

Pre-analysis Plan: Voter Mobilization in Sierra Leone

20 February 2019

K. Grieco, N.F. Meriggi and M. Voors

Registration timeline and Ethnical Approval

This pre-analysis plan is filed after treatment assignment and data collection, but before analysis. The most recent general election in Sierra Leone, the setting of this PAP, took place on March 7, 2018. This data is held on a secure server, where study field managers and one PI (Grieco) had access for monitoring data quality. Other PIs will be invited to access data as soon as PAP is lodged. We have received IRB clearance from the Office of the Sierra Leone Ethics and Scientific Review Committee (02/27/2018) and UCLA (4/2/2018, IRB#18-000401) for this trial.

Overview

This project evaluates the effectiveness of text message voter mobilization strategies in Africa. In the runup to the 2018 general election in Sierra Leone, the National Election Commission (NEC) initiated a voter mobilization campaign with the intention of increasing voter turnout. We evaluate a) the effectiveness of SMS reminders to vote and b) the relative effectiveness of development-oriented mobilization messages. We (block) randomize across 1293 polling stations – here referred to as *voter registrations centers* (VRCs)— throughout Sierra Leone.

This study makes three main contributions. First, we provide experimental estimates of the effects of a cost-effective SMS based voter-mobilization campaign in Africa. Second, we test if messages that motivate *why* voters should turnout, by linking voting to local or national development, are more effective than messages that simply remind voters to turnout. Third, we attempt to untangle if the turnout effects of development-oriented messages are driven by identity or material-based motivations.

Primary Hypotheses: Using SMS to increase turnout

Reminding Voters

Highly noticeable and salient reminders— such as SMS— can mobilize voters even when these reminders are impersonal (Dale and Strauss 2009; Malhotra et al. 2011). While there is a substantial “get-out-the-vote” literature, it is heavily focused on U.S elections. At the time of this PAP, we are aware of only two studies that evaluate the effect of SMS based voter-mobilization campaigns in Africa. Aker et al. (2017) combine SMS with door to door civic education campaign in the 2009 elections in Mozambique. They find an increase of about 5 percentage points in voter turnout. However, as SMS are bundled with canvassing this making it difficult to isolate the effects of impersonal SMS messages. If SMS alone are sufficient to mobilize voters, then this may be a cost-effective strategy to get people to come vote. Closest to our current study is Marx et al. (2017) who study the effects of SMS sent by the Kenyan Electoral Commission (IECB) in the days leading up the 2013 national election. They find small

(0.3 percentage points) but significant increases in turnout. Unfortunately, the Kenyan Electoral Commission suffered from equipment failure and questions of electoral legitimacy, potentially reducing people's trust in the institution and the effect of mobilization messages¹

- **Hypothesis 1:** SMS reminders increase voter turnout.

Motivating Voters: SMS about why voting matters

Beyond simply reminding registered voters to vote, appeals to *why* citizens should vote may be effective. Using a lab-in-the-field design Brinkley et al. (2016) find that motivational messages about why voting is important increases self-reported intention to register and to vote among South African youth.² We study the effects of motivational messages that link voting to either national or local development. We hypothesize that registered voters are more likely to turn out when reminded of this important reason to vote— namely the importance of selecting candidate that can bring development.

- **Hypothesis 2:** Development oriented SMS are more effective at increasing voter turnout than simple reminder messages

Untangling Effects: Identity

We attempt to untangle why appeals to development might increase voter turnout. Appeals to development may be effective because they activate components of an individual's identity. For example, when nationalist identity is high, people may be more motivated to turn out by appeals to the importance of voting for national development. In the same manner, people with strong local identity— in the context of Sierra Leone, ethnic identity— may be motivated to turn out by appeals to local development.

- **Hypothesis 3a:** Appeals to National Development are more effective at turning out voters with high levels of nationalist identity.
- **Hypothesis 3b:** Appeals to Local Development are more effective at turning out voters with higher levels of local (ethnic) identity.

Untangling Effects, Material Benefits

One the other hand, it could be that voters respond to development-oriented appeals to vote because it increases the salience of potential material gains that comes with development. For example, appeals to local development matter because voters are reminded of potential development projects that could materially impact their lives. By this logic, voters are more likely to be motivated to vote by appeals to local development if they live in the area (ward) where they vote, as they are more materially affected by who is elected to political office.

¹ “Unfortunately, the IEBC encountered numerous problems (including equipment failures) which led it to abandon its plan for a fully electronic tallying of the results in favor of a manual counting that seemed to lack transparency. In addition, some areas experienced outbursts of election-related violence”. (Marx et al, 2007: 2)

² These are the results of a large pilot study. The authors plan to expand this work into a field experiment targeting 6,000 youth in urban Gauteng, South Africa for the 2019 general elections.

Hypothesis 4: Appeals to Local Development are more effective at turning out voters who live in the ward where they vote.

Additional Analysis: Efficient Mobilization

How should mobilization campaigns allocate SMS messages across voting centers to achieve maximum returns to turnout? In most of rural Africa, voters live in dense social networks with high degrees of information sharing. In this context there may be diminishing turnout returns to sending additional text messages to voters registered at a given voting center.

An ideal way to gain insight into the relationship between SMS saturation rates and voter turnout would be to experimentally manipulate SMS saturation rates across voting centers. As a next-best approach we consider two proxies for SMS saturation rates: a) cell phone network coverage and b) percent of registered voters in a voting center that were sent SMS. We will describe and graph the relationship between treatment effects and these two variables.³

Intervention Description

In the weeks preceding the election, the National Election Commission (NEC) approached the International Growth Centre (IGC) in Freetown with plans for a voter mobilization campaign. NEC requested support from the IGC to implement and evaluate the campaign. NEC worked with the research team to develop mobilization messages.

The study area consists of 1293 VRCs across 52 constituencies in Sierra Leone (drawn randomly from a total of 3276 VRCs and 132 constituencies, about 40%). Each VRC was randomly assigned to one of four experimental conditions. Messages were sent out by the NEC as part of the GoVote campaign in the two days prior to election day. The first round of SMS was sent on the evening of 5 March 2018 and the second round was sent out on the evening of 6 March. Elections took place on 7 March.

Table 1 presents the message that voters received in each treatment group. The sender for all messages was “NEC GoVote”.

Table 1: SMS messages received by voters, by treatment group

Experimental Arm	Message Text	Sample (# VRCs)
Control	[No Message]	323
Treatment 1: Reminder	<i>On March 7 vote at [VRC Number and Polling Station Number]</i>	323
Treatment 2: Reminder + Local Political Info	<i>Your Local Councilor is important for developing your ward. Vote for the candidate who can bring development to your community! On March 7 vote at [VRC Number and Polling Station Number]</i>	324
Treatment 3: Reminder + National Political Info	<i>The President is important for developing your country. Vote for the candidate who can bring development to Sierra Leone! On March 7 vote at [VRC Number and Polling Station Number]</i>	323
		1293

³ We present these relationships as descriptive analysis and not hypothesis testing.

On 7 March 2018 voters in Sierra Leone voted on candidates for five political positions:

- President
- Member of Parliament (MP)
- District Council Chairperson
- Local Council
- Mayor

The top two presidential candidates moved to a runoff that was held on 31 March 2018; the highest polling candidate at the remaining four positions is elected. Mobilization messages were sent preceding the first election phase, but not before the presidential runoff

Randomization

VRCs were randomly allocated to one of four experimental conditions using blocked randomization (Appendix C). VRCs were blocked for randomization based on three variables:

- Total number of registered voters at VRC
- Percent of registered voters with listed cell phone number
- Percent of register voters who are male

BALANCE TABLE

Table 2: Balance

	T1	T2	T3	T4	f_stat	pval
VRC Population	987.148	993.625	996.383	982.384	0.031	0.993
VRC Phone Share	0.347	0.356	0.352	0.345	0.093	0.964
VRC Male Share	0.464	0.464	0.464	0.463	0.021	0.996

Table 2 reports means across four treatment arms for our three blocking variables: Population in VRC, Percent of Registered voters at VRC with a listed phone number, and percent of registered voters who are male. Table 2 also reports the F statistic testing the null hypothesis of equality of means across treatment conditions and the associated p-value. All p values are near 1 and we cannot reject the null that our blocking covariates are balanced across treatment arms.

Data

We use four sources of data:

- A. *Voter Registration Data* – We use this data to create a measure of *vrc turnout percent* at each polling station, our primary dependent variable. This data also contains information on occupation, age, gender, listed cell phone number, and the current residence of voters. We used this data for blocking and will use it for heterogeneous effects analysis.
- B. *Official Voting Data* – Provides polling station level data on voting outcomes. This provides the numerator value, valid votes, for vrc turnout percent.
- C. *Cell Phone Network Coverage* – Provides network coverage for Orange, a UK based telecommunications company operating in Sierra Leone. This data maps coverage area polygons. We use this data to determine if voting stations are in network coverage area.
- D. *Afrobarometer* – We use question 88B on the 6th round of Afrobarometer survey in Sierra Leone as a measure for national vs local identity. The responses are geo-coded. More details on variable construction can be found in Appendix A.

Analysis

Our estimand is the average treatment effect (ATE), ie the difference in voter turnout (at VRC) between treatment arms.

To test **Hypothesis 1**, we estimate:

$$Y_{vrc} = \alpha_0 + \alpha_1(T1_{vrc} + T2_{vrc} + T3_{vrc}) + \theta X_j + \varepsilon_{vrc} \quad (1)$$

where Y_{vrc} is the percent voter turnout at the VRC level; $T1$ is the simple SMS reminder, $T2$ is Local Development treatment and $T3$ is National Development treatment; X_j vector of block dummies; ε_{vrc} is the usual idiosyncratic error term. α_1 is the estimate for the combined effect of any of the three SMS treatments, ie the effect on voting turnout in a VRC where some citizens received a SMS reminder. Hypothesis 1 predicts the rejection of this null hypothesis that $\beta_1 = 0$, and that turnout is higher in VRCs that received SMS messages.

To test **Hypothesis 2**, we estimate:

$$Y_{vrc} = \beta_0 + \beta_1 T1_{vrc} + \beta_2 (T2_{vrc} + T3_{vrc}) + \theta X_j + \varepsilon_{vrc} \quad (2)$$

In equation 2, β_1 is the estimated effect of reminder SMS messages and β_2 is the estimated effect of development-oriented turnout messages. Hypothesis 2 predicts the rejection of the null hypothesis that $\beta_1 = \beta_2$, and that turnout is higher in VRCs that received development-oriented SMS messages, compared to VRCs that received reminder SMS messages.

To test **Hypothesis 3**, we estimate heterogeneous treatment effects for National Development message:

$$Y_{vrc} = \delta_0 + \delta_1 T3_{vrc} + \delta_2 ID + \delta_3 (T3_{vrc} * ID_{vrc}) + \theta X_j + \varepsilon_{vrc} \quad (3)$$

where Y_{vrc} is the percent voter turnout at the VRC level; T3 is National Development treatment; ID is measure for national vs ethnic identity; X_j is a vector of block dummies; ε_{vrc} is the error term.

Hypothesis 3a tests if $\delta_3 = 0$, where β_3 is the interaction term between *National Development* Treatment and Identity. Higher levels of national identity correspond with higher levels of the *Identity* variable. Hypothesis 3a predicts this equality is rejected and $\delta_3 > 0$, meaning treatment effects of National Development treatment are higher for higher levels of national identity.

Hypothesis 3b looks at heterogeneous effects for *Local Development* Treatment:

$$Y_{vrc} = \eta_0 + \eta_1 T2_{vrc} + \eta_2 ID + \eta_3 (T2_{vrc} * ID_{vrc}) + \theta X_j + \varepsilon_{vrc} \quad (4)$$

Hypothesis 3b tests if $\eta_3 = 0$, where β_3 is the interaction term between *Local Development* Treatment and Identity. Higher levels of ethnic identity correspond with lower levels of the *Identity* variable. Hypothesis 3b predicts this equality is rejected and $\eta_3 < 0$, meaning treatment effects are higher for higher levels of ethnic identity (lower levels of *identity* variable).

Hypothesis 4 looks at heterogeneous effects for *Local Development* Treatment:

$$Y_{vrc} = \tau_0 + \tau_1 T2_{vrc} + \tau_2 Live_ward_{vrc} + \tau_3 (T2_{vrc} * Live_ward_{vrc}) + \theta X_j + \varepsilon_{vrc} \quad (3)$$

Where $Live_ward_{vrc}$ is a measure of the percent of registered voters currently living in the ward where they vote. Hypothesis 4 tests if $\tau_3 = 0$, where β_3 is the interaction term between *Local Development* Treatment and the percent of registered voters who live in the ward where they vote. Hypothesis 4 predicts this equality is rejected and $\tau_3 > 0$, meaning treatment effects for Local Development treatment are higher for VRCs with a higher percentage of registered voters living in ward where they vote.

References

- Aker, Jenny C., Paul Collier, and Pedro C. Vicente. "Is information power? Using mobile phones and free newspapers during an election in Mozambique." *Review of Economics and Statistics* 99.2 (2017): 185-200.
- Arceneaux, Kevin, and David W. Nickerson. "Who is mobilized to vote? A re-analysis of 11 field experiments." *American Journal of Political Science* 53.1 (2009): 1-16.
- BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L., Runfola, D. 2017. Geocoding Afrobarometer Rounds 1 - 6: Methodology & Data Quality. Williamsburg, VA. AidData.
- Berinsky, A. J., et al. "Disengaged by Choice? A Research Agenda for Understanding Low Urban Youth Turnout in South Africa." (2016) working paper
- Dale, Allison, and Aaron Strauss. "Don't forget to vote: text message reminders as a mobilization tool." *American Journal of Political Science* 53.4 (2009): 787-804.
- Gerber, Alan S., and Donald P. Green. "The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment." *American Political Science Review* 94.3 (2000): 653-663.
- Marx, Benjamin, Vincent Pons, and Tavneet Suri. The Perils of Voter Mobilization. No. w23946. National Bureau of Economic Research, 2017.
- Malhotra, Neil, et al. "Text messages as mobilization tools: The conditional effect of habitual voting and election salience." *American Politics Research* 39.4 (2011): 664-681.

APPENDIX

Appendix A: Variable Construction

Appendix B: SMS Message Screen Shots

Appendix C: Randomization

Appendix A: Variable Construction

Outcome Variables

VRC Turnout - Ratio of total votes cast to total registered voters, at each Voter Registration Center (VRC).

Explanatory Variables

Identity - This variable is constructed using data from the 6th round of Afrobarometer survey in Sierra Leone. The 6th round contains 1190 observations, of which 509 are coded with village-level precision across 66 locations. We match each polling station to the closest location coordinate. The polling station identity variable is the average response to question 88B, where a 5 is highest level of attachment to national identity and 1 is highest level of attachment to ethnic identity. In Sierra Leone ethnicity is to a large degree location specific. Two exceptions are made for matching polling stations to Afrobarometer survey responses. First, urban and rural polling stations must be matched to urban and rural Afrobarometer locations, regardless of distance. Second, we do not match polling stations to survey response locations that are more than 15 kilometers away.⁴

Afrobarometer Question 88B: Let us suppose that you had to choose between being a Sierra Leonean and being a _____ [R's ETHNIC GROUP]. Which of the following statements best expresses your feelings?

Notes: [If respondent did not identify any group on Q87 – that is, IF they REFUSED to answer (9998), DIDN'T KNOW (9999), or said "Sierra Leonean only" (9990) – then circle 7

Answer categories:

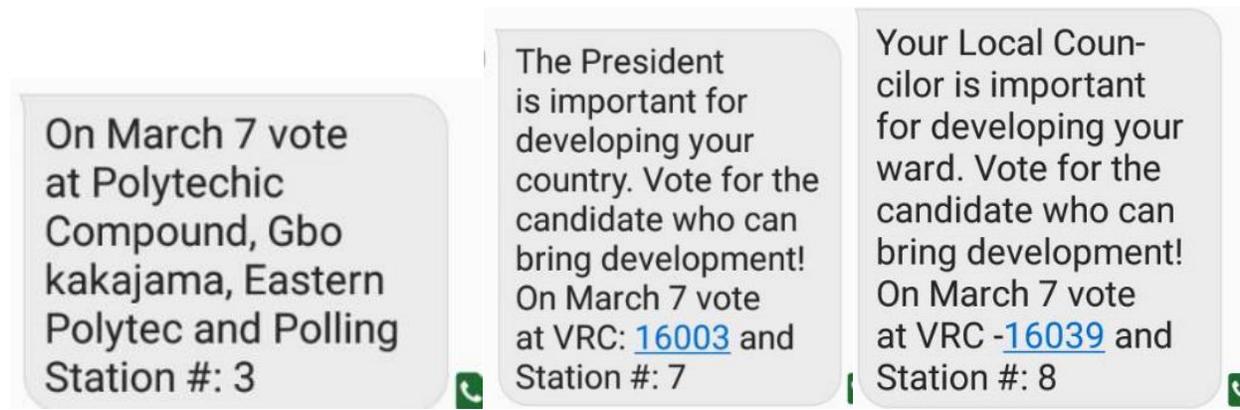
- = Not Applicable for questions 88A and 88B and continue to question 89.]
- I feel only Sierra Leonean 5
- I feel more Sierra Leonean than _____ [insert R's ethnic group] 4
- I feel equally Sierra Leonean and _____ [insert R's ethnic group] 3
- I feel more _____ [insert R's ethnic group] than Sierra Leonean 2
- I feel only _____ [insert R's ethnic group] 1
- Not applicable [Do not read] 7
- Don't know [Do not read] 9

Network Coverage - A voting center is coded "in-coverage" if it falls within a coverage polygon from the Orange network coverage data.

Vote "in-ward" - Voter registration data includes information on current voter residence. Voters are coded as living "in-ward" if their current residence matches a list of villages for that ward. Given differences in village spellings, we fuzzy match. For robustness, we'll code matches at multiple levels of maximum match distances.

⁴ Ofcourse, this is an arbitrary cutoff. For robustness, we will explore other cutoffs.

Appendix B: SMS Message Screen Shots



Treatment 1: reminder

Treatment 2: Reminder + Local

Treatment 3: Reminder + National