

CSAE Working Paper WPS/2012-14

Interpersonal Influence Regarding the Decision to Vote Within Mozambican Households

Ana Sílvia de Matos Vaz
University of Oxford

September 11, 2012

Abstract

Voter education is crucial to promote voters' participation. The 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 the delivery of the newspaper has almost no effect on the other people in the household. The hotline intervention affects the preferences and behavior of the other individuals, but not their information. Finally, the civic education campaign only affects the behavior of other people in the household. This paper shows 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 statistical evidence that social control increases voter turnout.

Contents

1	Introduction	2
2	Conceptual Framework and Testing Strategy	4
2.1	Model of Human Behavior subject to Interpersonal Influence . .	4
2.2	Testing Strategy	6
2.2.1	Reduced Form Specification	6
2.2.2	Structural Form Specification	9
3	Data and Descriptive Statistics	13
3.1	Treatments ¹	13
3.2	Measurement ²	14
3.3	Descriptive Statistics	17
4	Econometric Results	19
4.1	Results of Reduced Form Specification	19
4.1.1	Results	19
4.1.2	Robustness	22
4.2	Results of Structural Form Specification	23
4.2.1	Results	23
4.2.2	Robustness	25
5	Conclusion	26

¹Based on section "3.1 Treatments" of *Aker, Collier and Vicente, 2011*.

²Based on section "3.3 Measurement" of *Aker Collier and Vicente, 2011*.

1 Introduction

Voter education is key to guarantee democratic elections.³ The question that 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.⁴ Fortunately, there are factors that mitigate the effects of the limited access to mass media. First, the mobile phone is becoming a crucial tool of sharing information in the African context, with important repercussions in the political sphere (*Heacock, 2009, Ekine et al., 2010, Aker, Collier and Vicente, 2011*). Second, one of the most efficient means of spreading ideas and opinions is 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.⁵ In these circumstances, even when only a minority has direct access to the media, it is still possible that the information and opinions transmitted by the media reach the majority.

This paper focuses on the role of interpersonal influence in the diffusion of the effects of voter education campaigns. 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 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

³Some 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 Príncipe, 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 (2010)* show that different types of voter education interventions increase individuals' political participation and their information about politics.

⁴*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.

⁵Based 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.

people who cohabitate 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 almost has no effect on the other people in the household. The hotline intervention, which affects the information, interest and turnout of the treated individuals, only affects the interest and turnout of the other individuals in the household. The civic education campaign, which mainly affects treated individuals' turnout, only affects the turnout of other people in the household.

This paper shows that the transmission of 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 works through persuasion: it mainly pressures treated individuals to vote, who in their turn pressure the people around them to vote as well.

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)*, *Fafchamps and Vicente (2011)*, and *Gine and Mansuri (2011)*. 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 identifies the reinforcement and diffusion network effects triggered by a randomized campaign against political violence in Nigeria. The third estimates the peer effects associated with geographical distance and friendship on female turnout triggered by a randomized voter awareness campaign. 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. Furthermore, by examining how the effects of voter education campaigns spillover to individuals living

in the same household, this paper also adds to the literature on interpersonal influence.

The paper is organized as follows. Section 2 begins by developing a simple model of human behavior subject to interpersonal influence and describing the testing strategy. The experimental design, the data and the descriptive statistics, including balance tests, are discussed in Section 3. Subsequently, Section 4 presents the econometric results and robustness tests. Section 5 concludes.

2 Conceptual Framework and Testing Strategy

The starting point of this analysis is that voter education campaigns, even when targeted at one person per household, tend to affect more people than those who were directly treated (e.g. *Nickerson, 2008*). This happens because the effect of the campaigns is likely to spread to the people with whom the treated individual interacts, as she influences them. This section begins by proposing a conceptual framework to examine these spillovers effects. It develops a very simple model to explain how the interaction with other people may affect an individual's behavior. The section ends with the presentation of the testing strategy.

2.1 Model of Human Behavior subject to Interpersonal Influence

Human behavior is modeled as the result of a sequential process, where at each step the individual may be subject to the influence of people with whom she interacts. 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, it is possible that, 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 control; i.e., by exerting pressure, by promising a reward or

threatening with a punishment, other people may lead the individual to adopt a behavior that otherwise he would not.⁶⁷ Under this model it is possible that influence at the level of information may lead to a change in behavior, as the change in information may result in a change of preferences, which may trigger a change in behavior.

I assume that the individual chooses an action x , from a set of available actions X , to maximize the pay-off function:

$$\max_x E[U(y(x); N) | \Omega(M; F)] - c(x) - d|\bar{x} - x| \quad (1)$$

This pay-off function corresponds to the individual's valuation of the consequences of performing action x minus its direct costs and costs associated with social sanctions of performing such action.

The function $U(y(x); N)$ reflects the valuation of specific consequences y of performing action x . In the case of voting, the valuation of the consequences of voting may also include "expressive" components to vote, as the value of doing one's duty as a citizen (*Fiorina, 1976, Aldrich, 1993, Engelen, 2006*).⁸ This valuation may be affected by the opinions of other people with whom she interacts; this is reflected in the shift parameter N . It is assumed that other people make the individual care more about consequences y of behavior x , $\frac{\partial^2 U}{\partial y \partial N} > 0$.

When deciding which action to perform, the individual maximizes the expected value of the consequences of action x given her information set, Ω . The individual gathers information in two ways: directly, from the media; and indirectly, from other people with whom she interacts.⁹ So, the individual's in-

⁶A similar categorization of social influence was proposed by *Deutsch 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).

⁷In the context of the interventions discussed in this paper, an example of influence at information level would be the treated individual teaching the other household members the names of the presidential candidates. An example of influence at the level of preferences would be the treated individual convincing the 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 the other household members to vote.

⁸Note that if the function would only consider the instrumental value of voting - voting to make one's favorite candidate win - its value would probably be very close to zero. For a brief review of the literature on the paradox of voting see *Feddersen, 2004*.

⁹*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.

formation set is an increasing function of her access to media (M) and the information she obtains from others (F), $\frac{\partial \Omega}{\partial M} > 0$ and $\frac{\partial \Omega}{\partial F} > 0$.

The term $c(x)$ represents the direct costs of performing action x . In the case of voting this term may include the transportation cost of going to the pooling station and the opportunity cost of the time used in going to vote.

Finally, the term $d|\bar{x} - x|$ represents the costs associated with the sanctions that other people may impose to the individual when her behavior deviates from the norm $\bar{x}$ (e.g. *Akerlof, 1997*).

2.2 Testing Strategy

2.2.1 Reduced Form Specification

I am interested in measuring influence within households and identify at which level it takes place: information, preferences and/or behavior. 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 $y_{i,l}$, $z_{i,l}$, and $w_{i,l}$ be measures of electoral information, preferences, and behavior, respectively, of the main respondent living in household i , in village l . Let $T_i = 1$ if the household is located in a treated area, 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:

$$\begin{aligned} y_{i,l} &= \alpha_1 + \beta_1 T_l + u_{i,l} \\ z_{i,l} &= \alpha_2 + \beta_2 T_l + v_{i,l} \\ w_{i,l} &= \alpha_3 + \beta_3 T_l + e_{i,l} \end{aligned} \tag{2}$$

These regressions can also be estimated with village or individual fixed effects

to net out possible time effects and individual unobservables. Coefficients β_1, β_2 and β_3 are average treatment effects on information, preferences, and behavior, respectively.

Estimating average treatment effects is not the focus of this paper, however. This issue is covered in detail in *Aker, Collier and Vicente (2011)*. Here we 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 and denoted by the upper scrip S). I estimate the average indirect effect of the campaign on these non-treated individuals living in households in treated areas. Given the randomized nature of the data, the spillover (average) effects associated with the interventions can be estimated using the following regressions:

$$\begin{aligned} y_{i,l}^S &= \alpha_1^S + \beta_1^S T_l + u_{i,l} \\ z_{i,l}^S &= \alpha_2^S + \beta_2^S T_l + v_{i,l} \\ w_{i,l}^S &= \alpha_3^S + \beta_3^S T_l + e_{i,l} \end{aligned} \tag{3}$$

I also estimate these effects controlling for individual characteristics and geographical location (province dummies).¹⁰ 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 informational 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

¹⁰ According to *Duflo et al. (2007)*, "controlling for variables not affected by the treatment may reduce the variance of the outcome, leading to more precise estimates". Nevertheless, I keep the number of covariates as low as possible in order to minimize the potential bias due to the "covariance between the heterogeneity in the treatment effects and the squares of the included covariates" (*Deaton, 2009*, p. 36).

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 ($\beta_3^S > 0$ and $\beta_2^S = 0$) one will conclude that the spillover effect is the result of social control exerted by the treated individual on the 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 $\beta_2^S > 0$), one cannot separate what part of the change of 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 reduced form specification only allows to identify 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. The delivery of the newspaper consists in a basic information treatment. As a result, I expect that it should primarily affect the information set of treated individuals. Therefore, I also expect that an eventual spillover effect of this treatment it is more likely to be related to information outcomes.

The hotline intervention aims at involving treated individuals in the electoral process, allowing them to contribute to the process. So, I expect treated individuals to feel more empowered, to believe more in the electoral process and to feel more willing to participate. In this case, they will be more likely to influence the preferences of other household members than their information set.

Finally, the civic education intervention, despite including an information component (providing some elementary electoral information via cell phone messages), consists basically in persuading treated individuals to vote. Therefore, it is more likely that individuals that received this treatment may influence the preferences and behavior than the information set of other people.

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, Campbell 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, while information regarding the elections and the voting procedure are more relevant to the decision of whether to vote, 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 to change voting patterns. All the leaflets distributed 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 sent to hotline treated subjects that 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 cohabitate 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.

2.2.2 Structural Form Specification

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

first, the interventions affect the treated individuals; and then these treated individuals influence the other people in the household. *Aker, Collier and Vicente (2011)* provide evidence of the first step. The analysis based in the reduced form specification provides evidence that the effects of the interventions reached the other people in the household, but does not necessarily show that the treated individuals influenced these other people.

In order to directly test if the treated individuals influence the other members of the household, I regress the second interviewee’s outcomes on each outcome of the primary interviewee 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 interventions, I instrument main respondent’s outcomes using the three interventions. The purpose of this exercise is to identify what type of effects on the treated individuals – either on information set, or preferences, or behavior – lead to an effect on the second interviewee, and to analyze into what extent the treated individuals “simplify” the interventions’ effects when they transmit it to the other people in the household. For instance, the interventions may have expanded the treated individuals’ information set, and those individuals, based on that new information, may opt by influencing 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 365 households 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 when using an instrumental approach as a robustness check.

From now on, let $y_{i,l}^P$ and $z_{i,l}^S$ 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 village was selected for the corresponding treatment, newspaper (N_l), hotline (H_l) and civic education (C_l). In the first step, I regress main respondent’s post-treatment outcome of interest ($y_{i,l}^P$) on the interventions (N_l , H_l and C_l); controlling for main respondent’s demographic characteristics ($\mathbf{X}_{i,l}^P$) and geographical location ($\mathbf{G}_l$) (equation 4).

$$y_{i,l}^P = \delta_1 + \beta_n N_l + \beta_h H_l + \beta_c C_l + \gamma_1 \mathbf{X}_{i,l}^P + \tau_1 \mathbf{G}_l + \epsilon_{i,l} \quad (4)$$

I estimate the first step equation using all panel observations (≈ 1150). I estimate it using OLS.

As I intend to analyze the transmission of the interventions' effects by the treated individuals to the other people in the household, I am only interested in outcomes of the treated individuals that were affected by the interventions. Therefore, I exclude the main respondents' outcomes for which all the interventions (N_l , H_l and C_l) failed a significance test (t-test) and also a joint significance test (F test), both of at least 10% significance level.

Using the estimation results of this step, I predict the post-treatment outcomes of interest for all main respondents.¹¹

In the second step, I regress each outcome of the second interviewees ($z_{i,l}^S$) on each predicted outcome of the main respondent living in the same household ($\widehat{y_{i,l}^P}$). In this step I control for the set of controls that were included in the first step ($\mathbf{X}_{i,l}^P$ and $\mathbf{G}_l$) and also for second interviewee's characteristics ($\mathbf{X}_{i,l}^S$).

$$z_{i,l}^S = \delta + \beta \widehat{y_{i,l}^P} + \gamma \mathbf{X}_{i,l}^P + \tau \mathbf{G}_l + \rho \mathbf{X}_{i,l}^S + \epsilon_{i,l} \quad (5)$$

I estimate this equation using OLS. I use Hardin's sandwich cluster-robust covariance estimator to obtain the covariance matrix on which I base my inference (*Hardin, 2002*).

The coefficient β measures the effect of the main respondent's outcome $y_{i,l}^P$ on the second interviewee's outcome $z_{i,l}^S$.

The main caveat of the specification above is the risk that the second interviewee's outcomes are directly affected by the interventions. In such circumstances, the coefficient β will be biased upwards as it will also capture the direct effect of the intervention on the second interviewee's outcome. With the purpose of overcoming this limitation, I run a second test: I estimate a new two-step model in which I control for the direct effect of the interventions on the second interviewees' outcomes. Since I control for the interventions in the regressions of the second interviewee's outcomes, I can no longer rely on the interventions as instruments to estimate the main respondent's outcome. Therefore, I estimate the main respondent's outcomes using as instruments interactions between

¹¹The set of covariates $\mathbf{X}_{i,l}^P$ 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.

the interventions and the set of characteristics that *Aker, Collier and Vicente (2011)* identified as originating heterogeneous responses to the treatments.

In the first step of this new model I regress the post-treatment outcome of the main respondent ($y_{i,l}^P$): on the three interventions (N_l , H_l and C_l); on the set of characteristics that generate heterogeneous effects ($\mathbf{Q}_{i,l}^P$); on interactions between the demeaned equivalent of those characteristics ($\widetilde{q}_{i,l}^P = q_{i,l}^P - \frac{1}{n} \sum_{i=1}^n q_{i,l}^P$, where n is the total sample size) and the corresponding intervention for which they generated heterogeneous effects; on main respondent's demographic characteristics ($\mathbf{X}_{i,l}^P$); and on province dummies ($\mathbf{G}_l$). Adding the interactions guarantees that the main respondent's predicted outcome varies with the treatment but also with the individual's characteristics, avoiding multicollinearity issues in the second step regression. The set of main respondents' characteristics identified by *Aker, Collier and Vicente (2011)* includes gender ($gender_{i,l}^P$), literacy ($lit_{i,l}^P$), whether the individual works at home ($dom_{i,l}^P$), whether the household owns cattle ($cat_{i,l}^P$) and whether the household owns a cell phone ($cell_{i,l}^P$). Based on the conclusions of *Aker, Collier and Vicente (2011)*, I include interaction terms between the newspaper intervention and gender, literacy, domestic worker, cell phone ownership and cattle ownership; I also include interactions between the hotline and cell phone ownership and between civic education and gender (equation 6).

$$\begin{aligned}
y_{i,l}^P = & \delta_1 + \beta_{1n}N_l + \beta_{1h}H_l + \beta_{1c}C_l + \gamma_1\mathbf{X}_{i,l}^P + \theta_1\mathbf{Q}_{i,l}^P + \\
& + \kappa_g \cdot N_l \cdot \widetilde{gender_{i,l}^P} + \kappa_{lit} \cdot N_l \cdot \widetilde{lit_{i,l}^P} + \kappa_{dom} \cdot N_l \cdot \widetilde{dom_{i,l}^P} + \\
& + \kappa_{cell} \cdot N_l \cdot \widetilde{cell_{i,l}^P} + \kappa_{cat} \cdot N_l \cdot \widetilde{cat_{i,l}^P} + \lambda_{cell} \cdot H_l \cdot \widetilde{cell_{i,l}^P} + \\
& + \pi_g \cdot C_l \cdot \widetilde{gender_{i,l}^P} + \tau_1\mathbf{G}_l + \epsilon_{i,l}
\end{aligned} \tag{6}$$

I estimate this regression using all main respondents for which I have panel data (≈ 1150). I estimate it using OLS.

As before, I exclude the main respondents' outcomes for which all the instruments, in this case the interactions with the interventions, failed a significance test (t-test) and also a joint significance test (F test), both of at least 10% significance level.

In the second step, I regress the second interviewee's outcome (z_{il}^S) on the predicted outcome of the main respondent living in the same household ($\widetilde{y_{i,l}^P}$).

In this regression I control for all covariates included in the first step regression with the exception of the interaction terms, and also for second interviewee's demographics characteristics ($\mathbf{X}_{i,l}^S$) (equation 7).

$$z_{i,l}^S = \delta + \mu \widehat{y_{i,l}^P} + \beta_n N_l + \beta_h H_l + \beta_c C_l + \gamma \mathbf{X}_{i,l}^P + \theta \mathbf{Q}_{i,l}^P + \tau \mathbf{G}_l + \rho \mathbf{X}_{i,l}^S + \varepsilon_{i,l} \quad (7)$$

I estimate this equation 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 $y_{i,l}^P$ on the second interviewee's outcome $z_{i,l}^S$ after controlling for the direct effect of the interventions on the second interviewee.

3 Data and Descriptive Statistics

I use data collected in the context of a randomized field experiment conducted during the October 2009 general elections in Mozambique.¹² During this experiment, three different interventions aimed at increasing citizens' political participation 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 design¹³; thirdly, and finally, I present the descriptive statistics of the sample under analysis.

3.1 Treatments¹⁴

The newspaper intervention consisted in the distribution of the free newspaper *A Verdade*, once a week, since the baseline survey until the week after the elections.¹⁵ It was the first time this newspaper was distributed in the treated locations, because until then it was only distributed in the city of Maputo. The

¹²This project was coordinated by Pedro Vicente, Jenny Aker and Paul Collier; and it was funded by the International Growth Centre, a consortium between the LSE and Oxford financed by the Department for International Development (UK), and Trinity College Dublin.

¹³More detailed descriptions can be found in *Aker, Collier and Vicente (2011)*.

¹⁴Based on section "3.1 Treatments" of *Aker, Collier and Vicente, 2011*.

¹⁵All treated individuals received at least one issue, but it was not possible to ensure that all individuals would receive all issues.

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 1*), very similar to the one that the hotline intervention used.

The hotline treatment consisted in the distribution of a leaflet (see *Figure 2*) inviting people to send messages reporting electoral problems that occurred in their locations during the electoral campaign and the elections-day. Then, during the 14 days prior to the elections and during the elections-day, 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 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. First, through the distribution of a leaflet (see *Figure 3*) during the baseline survey, that explained thoroughly each step of the voting procedure.¹⁶ Then, during the 14 days before the elections and during the elections-day experimental subjects were sent cell phone messages (5 a 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.

3.2 Measurement¹⁷

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.¹⁸ Given the importance of the access to

¹⁶This leaflet was made available by the National Electoral Commission.

¹⁷Based on section "3.3 Measurement" of *Aker Collier and Vicente, 2011*.

¹⁸The 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

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.¹⁹ Despite this, during the baseline survey seven locations had to be replaced by the closest pooling station, because when the interviewers' team arrived there they found that there was no cell phone coverage.

Of the 161 polling locations selected for the study, 40 were randomly assigned to each of the three treatments²⁰, and 41 locations serve as the control group, with no treatment administered.²¹ 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, 2011*, p. 13).

Sampling the households in each location followed standard household representativeness, n^{th} calls (*Aker, Collier and Vicente, 2011*, p. 13). In each selected household, the primary interviewee was the household head or his or her partner. Once selected the potential interviewee, the interview, and hence the delivery of the treatment, was only conducted if the respondent had access to a cell phone.²²

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, 11 per experimental location. Note that in each treatment location, of the 11 respondents, on average, 2 did not receive the treatment. The post-elections survey started after the election results were made public, and lasted from No-

(*Aker, Collier and Vicente, 2011*, p. 12).

¹⁹In order to identify all the pooling stations that were more than 5 km away from a cell phone antenna, *Aker, Collier and Vicente (2011)* used data made available by the two cell phone operators (MCell and Vodacom) on the precise geographical location of each of their antennae (p.12).

²⁰During 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.

²¹It 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 one. Therefore there is an extra control location.

²²Note 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.

vember 16 until December 9. During this survey, in each household the enumerators interviewed the experimental subject and a second adult (who did not receive the treatment). The selection of this second interviewee alternated between the spouse, a son, a daughter, other male member of the household and other female member of the household. The post-elections survey included 1154 main interviewees and 518 second interviewees, but only in 365 households it was possible to interview both, the main and a second, interviewees. As the goal of this paper is to study the spillover of these interventions to the people living with treated subjects, the analysis focuses mainly on the data regarding the second person interviewed in the household.

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 the politics (*Aker, Collier and Vicente, 2011*). For logistic reasons, the second interviewees' questionnaire (only conducted in the post-elections survey) was a shorter version of the main respondents' questionnaire. A summary of the second interviewee's outcome variables is presented in *Table Survey Outcomes*. Based on the conceptual framework, I divided these variables into three categories: basic electoral information, interest in elections and turnout. I use a set of indicators of the level of information / awareness about basic electoral elements (e.g. names of presidential candidates, names of parties running for election, etc.) as proxies of the relevant information set for the individual's decision of voting 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, there is not a specific measure of the respondents' preferences towards voting, but there are measures of their interest in the elections, which may be a reasonable approximation of their interest in voting. There is also a measure of the respondents' interest in public matters (politics and government).

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.²³

²³The 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 may be 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

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.²⁴ 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 (2011)*. 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.

3.3 Descriptive Statistics

The descriptive statistics of the demographic characteristics of the sample of second interviewees is displayed in *Table 1*. This table includes two sets of results: the first was computed without sampling weights and the second with sampling weights. Looking at the analysis without sampling weights, we find that in all four groups most of the second interviewees are also women, 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. In the four groups the average household size is between 6 and 7 people. The most common occupation among second interviewees is agriculture. The majority of second interviewees are either the spouse or children 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

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.

²⁴According to the adjusted turnout those respondents who have reported to have voted but answered wrongly the questions regarding the number of ballot papers, the number of ballot boxes and whether the ballot papers had photos or not, were considered as not having voted and, thus, assigned a zero.

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 the daughter of the main respondent. 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.

In the computation of these statistics, as well as in the following empirical analysis, I adjust the standard errors by clustering at local level (enumeration area). The sampling weights for the sample of second interviewees were computed taking into consideration the second interviewee’s representability at the level of the household²⁵ and her representability at the level of her village and province²⁶ (the construction of the sampling weights is described in Appendix).

The descriptive statistics and balance tests of the baseline and post-elections samples of main respondents are presented in *Table A1* in Appendix. At standard significance levels, there are only three differences between treated and control groups in the post-elections sample.²⁷ In the baseline sample there is only one difference. As in the case of second interviewees, it seems that the observables are balanced across treatment and control groups.

One potential weakness of my econometric analysis is related with the attrition in the sample of second interviewees. In several houses a second person could not be interviewed because she was not available at the moment that the house was visited. Unfortunately, if the selection into the sample is correlated with the outcomes of interest, for instance if people who spend more time at home are more likely to be included in the sample and to be influenced by other people in the household, my results may not be generalized to the rest of the population.

²⁵Accounting for the fact that in each household, regardless of the number of potential second interviewees available, only one was selected.

²⁶Accounting for the fact that it was not possible to interview a second person in every household where there was a potential second interviewee (because the particular individual was travelling, or was working far away, etc.).

²⁷First, in the newspaper group there is a higher percentage of individuals who are either married or in an union than in the control group (difference of 7.5 percentage points). Second, in the civic education area there is a lower percentage of people working as artisans than in control areas (difference of 3.4 percentage points). And, third, also in the civic education group there is a higher percentage of public servants than in control areas (difference of 3.2 percentage points).

4 Econometric Results

4.1 Results of Reduced Form Specification

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

4.1.1 Results

In the case of information and interest proxies, I used logit estimation for dummy outcome variables, ordered logit estimation for categorical outcome variables and OLS for quantitative variables. In the case of turnout, to facilitate comparisons across measures I used OLS estimation for all proxies.

Table 2 presents the estimates of the indirect impact of the different interventions on information outcome variables. The only evidence of information sharing in treated households occurs in civic education areas, where I find a positive effect on second interviewee’s knowledge about the concepts of null and blank votes. However, when I estimate the same regressions using sampling weights, this effect no longer exists (*Table A2* in Appendix).

Table 3 displays the estimates of the indirect impact of the interventions on interest outcome variables. I find a positive significant effect on the second interviewees’ interest in provincial assemblies’ elections in the hotline areas. When I estimate the same regressions using sampling weights I also find hotline effects on interest in parliament elections and in public matters (*Table A3* in Appendix). As this increase in interest is not accompanied by an increase in the information set, according to my conceptual model, this constitutes evidence of influence directly at the level of preferences.

The estimates of the indirect effects on individual turnout measures are shown in *Table 4*. All our measures put the average turnout of second interviewees living in control areas above 69 percent (intercept term of the regressions with no controls). As expected, the self-reported turnout gives the highest average, 85.7 percent. Despite the high turnout in the control group, in civic education areas I find a significant positive effect on all measures of voter turnout. This effect varies between 7.0 and 12.6 percentage points. In the hotline group, I also find evidence of a positive indirect effect on the measure interviewers’ assessment (9.7 percentage points when using no controls and 8.0 percentage points when using controls) and on the measure based on the inked finger (10.5 percentage points when using no controls). Most of these effects are corroborated

rated when I estimate the same regressions using sampling weights (*Table A4* in Appendix). This evidence suggests that in civic education areas the transmission of the effects occurs through social control, because there is only evidence of an indirect effect on behavior but no effects on information or preferences. On the other hand, in the hotline group, the increase in individual turnout may be a result, at least partially, of the significant increase in the interest in elections.

These findings show that there is a spillover of the interventions' effects to the people who cohabitate with treated individuals. In addition, and as expected, they show that the three different treatments do not trigger the same type of influence. In households in newspaper areas (information intervention) there is almost no evidence of influence. In households in hotline areas (empowerment campaign) there is evidence of influence directly on interest (proxy of preferences), as well as a change on individual turnout. This last effect may result from the change in preferences and not necessarily from influence directly on behavior. In households in civic education areas (call-out the vote campaign), there is only consistent evidence of influence directly at the level of behavior, thus, influence is exerted through social control. This means 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.

Aker, Collier and Vicente (2011) studied the interventions' effects on experimental subjects. I extend their analysis by focusing on the indirect effects of the interventions on individuals who cohabitate with the experimental subjects. Since *Aker, Collier and Vicente* (from now on ACV) studied the subjects that supposedly exert influence, it is relevant to investigate how my findings relate to the ones that they report.

In newspaper intervention areas, ACV report a significant positive effect on information outcomes (namely, the ability to enumerate the presidential candidates and to explain the concept of abstention) and on individual turnout. Combining their findings with mine, I conclude that 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 hotline areas, ACV found significant positive effects on information outcomes (namely, knowledge of the presidential candidates and ability to name political parties running for the Parliament), on interest about the elections and on individual turnout. The present study shows that in these areas only the effect on interest and, to some extent, on turnout was transmitted to the

non-treated people in the household. Furthermore, in the words of *Aker, Collier and Vicente*, "the hotline was most effective in raising interest and information about the elections" (p. 25). This evidence is consistent with the fact that only in these areas I found an increase in the interest in elections of the second interviewees who lived in treated areas.

In civic education areas, ACV report a positive effect on individual turnout, but no effect on interest about elections and a significant effect only in one of the information outcomes (ability to name the presidential candidates). This evidence is consistent with my finding that civic education has an indirect effect only on turnout. The combined evidence from the two analyses suggests that the civic intervention works mainly through persuasion, it pressures targeted individuals to vote, who in their turn exert pressure on the people around them to also vote.

Contrary to what standard economic models would assume, the transmission of the interventions' effects is not driven by information sharing, but by the 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, mainly contribute 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.²⁸ 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 intervention²⁹), are transmitted to people who cohabitate with treated 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, they 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

²⁸Another exception is the study of *Gine and Mansuri (2011)*, who provide evidence of a field experiment on female voter turnout in Pakistan.

²⁹The civic education intervention is basically a get out the vote campaign, reinforced by cell phone messages.

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

4.1.2 Robustness

In this subsection I examine the robustness of the results presented above.

One main caveat of the empirical analysis is that the effect of the intervention on the non-targeted individuals 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 (114 individuals) reported to have access to the same number that the main respondent; and 23 percent of them (26 individuals) have claimed to own that cell phone. 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 same regressions excluding the second interviewees who claimed to own the cell phone to which the interventions' messages were sent. Using this subsample, all my findings are corroborated and become more consistent. In hotline areas I find significant effects in more interest turnout measures; and in civic education areas I find a positive effect in all turnout measures regardless of whether I use controls or not (*Table A5* in Appendix).

Of the 11 households/main respondents selected in treated villages, on average two did not receive the treatment. However, when estimating the indirect effects of the interventions I considered all the households in treated villages. The reason for this was that the randomization of the interventions is probably more reliable than the random selection of treated households within each village. In order to assess how sensitive are my conclusions to this option, I run

the same regressions excluding the second interviewees who lived in treated areas but in untreated households. All main findings regarding effects on interest and turnout are corroborated (see *Table A6* in Appendix). I find no significant effects on information outcomes.

4.2 Results of Structural Form Specification

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

4.2.1 Results

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 main respondent's outcome variables that passed the significance tests³⁰ in the first step regressions are: three information indicators (knowledge of names of candidates, knowledge of names of parties and knowledge of types of elections), interest in provincial assemblies' elections, and all turnout measures.³¹ The estimated coefficients of the interventions in the first step regressions are presented in *Table A7* in Appendix.³²

Table 5 presents the estimates of coefficient β in equation 5. Each cell of this table corresponds to a different second step regression. The value presented in cell (a, b) - in row a and column b - is the estimated coefficient of main respondent's predicted 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 main respondent's outcomes had a significant effect.

These results show no evidence of information sharing, since none of the main respondents' information outcomes affects second interviewees' information. This corroborates my argument that the transmission of the interventions' effects is not driven by information sharing. Nevertheless, it seems that main

³⁰The significance tests set above were a joint significant F-test of the three interventions and/or at least one the interventions passed the significance t-test.

³¹With the exception of the interest variable, in the regressions of all these outcome variables there was at least one intervention that was significant. In the case of interest in provincial assemblies' elections the interventions passed the joint significance F-test, but none of their coefficients passed the significance t-test.

³²The results displayed in *Table A7* are not exactly identical to the ones reported in *Aker, Collier and Vicente (2011)* because I use a different specification.

respondent’s information affects second interviewee’s interest in elections. This evidence suggests that the main respondent may simplify the message she receives. It may be the case that due to the information she received, she changes her preferences; and then shares with the other people in the household these new preferences, but not the new information.

Regarding the impact of main respondents’ preferences, I find no evidence of sharing of opinions, as main respondents’ interest in elections does not affect second interviewees’ interest.

In terms of behavior, main respondent’s turnout affects second interviewee’s interest in elections and turnout. Second interviewees’ turnout is not affected by any other main respondent’s outcome aside from turnout (information or preferences). This evidence suggests 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.

I believe that this evidence corroborates the conclusions that I presented in the previous section. Using this 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.

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 outcomes and uses the interactions between the interventions and main respondent’s characteristics as instruments.

The estimates of the coefficients of main respondent’s predicted outcomes in the second step regressions, coefficient μ in equation 7, are presented in *Table 6*. As explained for *Table 5*, each cell of this table corresponds to a different second step regression of one of second interviewee’s outcomes on one of main respondent’s predicted outcomes. This table only includes second interviewees’ outcomes in which at least one of main respondent’s predicted outcomes had a significant effect. The estimated coefficients of the interaction terms used as instruments in the first step regression are displayed in *Table A8* in Appendix.³³

³³This table only includes the main respondent’s outcomes in which regressions the interactions passed the joint significance F-test and/or at least one the interactions passed the significance t-test.

The estimates displayed in *Table 6* basically show two types of significant effects: (1) main respondent’s interest affects second interviewee’s turnout; and (2) main respondent’s turnout affects second interviewee’s turnout. So, there is evidence of change in the second interviewee’s behavior while her interest (proxy for preferences) remains unchanged. According to my conceptual model, this constitutes evidence that influence assumes the form of social control.

In short, this evidence supports my main conclusions. There is no evidence of information sharing and the campaigns’ effects seem to be transmitted to second interviewees, at least partially, through social control.

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 results from the combination of the influence effects in the different treatment areas.

4.2.2 Robustness

An alternative to the estimation of the two step model would be to adopt an instrumental variable approach. As a robustness check, I report the 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.

Table 7 reports the estimates of the parameters of interest. Similarly to *Table 5*, 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.

5 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 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 control. So, this paper provides statistical evidence that social control increases turnout. 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.

Contrary to what most economic models would assume, these results show that the transmission of the interventions' effects does not occur through information sharing, but mostly through social control. 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 paper demonstrates that if one accounts for the role of interpersonal influence, the reach of voter education interventions is much greater than the treated population. It would be worth investigating what are the determinants of interpersonal influence, i.e., what makes some people more capable of influencing others. Knowing that would contribute to increasing the reach of this type of interventions, by improving its targeting.

Finally, there may be the possibility that, in the context of democracies where the political forces are largely unbalanced in favor of the party in power, mobilization campaigns that work through persuasion may contribute to strengthen the incumbent position. This should be an object of further research.

References

- [1] Agadjanian, Victor (2001). "Negotiating Through Reproductive Change: Gendered Social Interaction and Fertility Regulation in Mozambique", *Journal of Southern African Studies*, Vol. 27, No. 2, pp. 291-309
- [2] Aldrich, John H. (1993). "Rational Choice and Turnout", *American Journal of Political Science*, Vol. 37, No. 1, pp. 246-278.
- [3] Aker, Jenny, Collier, Paul and Vicente, Pedro C. (2010). "Voter Education Using Cell Phones: Evidence from a Field Experiment in Mozambique".
- [4] Akerlof, George A. (1997). "Social Distance and Social Decisions", *Econometrica*, Vol. 65, No. 5, pp. 1005-1027.
- [5] Arthur, W. Brian and Lane, David A. (1993). "Information Contagion", *Structural Change and Economic Dynamics*, Vol. 4, No. 1, pp. 81-104
- [6] Bartels, Larry M. (2000). "Partisanship and Voting Behavior, 1952-1996", *American Journal of Political Science*, Vol. 44, No. 1, pp. 35-50
- [7] Beck, Paul A., Dalton, Russell J., Green, Steven and Huckfeldt, Robert (2002). "The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices", *American Political Science Review*, Vol. 96, No. 1, pp. 57-73
- [8] Bruhn, M. and D. McKenzie (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments", *American Economic Journal: Applied Economics*, 1(4), 200-32.
- [9] Collier, P., and P. C. Vicente (2009): "Votes and Violence: Evidence from a Field Experiment in Nigeria," Discussion paper, Oxford University and BREAD, Working Paper.
- [10] Davenport, Tiffany. (2010). Public accountability and participation: The effects of a feedback intervention on voter turnout in a low salience election. *Political Behavior*, No. 32, No. 3, pp.337-368
- [11] 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.

- [12] Deutsch, Morton and Harold, B. Gerard (1955). "A Study of Normative and Informational Social Influences Upon Individual Judgement", *Journal of Abnormal & Social Psychology*, Vol. 51, No. 3, pp. 629-636.
- [13] Duflo, Esther, Rachel Glennerster, and Michael Kremer (2007), "Using Randomization in Development Economics Research: A Toolkit", In *Handbook of Development Economics*, eds. T. Paul Schultz, and John Strauss, Vol. 4, pp. 3895–3962, Elsevier.
- [14] Ekine, S. (ed.) (2010). *SMS Uprising: Mobile Activism in Africa*, Oxford: Fahamu Books.
- [15] Engelen, Bart (2006). "Solving the Paradox – The Expressive Rationality of the Decision to Vote", *Rationality and Society*, Vol. 18, No. 4, pp. 419-441.
- [16] Fafchamps, M., and P. C. Vicente (2009): "Political Violence and Social Networks: Experimental Evidence from a Nigerian Election," BREAD, Working Paper 239.
- [17] Feddersen, Timothy J. (2004). "Rational Choice Theory and the Paradox of Not Voting", *The Journal of Economic Perspectives*, Vol. 18, No. 1 (Winter, 2004), pp. 99-112.
- [18] Fiorina, Morris P. (1976). "The Voting Decision: Instrumental and Expressive Aspects", *The Journal of Politics*, Vol. 38, No. 2, pp. 390-413.
- [19] Gine, X., and G. Mansuri (2011): "Together we will : experimental evidence on female voting behavior in Pakistan," World Bank, Policy Research Working Paper 5692.
- [20] Hardin, James W. (2002). "The robust variance estimator for two-stage models", *Stata Journal*, Vol. 2, No. 3, pp. 253-266.
- [21] Heacock, Rebekah (2009). "Mobile Activism In African Elections A Comparative Case Study", DigiActive Research Series
- [22] Hole, Arne R. (2006). "Calculating Murphy-Topel variance estimates in Stata: A simplified procedure", *Stata Journal*, Vol. 6, No. 4, pp. 521-529.
- [23] Huckfeldt, Robert, and Sprague, John (1995). *Citizens, Politics and Social Communication: Information and Influence in an Election Campaign*. New York: Cambridge Univ. Press.

- [24] Knack, Stephen (1992). "Civic Norms, Social Sanctions, and Voter Turnout", *Rationality and Society*, Vol. 4, No. 2, pp. 133-156.
- [25] Lazarsfeld, Paul F., Berleson, Bernard and Gaudet, Hazel (1944). *The People's Choice: How The Voter Makes Up His Mind in a Presidential Campaign*, New York: Duell, Sloan and Pearce.
- [26] Miller, Warren E. (1991). "Party Identification, Realignment, and Party Voting: Back to the Basics", *The American Political Science Review*, Vol. 85, No. 2, pp. 557-568.
- [27] Montgomery, Mark R. and Casterline, John B. (1996). "Social Learning, Social Influence, and New Models of Fertility", *Population and Development Review*, Vol. 22, Supplement: Fertility in the United States: New Patterns, New Theories, pp. 151-175.
- [28] MOZAMBIQUE Food Security Outlook, July to December 2009. Available at http://pdf.usaid.gov/pdf_docs/PNADS377.pdf, Version of 11 February 2011.
- [29] Murphy, Kevin M. and Topel, Robert H. (1985). "Estimation and Inference in Two-Step Econometric Models", *Journal of Business & Economics Statistics*, Vol. 3, No. 4, pp. 370-379.
- [30] Nickerson, David W. (2008). "Is Voting Contagious? Evidence from Two Field Experiments", *American Political Science Review*, Vol. 102, No. 1, pp. 49-57.
- [31] 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*, Vol. 32, No. 3, pp. 369-386
- [32] Stokes, Donald E., Campbell, Angus and Miller, Warren E. (1958). "Components of Electoral Decision", *The American Political Science Review*, Vol. 52, No. 2, pp. 367-387
- [33] Temin, Jonathan and Smith, Daniel A. (2002). "Media Matters: Evaluating the Role of the Media in Ghana's 2000 Elections", *African Affairs*, No. 101, pp. 585-605

- [34] Vicente, Pedro and Wantchekon, Leonard (2009). "Clientelism and vote buying: lessons from field experiments in African elections", *Oxford Review of Economic Policy*, Vol. 25, NO. 2, pp. 292–305
- [35] Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*, MIT Pres.

Figure 2: Hotline leaflet

MELHORAMOS AS ELEIÇÕES!

REPORTE PROBLEMAS DURANTE **A CAMPANHA ELEITORAL**



ENVIE MENSAGENS SMS FORMATO

LOCAL *espaço* PROBLEMA

POR EXEMPLO

“EPC Quelimane distúrbios no comício”

PARA

82 1112 OU 84 13333



APÓIO:

vodacom

mcel

OBSERVATÓRIO ELEITORAL

AMODE, CCM, CEDE, CEM, CISLAMO, FECIV, LDH, OREC

AMODE (Associação Moçambicana para o Desenvolvimento) CCM (Conselho Cristão de Moçambique) Comissão Episcopal de Justiça e Paz da Igreja Católica LDH (Liga Moçambicana dos Direitos Humanos) CEDE (Centro de Estudos de Democracia e Desenvolvimento) CISLAMO (Conselho Islâmico de Moçambique) FECIV (Instituto de Educação Cívica) OREC (Organização para Resolução de Conflitos)

MELHORAMOS AS ELEIÇÕES!

REPORTE PROBLEMAS DURANTE **O DIA DAS ELEIÇÕES**



ENVIE MENSAGENS SMS FORMATO

LOCAL *espaço* PROBLEMA

POR EXEMPLO

“EPC Quelimane estação de voto mudou”

PARA

82 1112 OU 84 13333



APÓIO:

vodacom

mcel

OBSERVATÓRIO ELEITORAL

AMODE, CCM, CEDE, CEM, CISLAMO, FECIV, LDH, OREC

AMODE (Associação Moçambicana para o Desenvolvimento) CCM (Conselho Cristão de Moçambique) Comissão Episcopal de Justiça e Paz da Igreja Católica LDH (Liga Moçambicana dos Direitos Humanos) CEDE (Centro de Estudos de Democracia e Desenvolvimento) CISLAMO (Conselho Islâmico de Moçambique) FECIV (Instituto de Educação Cívica) OREC (Organização para Resolução de Conflitos)

Figure 3: Civic education leaflet by CNE/STAE

1 só dia

28
Outubro 2009





Por Eleições Livres, justas e transparentes



PASSOS DE Votação

4^{as} ELEIÇÕES GERAIS (Presidenciais e Legislativas) e
1^{as} ELEIÇÕES para as Assembleias Provinciais

A votação é feita da seguinte forma:

- O eleitor ao apresentar-se perante a Mesa da Assembleia de Voto mostra as suas mãos aos MMV e entrega o seu cartão de eleitor.
- Reconhecido o eleitor e verificada a sua inscrição, o presidente da mesa entrega-lhe três boletins de voto (um para eleição do Presidente da República, outro dos deputados da Assembleia da República e outro para eleição dos Membros da Assembleia Provincial).
- De seguida o eleitor dirige-se à cabine de voto e sozinho, assinala com uma cruz ou com a impressão digital, caso não saiba escrever, no quadrado correspondente à candidatura em que vota e, por fim, dobra o boletim em quatro partes.
- Perante a mesa, o eleitor deposita cada boletim na uma correspondente; (Cada uma contém uma tampa com uma cor diferente da outra, com indicação PR¹ Presidente da República, AR² Assembleia da República e AP³ Assembleia Provincial, respectivamente). O 1^o escrutinador confirma a votação, rubricando o caderno de recenseamento eleitoral na coluna própria e no espaço correspondente ao nome do eleitor.
- Depois de exercer o seu direito de voto, o eleitor mergulha o dedo indicador direito em tinta especial (esta tinta não desaparece durante vários dias).
- Finalmente, o eleitor recebe o seu Cartão de Eleitor e retira-se da Assembleia de Voto.

NB: Se o eleitor inadvertidamente inutilizar um boletim de voto, deverá pedir outro ao presidente da mesa, devolvendo-lhe o inutilizado.

Table Survey Outcomes

Measure	Indicators	Description / Phrasing of the question	Range / Scale
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 - 5
	knows what elections took place	Do you know what elections took place in October 2009?	0 - 1
	null vote	Do you know what "null vote" means?	1 - 3
	blank vote	Do you know what "blank vote" means?	1 - 3
	electoral abstention	Do you know what "electoral abstention" means?	1 - 3
	secret vote	Do you believe that someone else can know for whom you voted for?	1 - 3
	presidential mandate	How long is the presidential mandate?	0 - 1
	interest in presidential elections	With how much interest did you follow the 2009 presidential elections?	1 - 4
	interest in parliamentary elections	With how much interest did you follow the 2009 parliamentary elections?	1 - 4
Interest in elections	interest in provincial assemblies' elections	With how much interest did you follow the 2009 provincial assemblies' elections?	1 - 4
	interest in public matters	With how much interest do you follow public matters?	1 - 4
	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.	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, ballot boxes and whether the ballot papers had photos or not.	0 - 1
Turnout	Based on ink finger	What finger was inked after voting?	0 - 1
	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

Table 1a: Differences across treatment and control areas - Sample of second respondents

		Without sampling weights				With sampling weights			
		Control	Civic Education	Hotline	Newspaper	Control	Civic Education	Hotline	Newspaper
Basic Demographics	female	0.623	-0.068 (0.061)	-0.035 (0.060)	-0.053 (0.062)	0.546	-0.008 (0.096)	0.116 (0.084)	0.131 (0.092)
	age	31.520	-0.004 (1.732)	0.756 (1.712)	-0.981 (1.756)	30.996	1.175 (2.736)	2.281 (2.564)	-1.070 (2.049)
	single	0.443	0.089 (0.062)	-0.122** (0.059)	0.020 (0.063)	0.432	0.009 (0.097)	-0.163* (0.086)	-0.016 (0.091)
	married or in a union	0.504	-0.082 (0.062)	0.124** (0.060)	-0.033 (0.063)	0.519	-0.020 (0.092)	0.161* (0.086)	0.002 (0.087)
	no school	0.244	-0.049 (0.052)	-0.076 (0.049)	-0.023 (0.053)	0.131	0.053 (0.078)	0.012 (0.059)	0.066 (0.071)
	primary school	0.221	0.021 (0.053)	0.049 (0.053)	0.033 (0.054)	0.215	0.065 (0.076)	0.081 (0.080)	0.040 (0.075)
	incomplete secondary school	0.244	-0.033 (0.052)	-0.033 (0.051)	0.010 (0.055)	0.268	-0.030 (0.077)	-0.098 (0.069)	0.030 (0.083)
	complete secondary school	0.061	0.048 (0.035)	0.005 (0.030)	-0.020 (0.028)	0.155	-0.063 (0.065)	-0.076 (0.068)	-0.137** (0.056)
	household head	0.740	0.041 (0.053)	0.041 (0.052)	0.022 (0.055)	0.687	0.094 (0.081)	0.054 (0.087)	0.084 (0.088)
	household size	6.271	0.389 (0.338)	0.298 (0.312)	0.360 (0.340)	7.601	0.545 (0.653)	-0.040 (0.442)	0.224 (0.528)
Relationship to primary interviewee	spouse	0.473	-0.060 (0.062)	0.027 (0.062)	-0.131** (0.063)	0.396	-0.077 (0.084)	0.054 (0.080)	-0.116 (0.084)
	daughter	0.156	-0.093** (0.039)	-0.053 (0.041)	-0.037 (0.044)	0.134	-0.089** (0.041)	-0.049 (0.044)	0.002 (0.054)
	son	0.086	0.057 (0.040)	0.047 (0.039)	0.025 (0.038)	0.087	0.102* (0.061)	0.017 (0.041)	0.009 (0.044)

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 1b: Differences across treatment and control areas - Sample of second respondents

	Without sampling weights				With sampling weights			
	Control	Civic Education	Hotline	Newspaper	Control	Civic Education	Hotline	Newspaper
Occupation	agriculture	0.323 (0.059)	-0.017 (0.057)	-0.009 (0.059)	0.250	0.026 (0.087)	0.043 (0.089)	0.140 (0.092)
	retail informal sector	0.046 (0.023)	-0.022 (0.023)	0.006 (0.027)	0.083 (0.026)	-0.074* (0.038)	-0.019 (0.045)	-0.062 (0.039)
	artisan	0.023 (0.016)	-0.015 (0.016)	-0.008 (0.017)	0.005 (0.016)	-0.004 (0.004)	0.004 (0.008)	-0.004 (0.004)
	unskilled worker	0.015 (0.014)	-0.007 (0.017)	0.007 (0.019)	0.023 (0.019)	-0.016 (0.020)	0.006 (0.025)	0.013 (0.027)
	wage worker	0.031 (0.024)	0.018 (0.022)	0.006 (0.019)	0.025 (0.019)	0.024 (0.030)	0.012 (0.027)	-0.023 (0.016)
	teacher	0.023 (0.019)	0.001 (0.022)	0.010 (0.021)	0.028 (0.021)	-0.001 (0.023)	0.013 (0.027)	0.011 (0.032)
	public servant	0.038 (0.019)	-0.030 (0.025)	0.006 (0.021)	0.049 (0.021)	-0.039 (0.032)	-0.020 (0.033)	-0.016 (0.041)
	student	0.131 (0.041)	-0.018 (0.042)	0.010 (0.046)	0.153 (0.046)	-0.059 (0.063)	-0.029 (0.067)	0.025 (0.073)
	domestic	0.123 (0.041)	-0.002 (0.036)	0.001 (0.042)	0.076 (0.042)	0.024 (0.046)	0.016 (0.054)	0.054 (0.057)
	house	0.893 (0.041)	-0.034 (0.038)	0.005 (0.043)	0.869 (0.043)	0.017 (0.066)	-0.002 (0.068)	-0.007 (0.066)
Assets and expenditure	land	0.656 (0.060)	-0.063 (0.058)	-0.058 (0.061)	0.616 (0.061)	0.053 (0.092)	0.086 (0.088)	0.002 (0.094)
	cattle	0.260 (0.054)	-0.025 (0.055)	0.040 (0.055)	0.275 (0.055)	0.031 (0.082)	0.106 (0.082)	0.074 (0.092)
	cell phone	0.679 (0.058)	0.024 (0.055)	0.072 (0.058)	0.783 (0.058)	-0.013 (0.076)	0.048 (0.077)	-0.031 (0.087)
	expenditure	112.460 (12.563)	-5.247 (16.935)	33.937* (18.446)	149.646 (18.446)	-37.650* (19.346)	11.475 (29.408)	13.805 (30.018)

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: Indirect effects on information outcomes

	Names of candidates	OLS										Concept of blank vote				Concept of abstention		Belief the vote is secret		Duration of pres. mandate	
		Names of parties		Types of elections		Concept of null vote		Concept of blank vote		Concept of abstention		Belief the vote is secret		Duration of pres. mandate							
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)				
Newspaper	-0.059 (0.107)	-0.066 (0.095)	-0.067 (0.219)	-0.157 (0.172)	0.077 (0.158)	-0.013 (0.110)	-0.108 (0.286)	-0.208 (0.272)	-0.125 (0.314)	-0.077 (0.293)	0.235 (0.288)	0.305 (0.327)	-0.272 (0.349)	-0.529 (0.370)	-0.266 (0.326)	-0.076 (0.338)					
Hotline	0.015 (0.094)	0.025 (0.082)	0.006 (0.228)	0.052 (0.179)	0.095 (0.151)	0.024 (0.106)	0.047 (0.273)	0.239 (0.257)	0.062 (0.283)	0.331 (0.263)	0.153 (0.327)	0.260 (0.353)	0.495 (0.378)	0.409 (0.396)	-0.226 (0.305)	-0.294 (0.310)					
Civic education	-0.008 (0.088)	-0.007 (0.084)	-0.011 (0.195)	0.064 (0.174)	0.137 (0.144)	0.080 (0.119)	0.182 (0.245)	0.416* (0.243)	0.070 (0.273)	0.414* (0.247)	-0.052 (0.313)	-0.014 (0.343)	0.121 (0.391)	0.150 (0.416)	0.015 (0.300)	0.181 (0.323)					
Adjusted R-squared	-0.004	0.083	-0.005	0.212	-0.003	0.218	0.002	0.082	0.001	0.083	0.002	0.110	0.009	0.036	0.003	0.070					
No. of observations	518	462	518	462	518	462	518	462	518	462	518	462	518	462	518	462					
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes					

Note: Regressions include observations for second respondents living in treated and control areas. Regressions of number of presidential candidates, number of parties running for Parliament and number of elections are OLS. Regressions of knowledge about duration of presidential mandate and number of presidents already elected are logit. The regressions of the remainder indicators are ordered logit. In the second column for each outcome we control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table 3: Indirect effects on interest outcomes

	Presidential elections		Parliamentary elections		Provincial assemblies		Public matters	
	Ordered logit							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Newspaper	-0.251 (0.303)	-0.080 (0.290)	-0.149 (0.304)	0.052 (0.291)	-0.141 (0.298)	0.077 (0.294)	0.063 (0.260)	0.130 (0.249)
Hotline	0.169 (0.308)	0.301 (0.317)	0.369 (0.313)	0.545 (0.331)	0.404 (0.314)	0.703** (0.310)	0.074 (0.218)	0.102 (0.227)
Civic education	-0.089 (0.333)	0.005 (0.312)	0.204 (0.322)	0.321 (0.304)	-0.015 (0.317)	0.247 (0.277)	0.141 (0.265)	0.181 (0.250)
Adjusted R-squared	0.003	0.080	0.005	0.102	0.005	0.113	0.000	0.056
No. of observations	515	460	504	452	475	426	513	459
Controls	No	Yes	Yes	No	No	Yes	Yes	No

Note: Regressions include observations for second respondents living in treated and control areas. All regressions are ordered logit. In the second column for each outcome we control for demographic characteristics and province dummies. Standard errors reported; these are corrected by clustering at the enumeration area level.

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

Table 4: Indirect effects on turnout measures

	Self-reported		Adjusted self-reported		Finger		Index 1		Index 2		Interviewer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Newspaper	0.005 (0.052)	-0.007 (0.055)	0.061 (0.064)	0.065 (0.064)	0.029 (0.055)	0.011 (0.060)	0.014 (0.049)	0.002 (0.051)	0.029 (0.048)	0.015 (0.050)	0.047 (0.048)	0.046 (0.053)
Hotline	0.059 (0.042)	0.045 (0.044)	0.066 (0.056)	0.068 (0.054)	0.105** (0.047)	0.073 (0.049)	0.052 (0.039)	0.037 (0.041)	0.057 (0.040)	0.045 (0.041)	0.097** (0.044)	0.080* (0.045)
Civic education	0.066 (0.043)	0.080* (0.044)	0.089 (0.063)	0.126** (0.061)	0.099** (0.050)	0.090* (0.053)	0.061 (0.041)	0.078* (0.041)	0.070* (0.042)	0.091** (0.042)	0.106** (0.043)	0.111*** (0.041)
_const	0.857*** (0.037)	0.756*** (0.073)	0.697*** (0.045)	0.567*** (0.092)	0.773*** (0.037)	0.734*** (0.080)	0.799*** (0.034)	0.745*** (0.067)	0.766*** (0.034)	0.719*** (0.068)	0.721*** (0.033)	0.573*** (0.070)
Adjusted R-squared	0.003	0.040	-0.001	0.056	0.008	0.024	0.001	0.042	0.002	0.050	0.010	0.047
No. of observations	483	431	483	431	483	431	483	431	483	431	483	431
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Note: Regressions include observations for second respondents living in treated and control areas. All regressions are OLS. 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 1%; ** significant at 5%; * significant at 10%

Table 5: Model with interventions as instruments - Second step estimation results

Panel's predicted outcome	Second interviewee's outcomes		
	Interest		Turnout
	Parliamentary elections	Provincial assemblies	Interviewer
	(1)	(2)	(3)
Names of presidential candidates that knows	1.726 (1.529)	2.496 (1.842)	0.569 (0.527)
Names of parties that knows	1.137 (0.759)	1.634* (0.975)	0.248 (0.280)
Knows types of elections	1.125* (0.593)	1.420** (0.697)	0.333 (0.211)
Interest in provincial assemblies' elections	0.982 (0.618)	1.177 (0.759)	0.205 (0.212)
Self-reported turnout	3.752* (1.959)	4.934** (2.308)	1.144* (0.592)
Adjusted self-reported turnout	2.286* (1.196)	2.993** (1.393)	0.706* (0.410)
Finger measure	3.933 (2.519)	4.597 (2.829)	1.486** (0.738)
Index 1	3.969* (2.148)	5.180** (2.478)	1.259** (0.640)
Index 2	3.675* (2.060)	4.704** (2.332)	1.228** (0.598)
Interviewer assessment	2.945* (1.745)	3.733* (1.940)	1.037** (0.513)

Note: Each cell corresponds to a linear model regression of an outcome variable of the second interviewee on a predicted main respondent's outcome variable. The values displayed are the coefficients of predicted main respondent's outcome variables (displayed in rows) on the regressions 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 1%; ** significant at 5%; * significant at 10%

Table 6: Model with interactions as instruments - Second step estimation results

Panel's predicted outcome	Second interviewee's outcomes					
	Information	Turnout				
	Concept of blank vote	Self-reported	Finger	Index 1	Index 2	Interviewer
	(1)	(2)	(3)	(4)	(5)	(6)
Knows types of elections	0.570 (1.102)	0.034 (0.325)	-0.046 (0.322)	-0.043 (0.294)	-0.086 (0.283)	0.192 (0.263)
Knows concept of null vote	1.016 (0.668)	0.058 (0.215)	0.124 (0.249)	0.030 (0.209)	0.019 (0.201)	0.146 (0.176)
Interest in presidential elections	1.252 (0.777)	0.327 (0.245)	0.424 (0.274)	0.323 (0.226)	0.304 (0.215)	0.164 (0.214)
Interest in parliament elections	1.288** (0.590)	0.415* (0.224)	0.509** (0.259)	0.406* (0.209)	0.379* (0.200)	0.253 (0.202)
Interest in provincial assemblies' elections	1.145 (0.789)	0.425* (0.259)	0.509* (0.284)	0.428* (0.238)	0.413* (0.223)	0.253 (0.222)
Interest in public matters	0.960 (0.950)	0.293 (0.220)	0.310 (0.235)	0.314 (0.193)	0.311* (0.183)	0.075 (0.187)
Self-reported turnout	2.426 (1.531)	0.569 (0.493)	0.601 (0.507)	0.496 (0.468)	0.396 (0.473)	0.646 (0.430)
Adjusted self-reported turnout	1.418 (1.180)	0.586 (0.372)	0.620 (0.414)	0.522 (0.356)	0.444 (0.357)	0.594* (0.338)
Finger measure	2.549 (1.893)	0.481 (0.625)	0.547 (0.602)	0.369 (0.610)	0.272 (0.603)	0.693 (0.487)
Index 1	2.141 (1.494)	0.577 (0.479)	0.622 (0.493)	0.489 (0.457)	0.388 (0.461)	0.679 (0.426)
Index 2	1.741 (1.342)	0.584 (0.447)	0.633 (0.471)	0.486 (0.429)	0.391 (0.429)	0.673* (0.407)
Interviewer assessment	1.702 (1.231)	0.515 (0.404)	0.525 (0.411)	0.410 (0.383)	0.319 (0.379)	0.569 (0.352)

Note: Each cell corresponds to a linear model regression of an outcome variable of the second interviewee on a predicted main respondent's outcome variable. The values displayed are the coefficients of predicted main respondent's outcome variables (displayed in rows) on the regressions 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 1%; ** significant at 5%; * significant at 10%

Table 7: Instrumental variable approach - Second step estimation results

Panel's predicted outcome	Second interviewee's outcomes		
	Interest		Turnout
	Parliamentary elections	Provincial assemblies	Interviewer
	(1)	(2)	(3)
Knows types of elections	0.971* (0.538)	0.946* (0.531)	0.299 (0.217)
Knows concept of null vote	0.595 (0.426)	0.914* (0.498)	0.176 (0.133)
Knows concept of blank vote	1.037** (0.503)	0.685* (0.354)	0.343 (0.213)
Self-reported turnout	1.756 (1.232)	2.437* (1.411)	0.854* (0.457)
Adjusted self-reported turnout	1.056 (0.834)	1.080 (0.791)	0.474* (0.286)
Finger measure	2.191 (1.530)	2.394 (1.669)	1.048 (0.692)
Index 1	1.874 (1.472)	2.321 (1.466)	0.959* (0.556)
Index 2	1.779 (1.375)	1.928 (1.339)	0.902* (0.523)
Interviewer assessment	1.450 (0.902)	1.631 (1.043)	0.602* (0.323)

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) on the regressions of second interviewee's outcome variables (displayed in columns). Standard errors reported; these are corrected by clustering at the enumeration area level.

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

A Appendix: Construction of Sampling Weights

The sampling weights used in the analysis were computed in order to adjust for, first, unequal probability of selection in each household; and second, for total non-response.

In each household it was selected only one second interviewee regardless of the number of potential adults available; this means that an individual living in a household with few adults had a higher probability of being selected than an individual living in a household with a large number of adults. Therefore, one needs to compensate for these unequal probabilities of selection into the sample of second interviewees. Ideally, each second interviewee observation should be weighted by the number of adults in his or her household, excluding the treated individual (who, by definition, could not be selected into the second interviewees sample). In the absence of information about the number of adults living in the household, I used the total number of household members (including children) to adjust for the unequal probability of selection in each household. I am assuming that the number of adults is proportional to the total number of people in the household.

Unfortunately, it was not possible to interview a second person in all houses where there was potential adults available. The main reasons for that were that the people or were in the "machamba" (agricultural plot of land generally located several kilometers away from the family's home)¹, or had temporarily migrated to some other part of the country, or where at work and, for logistic reasons, the interviewer could not wait for them. To compensate for this total non-response I computed non-response adjustments weights. These weights are equal to the product of the reciprocal of the weighted response rate at the village level (the number of weighted second interviewees conducted in the village divided by the number of potential second interviewees available) by the reciprocal of the weighted response rate at the province level (the sum of potential second interviewees in villages where at least one second interview was conducted in that province, divided by the total number of potential second interviewees in the province).

The final sampling weight consists in the size of the household (weight for unequal probability of selection) multiplied by the reciprocal of the village and province response rates (non-response adjustment weight).

In mathematical terms:

$$w_{ilp} = w_{ilp}^s \times w_{ilp}^r$$

The sampling weight for individual i living in location l , in province p , is the product of the weight for his or her unequal probability of selection in the household (w_{ilp}^s) by the non-response adjustment weight (w_{ilp}^r).

¹As the pos-elections survey took place in the beginning of the rainy season the number of people in the "machambas" was significant, as most of the planting work takes place during that period (Seasonal Calendar in Mozambique Food Security Outlook, July to December 2009, available in http://pdf.usaid.gov/pdf_docs/PNADS377.pdf).

$$w_{ilp}^s = h_{ilp}$$

The weight for unequal probability of selection of each individual is the size of his or her household (h_{ilp}).

$$w_{ilp}^r = (r_l)^{-1} \times (r_p)^{-1}$$

The non-response adjustment weight is the product of the reciprocals of the response rates in the local l (r_l) and in the province p (r_p).

Let N_l represent the set of households in location l that were visited during the pre-elections survey and that include other adults besides the individual that was interviewed during the pre-elections survey; and $\bar{N}_l$ represent the set of households in location l where a second person, besides the individual who was interviewed during the pre-elections survey, was interviewed. The response rate in location l is equal to the sum of second interviews that were conducted in that location, weighted by the size of their households, divided by the total number of potential second interviewees existent in that location.

$$r_l = \frac{\sum (w_{ilp}^s \times 1_{\bar{N}_l})}{\sum (w_{ilp}^s \times 1_{N_l})}$$

Let L_p represent the set of locations in province p where the pre-elections survey was conducted; and $\bar{L}_p$ represent the set of locations in province p where at least in one household a second person, besides the interviewee in the pre-elections survey, was interviewed. The response rate in province p is equal to the sum of potential second interviewees in villages where at least one second interview was conducted divided by the total number of potential second interviewees in the province.

$$r_p = \frac{\sum (w_{ilp}^s \times 1_{N_l} \times \bar{L}_p)}{\sum (w_{ilp}^s \times 1_{N_l} \times L_p)}$$