(0.016)	(0.062)	(0.025)	(0.076)	(0.017)	(0.063)	(0.014)	(0.054)	(0.016)	(0.053)	(0.015)	(0.056)	(0.028)	(0.088)
No. of observations	704	704	704	704	704	704	704	704	704	704	704	704	704	704
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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

Table 3.15: Network effects on turnout measure index 1 using imputed data

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.017	0.012	0.016	0.075**	0.072**	0.072**
	(0.023)	(0.024)	(0.024)	(0.031)	(0.031)	(0.031)
Hotline	0.047**	0.045**	0.048**	0.067**	0.068**	0.066**
	(0.019)	(0.019)	(0.019)	(0.033)	(0.033)	(0.033)
Civic education	0.028	0.026	0.031	0.043	0.042	0.046
	(0.020)	(0.020)	(0.020)	(0.032)	(0.033)	(0.033)
Network	0.197***	0.196***	0.049*	0.242***	0.268***	0.064**
	(0.062)	(0.073)	(0.025)	(0.073)	(0.078)	(0.027)
Network x Newspaper	**-0.013**	**-0.023**	**-0.038**	**-0.029**	**-0.033**	**-0.036**
	(0.020)	(0.019)	(0.025)	(0.030)	(0.027)	(0.031)
Network x Hotline	**-0.037***	**-0.041****	**-0.032**	**-0.065****	**-0.062****	**-0.061**
	(0.020)	(0.018)	(0.023)	(0.025)	(0.024)	(0.038)
Network x Civic education	**-0.029**	**-0.037****	**-0.039***	**-0.037**	**-0.026**	**-0.038**
	(0.022)	(0.019)	(0.022)	(0.027)	(0.029)	(0.031)
constant	0.776***	0.798***	0.869***	0.762***	0.785***	0.873***
	(0.040)	(0.038)	(0.046)	(0.056)	(0.057)	(0.061)
No. of observations	1,514	1,514	1,514	704	704	704
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

3.7 Discussion

3.7.1 Conceptual framework

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

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

$$\max_{x_i} E_{\Omega_i} U(G(x_i, x_{-i}), x_i) - C(x_i) \tag{3.4}$$

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 $C(x_i)$ (with $\frac{\partial C}{\partial x_i} > 0$) is the total material cost of the action (e.g., transport cost, opportunity cost of time, cost of text messaging). To capture non-instrumental motivations – e.g., civic-mindedness – we allow x_i to enter the function U independently from the outcome of the voting process G. This could be related to either intrinsic or extrinsic motivations. The latter could be due to the subjective or social cost of deviating from a behavioral norm (e.g., *Akerlof, 1997*).

The first order condition is:

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

It illustrates how we expect a voter education campaign to influence voter participation. First,

the campaign can change voters' information set Ω_i. Distributing information about the electoral process may convince voters of the integrity of the 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 $\partial U/\partial x_i$ by raising civic-mindedness (and the emotional cost of not voting). Both channels are conducive to increasing voter participation: that is the direct impact we expect from the voter education campaign that we study.

Focusing on turnout, difficulties arise in the absence of non-instrumental motivations, i.e., if $\frac{\partial U}{\partial x_i} = 0$. Optimal turnout requires $E_{\Omega_i}\left[\frac{\partial U}{\partial G}\frac{\partial G}{\partial x_i}\right] = \frac{\partial C}{\partial x_i}$. Since a single vote has little effect on the electoral outcome – $\partial G(x_i, x_{-i})/\partial x_i$ is small – voting is not individually rational unless the cost of voting $C(x_i)$ is minimal. That constitutes the paradox of not voting that dates back at least to *Downs (1957)*. However 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 uncertainty and a large population eliminates the possibility that high turnout arises in equilibrium. Recently, *Myatt (2012)* recovered the idea that $\partial G(x_i, x_{-i})/\partial x_i$ depends on the perceived competitiveness of the election. Myatt considers a two-candidate election in which there is aggregate uncertainty about the popularity of each candidate. Despite an underdog effect through which higher turnout from the underdog compensates the advantage of the frontrunner, Myatt finds that turnout is high and that it peaks in elections that are expected to be close.

Apart from proposing a positive impact of voter education on political participation including voter turnout, we argue that the incumbent, i.e., the candidate that is expected to win in our case, will benefit most from this increased political participation. If the incumbent values the legitimacy provided by the electoral process, the election may then turn into a turnout contest across locations for clientelistic benefits: the incumbent can look at turnout per ballot station and reward those with high turnout after the election.[26]

[26] We can look at rigorous explanations for favoring incumbents in the two-candidate voting contest of *Myatt (2012)*. Take the event that the underdog has complete turnout, i.e., that all his supporters actually vote (note that the underdog effect proposed by Myatt may constitute a likely explanation for complete turnout by the supporters of the underdog). Then, an increase in the value of the election induced by voter education can only increase turnout for the incumbent. The same result can be sustained without assuming complete turnout for the

We now turn to peer effects, that is, to the social reinforcement and diffusion effects triggered by the campaign. Some of these effects may operate in the same direction as direct effects, as information and civic-mindedness produced by the campaign are cheaply transmitted to peers. See for instance *Montgomery and Casterline (1996)* on social learning. It is also possible that peers internalize the positive direct effects of the treatment on turnout and on voting for the incumbent. In this case, we may observe a free-riding peer effect, reducing the effect of the campaign on individuals who are most central geographically and socially. Free riding may arise because increased turnout favoring the incumbent makes the electoral race less close, triggering the kind of effect discussed in *Myatt (2012)*.

3.7.2 Free-riding vs. saturation effect

Our estimations of network effects on voting behavior show three main results. First, individuals with a larger social network have a higher propensity to vote, *ceteris paribus*. Second, the treatments, especially the hotline, have a positive average effect on individual turnout. Third, the effect of the campaign on turnout is smaller among individuals with a larger social network. Our preferred interpretation of the latter result is free-riding.

A potential criticism of this interpretation is that the negative coefficient of the term $\frac{1}{N}\sum_{j\neq i} g_{ij}T_i$ may reflect a voter saturation effect rather than free-riding. As individuals with a larger social network vote with a high probability, it may be more difficult for them to further increase their likelihood of voting. This, and not free-riding, could explain why the effect of the treatment on these individuals is weaker than on individuals with a smaller social network. To show this formally, let P_i represent individual i's propensity to vote in the absence of treatment. We now assume that turnout among the targeted follows:

$$y_i = P_i + \beta T_i + \gamma \frac{1}{N}\sum_{j\neq i} g_{ij}T_i + \pi P_i T_i + \varepsilon_i \tag{3.5}$$

where γ captures free-riding as before and a significantly negative π coefficient indicates voter saturation. A bias in the estimation of γ arises if P_i is correlated with network size $\frac{1}{N}\sum_{j\neq i} g_{ij}$.

underdog to begin with: namely by postulating an increase in the value of the election only for the supporters of the incumbent (this is in line with the clientelistic story that was mentioned).

To demonstrate, let $P_i = \alpha + \delta \frac{1}{N} \sum_{j \neq i} g_{ij}$ and replace P_i in (3.5):

$$y_i = \alpha + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} + (\beta + \alpha\pi) T_i + (\gamma + \delta\pi) \frac{1}{N} \sum_{j \neq i} g_{ij} T_i + \varepsilon_i \tag{3.6}$$

Comparing (3.6) with (3.2) it is immediately apparent that voter saturation – a negative π – can be misinterpreted as free-riding – a negative γ – when estimating regression (3.2).

The solution we propose is to estimate $\widehat{P}_i$ using individuals in untreated locations, and use it as control function to obtain separate estimates of γ and π. We obtain $\widehat{P}_i$ by regressing, using control individuals only, y_i on network size $\frac{1}{N} \sum_{j \neq i} g_{ij}$ and other characteristics known to affect turnout, such as gender and age. Because treatment is assigned randomly, $\widehat{P}_i$ is a consistent estimate of treated individuals' propensity to vote in the absence of treatment. We can thus estimate (3.5) on targeted individuals using $\widehat{P}_i$ in lieu of P_i. Since by design $T_i = 1$ for the targeted, the estimated regression boils down to:[27]

$$y_i = \beta + (1 + \pi) \widehat{P}_i + \gamma \frac{1}{N} \sum_{j \neq i} g_{ij} + \varepsilon_i \tag{3.7}$$

Voter saturation $\pi < 0$ requires that the coefficient of $\widehat{P}_i$ be less than one. Coefficient γ in regression (3.7) is estimated free of voter saturation bias.

[27] 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 they can be ignored.

Table 3.16: Estimates of network effects controlling for possible saturation effect using only sample of respondents targeted by the hotline

	Panel A: Chatting											
	Turnout measures											
	Self-reported		**Adjusted**		**Finger measure**		**Index 1**		**Index 2**		**Interviewer**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Network	-0.357**	-0.458**	-0.242	-0.344	0.320	0.281	-0.272**	-0.313**	-0.214*	-0.235*	-0.168	-0.208
	(0.172)	(0.198)	(0.212)	(0.217)	(0.295)	(0.305)	(0.118)	(0.137)	(0.121)	(0.134)	(0.277)	(0.296)
Predicted propensity to vote	1.738**	2.127**	0.919**	0.861*	-0.387	-0.450	1.439***	1.531***	1.186***	1.100**	1.019	1.239
	(0.761)	(0.867)	(0.432)	(0.450)	(0.889)	(0.913)	(0.491)	(0.581)	(0.429)	(0.493)	(0.947)	(1.035)
constant	-0.471	-0.847	0.225	0.288	1.110*	1.042	-0.211	-0.314	-0.015	0.034	0.115	-0.170
	(0.624)	(0.735)	(0.273)	(0.299)	(0.658)	(0.656)	(0.371)	(0.454)	(0.307)	(0.360)	(0.657)	(0.709)
Adjusted R-squared	0.023	0.034	0.024	0.064	0.013	0.029	0.032	0.028	0.033	0.036	0.008	0.045
No. of observations	229	229	229	229	229	229	229	229	229	229	229	229
	Panel B: Kinship											
	Turnout measures											
	Self-reported		**Adjusted**		**Finger measure**		**Index 1**		**Index 2**		**Interviewer**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Network	-0.491**	-0.650**	-0.316*	-0.454**	0.170	0.080	-0.405**	-0.486**	-0.348**	-0.394**	-0.435	-0.589
	(0.246)	(0.281)	(0.189)	(0.219)	(0.326)	(0.341)	(0.160)	(0.193)	(0.145)	(0.174)	(0.370)	(0.414)
Predicted propensity to vote	1.760**	2.127**	1.004**	0.898*	-0.170	-0.221	1.586***	1.688***	1.358***	1.259**	1.492	1.881
	(0.800)	(0.886)	(0.440)	(0.463)	(0.979)	(0.993)	(0.528)	(0.619)	(0.449)	(0.523)	(1.045)	(1.157)
constant	-0.514	-0.882	0.147	0.230	0.986	0.905	-0.341	-0.457	-0.154	-0.099	-0.231	-0.637
	(0.667)	(0.765)	(0.304)	(0.332)	(0.759)	(0.751)	(0.408)	(0.496)	(0.333)	(0.397)	(0.752)	(0.828)
Adjusted R-squared	0.022	0.035	0.014	0.065	-0.005	0.019	0.030	0.031	0.028	0.037	0.003	0.038
No. of observations	229	229	229	229	229	229	229	229	229	229	229	229

Table 3.16: Estimates of network effects controlling for possible saturation effect using only sample of respondents targeted by the hotline (cont.)

	Panel C: Geographical proximity											
	Turnout measures											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Network	-0.685*	-0.829**	-0.130	-0.130	0.075	0.081	-0.121	-0.088	-0.061	-0.030	-0.040	-0.024
	(0.392)	(0.410)	(0.082)	(0.087)	(0.057)	(0.061)	(0.101)	(0.113)	(0.079)	(0.080)	(0.069)	(0.075)
Predicted propensity to vote	8.165*	9.776**	1.395**	1.158**	-1.159	-1.181	1.934	1.484	1.065	0.585	1.264	1.425
	(4.429)	(4.668)	(0.601)	(0.590)	(0.789)	(0.786)	(1.310)	(1.470)	(0.921)	(0.933)	(2.037)	(2.182)
constant	-6.905	-8.561*	-0.322	-0.225	1.895***	1.773***	-0.813	-0.503	-0.047	0.294	-0.140	-0.426
	(4.284)	(4.548)	(0.521)	(0.505)	(0.683)	(0.667)	(1.157)	(1.293)	(0.789)	(0.780)	(1.596)	(1.700)
Adjusted R-squared	0.034	0.050	0.026	0.099	0.008	0.021	0.020	0.023	0.013	0.039	-0.008	0.015
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 and province dummies. 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 reported; these are corrected by clustering at the enumeration area level, but do not account for the fact that 'predicted propensity to vote' is an estimate.

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

Equation 3.7 was estimated with and without provincial dummies and individual controls, for different turnout measures y_i, separately for each treatment and each network measure. The results for the individuals targeted by the hotline are displayed in *Table 3.16*. The top panel uses chatting for g_{ij}, the second panel uses kinship, and the third panel uses geographical proximity. Point estimates of γ are negative in most regressions and they are significant in many of them. On the other hand, the point estimates of the coefficient of $\widehat{P}_i$ are in most cases larger than one. Similar results are obtained for the sample of individuals targeted by the civic education.[28] Regarding the targeted by the newspaper, the evidence is less clear. We find negative point estimates of γ practically only when considering geographical distance, and in most regressions none of the relevant coefficients is significant. Overall, this evidence corroborates the hypothesis of a free-riding effect in the hotline and civic education areas.[29]

3.7.3 Endogenous network 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 (3.2) and (3.3) estimate the combined exogenous and endogenous peer effects. In the setting of our experiment, endogenous peer effects correspond to the situation where subjects are less likely to vote when they know that more of their social or geographical neighbors will vote; it is a direct strategic substitution effect. In contrast, an exogenous peer effect would be when the treatment of individual j has an effect on the voting behavior of individual i, although the treatment did not affect j's behavior. This could happen, for instance, because j is already voting with high probability but now uses the campaign material to induce i to vote. Distinguishing between the two types of peer effects may help confirm free-riding.

To throw some light on the issue, we estimate a model that incorporates endogenous peer

[28] Contrary to what happens in the sample of respondents targeted by the hotline, the evidence of free-riding in civic education areas is particularly strong when using geographical proximity.

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

effects but assumes away exogenous effects as follows:[30]

$$y_i = \alpha + \beta T_i + \delta \frac{1}{N} \sum_{j \neq i} g_{ij} + \lambda \frac{1}{N} \sum_{j \neq i} g_{ij} y_j + \varepsilon_i \tag{3.8}$$

If peer effects estimated in regressions (3.2) and (3.3) are driven exclusively by exogenous peer effects, then we should find no evidence of endogenous peer effects, i.e., we should observe $\lambda = 0$.

Regression (3.8) cannot be estimated using OLS due to the reflection problem (e.g., *Manski 1993*). Therefore, we instrument the voting behavior of i's neighbors ($\frac{1}{N} \sum_{j \neq i} g_{ij} y_j$) with their treatment ($\frac{1}{N} \sum_{j \neq i} g_{ij} T_j$).[31] The equation is estimated using two-stages least squares and standard errors are clustered at the EA level. 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 voted as a result of treatment.

The full results for turnout measure index 1 are presented in *Table 3.17*. The estimates of regression (3.8) are displayed in columns (1)-(3). The estimates of a similar regression for untargeted respondents are displayed in columns (4)-(6). All regressions are estimated using follow-up observations only. We control for provincial dummies, EA characteristics and individual characteristics. The coefficient of $\frac{1}{N} \sum_{j \neq i} g_{ij} y_j$ is negative in all regressions, significantly so for reinforcement effects through geographical proximity.[32] This offers some additional support to our free-riding interpretation.

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

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

[32] The coefficient λ in equation 3.8 is more likely to be significant when two conditions are met for at least one of the three interventions: (i) a significant treatment effect ($\beta \neq 0$ in equation 3.1); and (ii) a significant reduced-form network effect ($\gamma \neq 0$ in equation 3.2). These two equations are seldom verified simultaneously in our data, which goes a long way in explaining why the evidence for endogenous peer effects is not stronger.

Table 3.17: Endogenous network effects on turnout measure index 1

	reinforcement effect (targeted vs. control)			diffusion effect (untargeted vs. control)		
	chatting	**kinship**	**proximity**	**chatting**	**kinship**	**proximity**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Newspaper	0.033	0.024	0.005	0.123**	0.113***	0.079**
	(0.049)	(0.043)	(0.018)	(0.055)	(0.042)	(0.039)
Hotline	0.103**	0.084**	0.023	0.118**	0.106***	0.060
	(0.050)	(0.039)	(0.026)	(0.054)	(0.040)	(0.051)
Civic education	0.063	0.032	0.012	0.065	0.040	0.032
	(0.046)	(0.041)	(0.022)	(0.053)	(0.049)	(0.044)
Network	2.226	3.137	0.554***	2.108	2.282	0.151
	(2.589)	(3.916)	(0.141)	(1.867)	(1.405)	(0.443)
Endogenous network effect	**-2.519**	**-3.701**	**-0.665*** **	**-2.375**	**-2.522**	**-0.140**
	(3.093)	(4.626)	(0.180)	(2.189)	(1.636)	(0.589)
constant	0.831***	0.879***	0.828***	0.728***	0.762***	0.781***
	(0.096)	(0.097)	(0.044)	(0.103)	(0.098)	(0.087)
No. of observations	845	845	845	381	381	381
Controls	Yes	Yes	Yes	Yes	Yes	Yes

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

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

3.8 Concluding remarks

Using a large-scale field experiment, we have investigated how voter education treatments affect political participation in the 2009 elections in Mozambique. Three types of interventions were tested: distribution of an independent newspaper; access to a text message hotline; and a civic education campaign. The interventions are shown to increase voter turnout and to increase information of targeted and untargeted individuals in treated locations. The focus of this paper is on peer effects. Using several measures of network centrality based on social and geographical connectedness, we estimate reinforcement and diffusion network effects by comparing control individuals with, respectively, targeted and untargeted individuals in treated locations. We find that network effects on political participation are consistently negative, i.e., connectedness decreases participation in face of the voter education interventions. This is particularly clear for the hotline. At the same time, information and interest in politics are positively transmitted across peers.

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

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

Bibliography

[1] Aker, J., Collier, P. and Vicente, P. C. (2013). "Is information power? Using cell phones during an election in Mozambique," Universidade Nova de Lisboa, mimeograph.

[2] Akerlof, G. A. (1997). "Social distance and social decisions," *Econometrica*, 65(5), 1005-1028.

[3] Angelucci, M. and De Giorgi, G. (2009). "Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption?," *American Economic Review*, 99(1), 486–508.

[4] Angelucci, M., De Giorgi, G., Rangel, M. and Rasul, I. (2010). "Family networks and school enrollment: Evidence from a randomized social experiment," *Journal of Public Economics*, 94(3-4), 197–221.

[5] Bandiera, O. and Rasul, I. (2006). "Social networks and technology adoption in Northern Mozambique," *Economic Journal*, 116(514), 862–902.

[6] Banerjee, A. V., Kumar, S., Pande, R. and Su, F. (2011). "Do informed voters make better choices? Experimental evidence from urban India," Harvard University, mimeograph.

[7] Bauer, M., Cassar, A., Chytilova, J. and Henrich, J. (2012). "Warfare experience during ontogeny increases egalitarian and parochial motivations," Charles University in Prague, mimeograph.

[8] Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). "Identification of peer effects through social networks," *Journal of Econometrics*, 150(1), 41-55.

[9] Bruhn, M. and McKenzie, D. (2009). "In pursuit of balance: Randomization in practice in development field experiments," *American Economic Journal: Applied Economics*, 1(4), 200-32.

[10] Chandrasekhar, A. G. and Lewis, R. (2012). "Econometrics of sampled networks," Massachusetts Institute of Technology, mimeograph.

[11] Collier, P., and Vicente, P. C. (2009), "Votes and violence: Evidence from a field experiment in Nigeria," BREAD, Working Paper 232.

[12] Dale, A. and Strauss, A. (2009). "Don't forget to vote: Text message reminders as a mobilization tool", *American Journal of Political Science*, 53(4), 787–804.

[13] Deaton, Angus (2009). "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," *Proceedings of the British Academy, 2008 Lectures*, 162, 123-160, Oxford: Oxford University Press

[14] Dhillon, A. and Peralta, S. (2002). "Economic theories of voter turnout," *Economic Journal*, 112, F332–F352.

[15] Downs, A. (1957). *An Economic Theory of Democracy*, New York: Harper and Row.

[16] Fafchamps, M. and Vicente, P. C. (2013). "Political violence and social networks: Experimental evidence from a Nigerian election," *Journal of Development Economics*, 101, 27-48.

[17] Feddersen, T. (2004). "Rational choice theory and the paradox of not voting," *Journal of Economic Perspectives*, 18(1), 99-112.

[18] 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.

[19] Gine, X. and Mansuri, G. (2011). "Together we will: experimental evidence on female voting behavior in Pakistan," World Bank, Policy Research Working Paper 5692.

[20] Kling, J. R., Liebman, J. B. and Katz, L. F. (2007): "Experimental analysis of neighborhood effects," *Econometrica*, 75(1), 83–119.

[21] Miguel, E. and Kremer, M. (2004). "Worms: Identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 72(1), 159–217.

[22] Macours, K. and Vakis, R. (2008). "Changing households' investments and aspirations through social interactions: Evidence from a randomized transfer program in a low-income country," World Bank, Policy Research Working Paper 5137.

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

[24] Montgomery, M. R. and Casterline, J. B. (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, pp. 151-175.

[25] Myatt, D. P. (2012). "A rational choice theory of voter turnout," London Business School, mimeograph.

[26] Nickerson, D. W. (2008). "Is voting contagious? Evidence from two field experiments," *American Political Science Review*, 102(1), 49–57.

[27] Palfrey, T. and Rosenthal, H. (1983). "A strategic calculus of voting," *Public Choice*, 41(1), 7-53.

[28] Palfrey, T., and Rosenthal, H. (1985). "Voter participation and strategic uncertainty," *American Political Science Review*, 79(1), 62-78.

[29] Pop-Eleches, C, Thirumurthy, H., Habyarimana, J. P., Zivin, J. G., Goldstein, M. P., De Walque, D., MacKeen, L., Haberer, J., Kimaiyo, S., Sidle, J., Ngare, D. and Bangsberg, D. R. (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.

[30] Schafer, J. L. and Graham, J. W. (2002). "Missing data: Our view of the state of the art," *Psychological Methods*, 7(2), 147–177.

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

[32] Vicente, P. C. (2007). "Is vote-buying effective? Evidence from a field experiment in West Africa," BREAD, Working Paper 161.

[33] Wantchekon, L. (2003). "Clientelism and voting behavior: Evidence from a field experiment in Benin," *World Politics*, 15(3), 399-422.

Chapter 4

Conclusion

This thesis documents the effect of social influence on voter's turnout in Mozambique and contributes to the understanding of how voter education campaigns impact turnout.

The empirical analysis conducted here makes three main contributions to the literature. First, it clearly document that the reach of voter education campaigns goes far beyond the targeted individuals. The effects of the campaign spread to other household members, friends, relatives and neighbors, through influence. These results also show that different types of voter education campaigns have different effects on treated individuals and trigger influence at different levels.

Second, the evidence presented here suggests that objective information is not as easily transmitted as social pressure or interest. It seems that interpersonal influence regarding voting within the household tends to assume the form of social pressure. This social pressure towards voting is especially effective in the case of younger individuals.

Third, the results reported in Chapter 3 show that the effect of the campaign on targeted and untargeted respondents depends on their social and geographical proximity to other targeted respondents. The results of this chapter show that the interventions increase voter participation on average, but much less so among individuals who are socially or geographically closer to other targeted individuals. This result is interpreted as evidence of free-riding on voting as a civic duty.

Further research is required in order to build on the findings of this experiment. As these

findings are based on a field experiment conducted in Mozambique, it would be worth investigating whether they can be extended to other developing countries. In particular, it would be very useful to examine if the pattern of positive direct effects and negative network effects on turnout also occurs in other countries. This is a curious pattern that should be the subject of further research.

The heterogeneous effects reported in Chapter 2 show that interpersonal influence within the household depends on the age of the targets of influence. Therefore, studies of the determinants of interpersonal influence, in different contexts, should start looking more at the impact of the characteristics of the targets of influence, and not just be focused on the characteristics of the influencers.

The findings reported in this thesis have implications for the design and evaluation of voter education campaigns. To begin, campaigns should be carefully designed to attain desirable outcomes and not simply seek to increase social pressure to vote. Interventions with a nudge component, like the civic education intervention, seem to work mainly through persuasion (raising the pressure to vote). The increased pressure associated with the low levels of political information and the large gap in popularity between the ruling party and the opposition may trigger a kind of herd behavior that favors the party in power. This implies that, in similar settings, this type of campaign may be used as a political tool to reinforce the incumbent's dominant position. In such contexts, perhaps interventions aimed at involving individuals in the elections, such as the hotline, should be preferred over campaigns based on nudging. Further, the design of the campaign should take into consideration the role of the characteristics of potential secondary targets as well as the role of social networks in the diffusion of the effects of the campaign. Finally, an accurate cost-benefit analysis of voter education campaigns should account for their spillover effects.

Appendix A

Tables

Table A.1: Model of probability of selecting a second interviewee

	Dependent variable: Household has information regarding a second interviewee		
Covariates	**(1)**	**(2)**	**(3)**
Gender	0.257***	0.250***	0.164
	(0.076)	(0.077)	(0.177)
Age	0.003	0.002	0.003
	(0.003)	(0.003)	(0.006)
Household size	0.058***	0.058***	0.077***
	(0.013)	(0.013)	(0.026)
Married or in an union	-0.107	-0.130*	-0.162
	(0.072)	(0.075)	(0.151)
Has completed primary school	-0.017	-0.047	0.118
	(0.078)	(0.082)	(0.169)
Has a job	-0.009	-0.021	-0.324*
	(0.089)	(0.089)	(0.177)
Province: Zambezia	-0.442***	-0.403***	-0.430***
	(0.120)	(0.125)	(0.125)
Province: Maputo	-0.349***	-0.323***	-0.368***
	(0.114)	(0.118)	(0.121)
Province: Gaza	-0.274**	-0.215*	-0.230*
	(0.108)	(0.118)	(0.118)
Newspaper	0.039	0.050	0.612
	(0.107)	(0.111)	(0.739)
Hotline	0.115	0.130	-0.083
	(0.108)	(0.112)	(0.889)
Civic education	0.039	0.065	-4.178***
	(0.115)	(0.119)	(0.774)
Untargeted household	-0.268**	-0.273**	-0.261**
	(0.106)	(0.109)	(0.113)
Knows candidates names		0.039	0.008
		(0.038)	(0.065)
Interest		-0.008	0.010
		(0.043)	(0.075)
Self-reported turnout		-0.078	0.062
		(0.235)	(0.391)
How often discusses politics		0.028	-0.021
		(0.046)	(0.093)
Newspaper x how often discusses politics			0.218*
			(0.126)
Hotline x has a job			0.534**
			(0.239)
Hotline x completed primary school			-0.375*
			(0.224)
Civic education x has a job			0.499*
			(0.262)
Civic education x self-reported turnout			4.239***
			(0.560)
Chi-squared statistic	72.48***	74.71***	1,019.87***
No. of observations	1,568	1,501	1,501
Includes interaction terms	No	No	Yes

Note: Regressions include observations for households where there were other adults besides the main respondent. The coefficients of some characteristics and interaction terms that are not statistically significant are omitted. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table A.2 : Model with interventions as instruments - Second step estimation results, using bootstrapping

Panel's predicted outcome	Second interviewee's outcomes					
	Interest in elections	Turnout				
		Self-reported	Finger measure	Index 1	Index 2	Interviewer's assessment
	(1)	(2)	(3)	(4)	(5)	(6)
Self-reported turnout	0.191*	0.058	0.072	0.050	0.053	0.079*
	(0.106)	(0.042)	(0.050)	(0.038)	(0.039)	(0.046)
Adjusted turnout	0.267*	0.086	0.104	0.075	0.082	0.116*
	(0.152)	(0.064)	(0.065)	(0.056)	(0.056)	(0.063)
Finger measure	0.217	0.082	0.092	0.071	0.075	0.099*
	(0.147)	(0.053)	(0.067)	(0.048)	(0.049)	(0.060)
Index 1	0.213*	0.063	0.080	0.055	0.060	0.089*
	(0.127)	(0.047)	(0.053)	(0.042)	(0.043)	(0.050)
Index 2	0.236**	0.072	0.089^{+}	0.064	0.070	0.102**
	(0.116)	(0.050)	(0.055)	(0.045)	(0.045)	(0.050)
Interviewer's assessment	0.177**	0.059	0.070^{++}	0.052	0.057*	0.080**
	(0.087)	(0.037)	(0.043)	(0.034)	(0.034)	(0.038)

Note: Each cell corresponds to a linear model regression of an outcome variable of the second interviewee on the standardized predicted main respondent's outcome variable. The values displayed are the coefficients of standardized predicted main respondent's outcome variables (displayed in rows) in the regression of second interviewee's outcome variables (displayed in columns). Standard errors reported between parenthesis; these were obtained by bootstrapping.

$^{+}$ significant at 11%; '++ significant at 10.5%; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.3: Model with interactions as instruments - First step estimation results

	Main respondent's outcomes									
	Information		Interest in elections	Turnout						Open letter
	About elections	About voting		Self-reported	Adjusted	Finger measure	Index 1	Index 2	Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Newspaper	0.111	0.096	-0.028	0.029	0.050	0.017	0.033	0.041	0.039	0.084*
	(0.088)	(0.093)	(0.108)	(0.037)	(0.048)	(0.042)	(0.035)	(0.036)	(0.036)	(0.050)
Hotline	0.138	0.055	0.079	0.076***	0.107**	0.056**	0.070***	0.075***	0.086***	-0.034
	(0.086)	(0.084)	(0.107)	(0.026)	(0.044)	(0.028)	(0.024)	(0.027)	(0.025)	(0.036)
Civic education	0.051	-0.022	0.072	0.043	0.079*	0.055*	0.041	0.050*	0.075***	0.053
	(0.093)	(0.085)	(0.093)	(0.028)	(0.041)	(0.029)	(0.025)	(0.027)	(0.027)	(0.049)
Newspaper x age	0.005	0.009**	0.000	0.003	0.000	0.002	0.002	0.002	0.004**	-0.002
	(0.005)	(0.004)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Newspaper x 12 years edu.	0.131	0.148	0.244	-0.079	0.063	-0.149	-0.041	-0.031	-0.126	-0.110
	(0.140)	(0.208)	(0.297)	(0.113)	(0.125)	(0.100)	(0.105)	(0.107)	(0.100)	(0.114)
Civic x 12 years education	0.277**	0.252	-0.001	-0.147	-0.208	-0.255	-0.147	-0.160	-0.184	0.124
	(0.133)	(0.207)	(0.273)	(0.134)	(0.177)	(0.155)	(0.123)	(0.122)	(0.126)	(0.168)
Hotline x farmer	-0.205	0.052	0.073	-0.086**	-0.054	-0.069	-0.090**	-0.075*	-0.078	-0.067
	(0.181)	(0.140)	(0.159)	(0.043)	(0.067)	(0.060)	(0.041)	(0.043)	(0.050)	(0.056)
Paper x domestic worker	-0.299	0.213	0.015	0.197***	0.169	0.187**	0.190***	0.179***	0.241***	0.047
	(0.193)	(0.142)	(0.210)	(0.070)	(0.108)	(0.084)	(0.064)	(0.065)	(0.069)	(0.111)
Civic x artisan	0.345	-0.045	0.454*	0.156**	0.250**	0.136*	0.188***	0.173**	0.265***	0.164
	(0.293)	(0.440)	(0.266)	(0.074)	(0.098)	(0.080)	(0.070)	(0.077)	(0.079)	(0.179)
Newspaper x artisan	0.377*	-0.200	0.151	0.182**	0.136	0.097	0.156**	0.083	0.241***	0.336***
	(0.209)	(0.279)	(0.270)	(0.078)	(0.128)	(0.118)	(0.075)	(0.079)	(0.089)	(0.087)
Civic x cattle	-0.073	-0.004	-0.067	-0.088	-0.092	-0.060	-0.088	-0.092	-0.111*	-0.021
	(0.162)	(0.152)	(0.156)	(0.063)	(0.080)	(0.074)	(0.057)	(0.059)	(0.065)	(0.070)
Newspaper x cattle	-0.118	-0.237	0.100	-0.148**	-0.239***	-0.157**	-0.154***	-0.172***	-0.221***	-0.030
	(0.172)	(0.161)	(0.171)	(0.059)	(0.077)	(0.064)	(0.055)	(0.054)	(0.059)	(0.088)
constant	0.090	0.142	0.028	0.793***	0.558***	0.672***	0.721***	0.669***	0.644***	0.180***
	(0.135)	(0.119)	(0.139)	(0.046)	(0.062)	(0.055)	(0.043)	(0.043)	(0.046)	(0.054)
Adjusted R-squared	0.043	0.089	0.004	0.030	0.041	0.015	0.037	0.041	0.049	0.017
F-statistic	1.41	1.44	0.51	2.97***	2.41**	2.62***	4.07***	3.83***	6.03***	2.35**
No. of observations	966	969	966	944	944	944	944	944	944	963
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include observations for main respondents living in treated and control areas. All regressions are OLS. In all regressions I control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level. The F-statistic reported refers to the null hypothesis that the coefficients of all interaction terms are jointly equal to zero.

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

Table A.4: Model with interactions as instruments - Second step estimation results

Covariates	Dependent variable: Second interviewee's turnout, measured by interviewer assessment											
	Sample of second interviewees living in targeted or control households						Excluding interviewees living in targeted households					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.031	0.035	0.040	0.030	0.028	0.029	0.007	0.005	0.020	0.003	0.003	0.003
	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)	(0.047)	(0.059)	(0.060)	(0.058)	(0.060)	(0.059)	(0.060)
Hotline	0.031	0.044	0.046	0.033	0.033	0.040	0.020	0.031	0.041	0.018	0.022	0.031
	(0.056)	(0.062)	(0.053)	(0.055)	(0.056)	(0.053)	(0.066)	(0.070)	(0.064)	(0.066)	(0.066)	(0.062)
Civic education	0.057	0.059	0.050	0.058	0.054	0.051	0.033	0.034	0.029	0.031	0.029	0.028
	(0.048)	(0.054)	(0.053)	(0.048)	(0.050)	(0.049)	(0.058)	(0.064)	(0.067)	(0.060)	(0.061)	(0.059)
Main respondent's turnout (standardized)												
Self-reported	0.052[+]						0.054					
	(0.033)						(0.037)					
Adjusted		0.047						0.055				
		(0.049)						(0.050)				
Finger measure			0.048						0.050			
			(0.033)						(0.042)			
Index 1				0.054[++]						0.056		
				(0.033)						(0.039)		
Index 2					0.056						0.057	
					(0.036)						(0.041)	
Interviewer's assessment						0.042*						0.050
						(0.025)						(0.031)
Adjusted R-squared	0.044	0.038	0.042	0.044	0.043	0.043	0.053	0.480	0.491	0.053	0.051	0.053
No. of observations	411	411	411	411	411	411	245	245	245	245	245	245
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include observations for second respondents living with targeted (in treated locations) and control respondents. All regressions are OLS and in all regressions I control for demographic characteristics and province dummies. Standard errors reported between parenthesis; these were obtained using Hardin's sandwich cluster-robust covariance estimator.

[+] significant at 11%; [++] significant at 10.5%; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.5: Heterogeneous effects associated with second interviewee's age - IV model

Covariates	Dependent variables: Second interviewee's turnout measures					
	Self-reported		Index 1		Index 2	
	(1)	(2)	(3)	(4)	(5)	(6)
Second interviewee's age	0.024**	0.014**	0.025**	0.014**	0.028**	0.015**
	(0.226)	(0.007)	(0.012)	(0.007)	(0.013)	(0.007)
Main respondent's turnout (standardized)						
Index 1	0.050		0.100		0.159	
	(0.226)		(0.213)		(0.221)	
Index 2		0.061		0.117		0.183
		(0.222)		(0.205)		(0.208)
Interaction main respondent's turnout and age						
Index 1 x age	**-0.017***		**-0.018***		**-0.020***	
	(0.010)		(0.010)		(0.011)	
Index 2 x age		**-0.008**		**-0.009**$^{++}$		**-0.009**$^{++}$
		(0.005)		(0.005)		(0.005)
Wald chi-squared	41.99***	35.12***	57.84***	49.26***	54.62***	57.71***
No. of observations	313	313	313	313	313	313
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include observations for second respondents living in treated and control areas. All regressions were estimated using a two stage instrumental variables model. In all regressions I control for demographic characteristics and province dummies. Standard errors reported between parenthesis; these are corrected by clustering at the enumeration area level.

$^{++}$ significant at 10.5%; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.6: Heterogeneous effects associated with second interviewee's education - IV model

Covariates	Dependent variables: Second interviewee's turnout measures					
	Self-reported	Adjusted	Finger measure	Index 1	Index 2	Interviewer's assessment
	(1)	(2)	(3)	(4)	(5)	(6)
Second interviewee's completed primary school	-0.291	-0.480	-0.074	-0.266	-0.334	-0.324
	(0.297)	(0.376)	(0.292)	(0.289)	(0.307)	(0.319)
Main respondent's turnout (standardized)						
Interviewer's assessment	0.115	0.223	0.224	0.152	0.198	0.280*
	(0.153)	(0.199)	(0.162)	(0.144)	(0.151)	(0.161)
Interaction main respondent's turnout and age						
Interviewer's assessment x education	**0.393**	**0.636***	**0.222**	**0.370**	**0.454***	**0.445**
	(0.262)	(0.345)	(0.263)	(0.258)	(0.276)	(0.293)
Wald chi-squared	24.44**	32.69***	45.06***	25.59**	27.51**	34.74***
No. of observations	313	313	313	313	313	313
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Regressions include observations for second respondents living in treated and control areas. All regressions were estimated using a two stage instrumental variables model. In all regressions I control for demographic characteristics and province dummies. Standard errors reported between parenthesis; these are corrected by clustering at the enumeration area level.

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

Table A.7: Regressions of turnout measures with interaction with second interviewee's age - Reduced form specification, using inverse probability weighting correction

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.021	0.027	0.051	0.064	0.026	0.033	0.030	0.036	0.047	0.054	0.061	0.081
	(0.060)	(0.057)	(0.068)	(0.066)	(0.065)	(0.064)	(0.056)	(0.053)	(0.055)	(0.052)	(0.057)	(0.053)
Hotline	0.068	0.051	0.019	0.006	0.098*	0.069	0.061	0.044	0.067	0.050	0.105*	0.091*
	(0.052)	(0.050)	(0.063)	(0.057)	(0.058)	(0.057)	(0.048)	(0.046)	(0.049)	(0.047)	(0.055)	(0.051)
Civic education	0.073	0.066	0.069	0.068	0.092	0.066	0.069	0.064	0.078	0.076	0.109**	0.093*
	(0.055)	(0.052)	(0.069)	(0.064)	(0.065)	(0.059)	(0.052)	(0.049)	(0.052)	(0.049)	(0.054)	(0.048)
Second respondent's age	0.006***	0.008***	0.011***	0.013***	0.004	0.006**	0.006***	0.007***	0.007***	0.008***	0.007***	0.009***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Newspaper x age	**-0.001**	**-0.001**	**-0.009****	**-0.010****	**-0.002**	**-0.003**	**-0.002**	**-0.002**	**-0.002**	**-0.003**	**-0.003**	**-0.003**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Hotline x age	**-0.003**	**-0.004***	**-0.007***	**-0.007***	**-0.001**	**-0.002**	**-0.004***	**-0.004***	**-0.005****	**-0.005****	**-0.004**	**-0.005**
	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Civic x age	**-0.004**	**-0.005**	**-0.011*****	**-0.011*****	**-0.004**	**-0.005**	**-0.005****	**-0.005****	**-0.006****	**-0.007****	**-0.003**	**-0.004**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
constant	0.558***	0.534***	0.325**	0.306**	0.522***	0.489***	0.532***	0.516***	0.478***	0.460***	0.308***	0.285***
	(0.116)	(0.116)	(0.136)	(0.135)	(0.127)	(0.129)	(0.108)	(0.108)	(0.107)	(0.108)	(0.112)	(0.111)
Adjusted R-squared	0.066	0.070	0.075	0.085	0.053	0.061	0.065	0.071	0.072	0.084	0.091	0.105
No. of observations	393	382	393	382	393	382	393	382	393	382	393	382
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for second respondents living with targeted (in treated locations) and control respondents. All regressions are OLS and use as sampling weights the inverse of the probability of the individual being selected into the second interviewees sample. In the second column for each outcome I control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table A.8: Regressions with an interaction with second interviewee's education and another with main respondent's education - Reduced form specification, using inverse probability weighting correction

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.013	0.020	0.045	0.061	0.036	0.032	0.026	0.031	0.045	0.048	0.056	0.075
	(0.054)	(0.053)	(0.062)	(0.064)	(0.064)	(0.064)	(0.051)	(0.050)	(0.051)	(0.050)	(0.051)	(0.048)
Hotline	0.060	0.048	0.029	0.010	0.109*	0.070	0.057	0.041	0.065	0.049	0.099*	0.081
	(0.048)	(0.046)	(0.056)	(0.054)	(0.057)	(0.056)	(0.045)	(0.044)	(0.046)	(0.045)	(0.054)	(0.050)
Civic education	0.058	0.057	0.045	0.052	0.083	0.061	0.055	0.055	0.064	0.065	0.089*	0.084*
	(0.050)	(0.048)	(0.065)	(0.060)	(0.062)	(0.060)	(0.047)	(0.046)	(0.048)	(0.046)	(0.051)	(0.045)
Main respondent has primary schooling	0.148*	0.175**	0.132	0.175*	0.145	0.172*	0.124	0.148*	0.104	0.132	0.199**	0.230***
	(0.086)	(0.085)	(0.103)	(0.104)	(0.095)	(0.096)	(0.081)	(0.081)	(0.081)	(0.081)	(0.088)	(0.083)
Second respondent has primary schooling	-0.122**	-0.018	-0.182**	-0.058	0.034	0.096	-0.108*	-0.015	-0.094	0.003	-0.084	0.005
	(0.060)	(0.068)	(0.089)	(0.098)	(0.085)	(0.087)	(0.055)	(0.064)	(0.059)	(0.069)	(0.073)	(0.064)
Newspaper x main education	**-0.275****	**-0.276*****	**-0.275****	**-0.261****	**-0.185**	**-0.177**	**-0.257****	**-0.250****	**-0.230****	**-0.222****	**-0.297*****	**-0.303*****
	(0.113)	(0.106)	(0.135)	(0.131)	(0.129)	(0.126)	(0.107)	(0.101)	(0.107)	(0.102)	(0.109)	(0.102)
Hotline x main education	**-0.224****	**-0.227****	**-0.326****	**-0.327****	**-0.212***	**-0.205**	**-0.197***	**-0.193***	**-0.188***	**-0.184***	**-0.234***	**-0.239***
`	(0.109)	(0.115)	(0.136)	(0.143)	(0.126)	(0.132)	(0.104)	(0.109)	(0.105)	(0.111)	(0.123)	(0.130)
Civic x main education	**-0.051**	**-0.108**	**-0.100**	**-0.178**	**-0.081**	**-0.156**	**-0.048**	**-0.098**	**-0.022**	**-0.078**	**-0.082**	**-0.166***
	(0.102)	(0.110)	(0.139)	(0.148)	(0.105)	(0.106)	(0.099)	(0.106)	(0.100)	(0.106)	(0.098)	(0.096)
Newspaper x second education	**0.335*****	**0.309*****	**0.429*****	**0.400*****	**0.211***	**0.245****	**0.314*****	**0.292*****	**0.302*****	**0.281*****	**0.338*****	**0.326*****
	(0.094)	(0.093)	(0.120)	(0.122)	(0.116)	(0.112)	(0.087)	(0.086)	(0.091)	(0.092)	(0.098)	(0.088)
Hotline x second education	**0.169***	**0.145**	**0.222***	**0.212***	**-0.023**	**-0.016**	**0.149***	**0.128**	**0.139**	**0.116**	**0.129**	**0.134**
	(0.090)	(0.096)	(0.121)	(0.123)	(0.116)	(0.118)	(0.086)	(0.092)	(0.090)	(0.096)	(0.112)	(0.112)
Civic x second education	**0.113**	**0.098**	**0.200**	**0.179**	**0.035**	**0.035**	**0.119**	**0.106**	**0.133**	**0.124**	**0.102**	**0.092**
	(0.091)	(0.098)	(0.143)	(0.153)	(0.108)	(0.102)	(0.087)	(0.094)	(0.092)	(0.100)	(0.094)	(0.082)
constant	0.862***	0.619***	0.791***	0.551***	0.690***	0.538***	0.804***	0.623***	0.772***	0.596***	0.687***	0.361***
	(0.071)	(0.105)	(0.080)	(0.119)	(0.087)	(0.130)	(0.067)	(0.100)	(0.068)	(0.101)	(0.076)	(0.098)
Adjusted R-squared	0.041	0.101	0.027	0.095	0.028	0.077	0.041	0.098	0.039	0.100	0.058	0.147
No. of observations	411	382	411	382	411	382	411	382	411	382	411	382
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for second respondents living with targeted (in treated locations) and control respondents. All regressions are OLS and use as sampling weights the inverse of the probability of the individual being selected into the second interviewees sample. In the second column for each outcome I control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table A.9: Regressions of turnout measures with interaction with age difference between main and second respondents - Reduced form specification

	Turnout											
	Self-reported		Adjusted		Finger measure		Index 1		Index 2		Interviewer's assessment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.012	0.015	0.053	0.061	0.030	0.035	0.022	0.024	0.037	0.039	0.053	0.070
	(0.052)	(0.050)	(0.062)	(0.057)	(0.057)	(0.056)	(0.049)	(0.047)	(0.048)	(0.046)	(0.046)	(0.046)
Hotline	0.064	0.051	0.039	0.031	0.098**	0.074	0.056	0.043	0.064	0.051	0.098**	0.088*
	(0.042)	(0.044)	(0.050)	(0.047)	(0.050)	(0.051)	(0.039)	(0.040)	(0.040)	(0.041)	(0.046)	(0.046)
Civic education	0.080*	0.083*	0.083	0.095*	0.105*	0.085	0.074*	0.079**	0.083*	0.090**	0.127***	0.113***
	(0.042)	(0.043)	(0.062)	(0.054)	(0.055)	(0.052)	(0.040)	(0.040)	(0.043)	(0.041)	(0.041)	(0.040)
Age difference	-0.004***	-0.002	-0.007***	-0.005**	-0.004*	-0.002	-0.004***	-0.003	-0.004***	-0.003*	-0.006***	-0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Newspaper x age difference	**0.003**	**0.002**	**0.005**	**0.004**	**0.005**	**0.004**	**0.003**	**0.002**	**0.003**	**0.002**	**0.005****	**0.004**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Hotline x age difference	**0.002**	**0.001**	**0.001**	**-0.001**	**0.001**	**0.001**	**0.002**	**0.001**	**0.002**	**0.001**	**0.003**	**0.002**
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Civic x age difference	**0.004****	**0.004****	**0.005****	**0.005***	**0.003**	**0.003**	**0.004****	**0.004****	**0.004****	**0.004****	**0.004****	**0.004****
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
constant	0.881***	0.709***	0.783***	0.653***	0.798***	0.613***	0.823***	0.710***	0.791***	0.680***	0.760***	0.493***
	(0.031)	(0.083)	(0.037)	(0.099)	(0.042)	(0.098)	(0.029)	(0.073)	(0.029)	(0.074)	(0.032)	(0.083)
Adjusted R-squared	0.021	0.055	0.028	0.091	0.014	0.056	0.021	0.063	0.024	0.078	0.046	0.088
No. of observations	407	396	407	396	407	396	407	396	407	396	407	396
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

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

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