

Pre-Analysis Plan for “Diminishing the Effect of Vote-buying on Electoral Outcomes in India: A Pilot RCT to Test the Effectiveness of Radio Messages”

Donald P. Green

Srinivasan Vasudevan

May 14, 2014

Overview:

Vote-buying is a widespread practice in developing countries that may undermine the accountability and responsiveness of elected officials, especially to the needs of the poor. One way to reduce the incidence of vote-buying is by diminishing its effectiveness in influencing electoral outcomes. Vote-buying may derive its effectiveness in influencing voter behavior by appealing to internalized norms of reciprocity, conveying information about candidates or because of lack of faith in the secret ballot. We designed and deployed a radio campaign during the April-May 2014 general election period in India.

Our radio campaign consisted of a series of three 60-second spots. Each radio spot consisted of a satirical vignette about the hidden social costs of vote buying. The radio spots highlighted the incentive of vote-buying candidates, which is to earn returns on their “investment” on getting elected. Further, it raised questions on the credibility of their election promises. Finally, it reminded voters of the secret ballot and urged them to vote for honest candidates and punish vote-buying candidates. Through the radio messages our goal was to diminish positive voter reciprocity (indeed we appeal to their vindictiveness by urging them to punish vote-buying candidates), undermine the signaling value of the bribes and remind voters of the secrecy of the ballot. The Appendix contains scripts of the radio spots aired.

The 2014 general elections were conducted in 9 phases across the country from April 10 to May 12. The results of the elections of all constituencies will be released on May 16. Election rules came into force once elections were declared on March 10. One of the rules sets aside a cooling off period of 48 hours before polling begins when political campaigning by candidates and parties is forbidden. Our (non-partisan) radio campaign was strategically timed during this period so candidates could not adapt their behavior in response to it. Thus, any effect of the campaign can be attributed to voter behavior alone.

We will use two main sources of data in our analysis. The first is interviews of local journalists on aspects of campaigning in constituencies they cover. We attempted to conduct all the interviews before the radio campaign started and did not, during the course of the interviews, reveal any information about the intervention. We aimed to get responses from two journalists for each constituency. The Appendix contains the script we used for the phone interviews of the journalists. The second main data source is the official polling-station level election results.

Research Design

Akashvani, also known as All India Radio (AIR), is the national public radio broadcaster of India. AIR's transmissions cover 95% of the country by area and 99% by population. It is the only radio broadcaster with comprehensive coverage in rural areas. As of 2014, it operated 194 stations which broadcast external advertising.

Of the 194 stations, 46 radio stations had transmitter power greater than 20 KW and therefore had large coverage areas that overlapped with neighboring radio stations. In order to minimize interference we eliminated these high-power stations. Further, for budgetary reasons we excluded 30 additional radio stations that belonged to the most expensive advertising categories. AIR broadcasts in the 22 nationally recognized regional languages and additional languages recognized by individual states. Since it is infeasible to run the campaign in so many languages, we restricted our sample to languages that could be understood by the largest listener populations – Hindi, Marathi, Telugu, Kannada and Odiya. The radio stations belong to 10 of the major states in India, which contain 67% of India's population. This restriction eliminated 47 additional stations from the sample. We additionally dropped 10 stations that covered more than 4 million people and finally 1 station to make the final sample size an even number. Thus the final number of radio stations we obtained was 60.

The radio stations range in power between 1 KW and 20 KW, with 40 radio stations of 6 KW power. Of the 60 stations, 50 are FM and 10 are AM stations. The expected radii of the coverage areas lie between 50 km and 125 km. The coverage areas of the radio stations contain parts of different constituencies ("segments"). This constituency segment is the level of analysis in this study.

We divided the 60 radio stations into 5 blocks based on the expected day of election in that region. Half of the radio stations in each block were randomly assigned to receive the radio

campaign (“treatment” group), and the other half were to not receive it (“control” group). Thus, the probability of assignment to treatment is equal to half and is the same in all 5 blocks. The Stata code used to randomly assign the radio stations to treatment and control groups is included in the Appendix.

We produced the 3 spots in each of the 5 languages and broadcasted the spots in the local language of the radio station. Each spot aired 16 times (8 primetime and 8 non-primetime) over a period of 3 consecutive days (starting 2 days prior to polling and including the day of polling). Thus each treatment radio station broadcasted 48 spots over 3 days.

There are 543 constituencies in the country, out of which 338 are in states in our sample. Constituencies that are not completely covered by any of the 60 stations will be excluded from the sample. We do not currently have a detailed mapping of the radio propagation zones, but we expect our final sample to consist of approximately 275 constituency segments. There are on average around 1.5 million registered voters, 15 candidates and 1500 polling stations per constituency. Our focus will be on turnout and on voting for main candidates.

Electoral returns are released at the polling station level. We have the locations for all polling stations from the Election Commission of India (ECI). We will aggregate the electoral returns from the polling stations that lie within each constituency segment. We identified the main candidates based on the journalist interviews. We expect this variable to be quite straightforward but will check it against other specifications such whether the candidate exceeds the legally set threshold of vote-share (one sixth) to avoid forfeiting the deposit amount for contesting.

Estimation

The outcome of primary interest in this study is the effect of the treatment on the vote-share of candidates, particular vote-buying candidates. The other outcome of interest is the effect of the treatment on voter turnout, particularly in constituencies with high levels of vote-buying. Since vote-buying cannot be observed directly, we use different proxies for vote-buying by candidates and for aggregate levels of vote-buying in constituencies. In order to improve the precision with which the average treatment effects are estimated, we will present both the unadjusted difference-in-means estimate of each average treatment effect as well as a regression-adjusted estimate using lagged outcome variables (i.e., vote-share or voter turnout in the previous

(2009) general election, respectively) as covariates. The regressions we will estimate are as follows.

Base specification:

$$y_{ij,2014}^k = \alpha_1^k + \alpha_2^k T_j + \epsilon_{ij}^k$$

Base specification with covariate adjustment:

$$y_{ij,2014}^k = \beta_1 y_{ij,2009}^k + \beta_2^k T_j + \beta_3^k B_5 + \epsilon_{ij}^k$$

where, $y_{ij,2014}^1$ and $y_{ij,2009}^1$ are the vote-shares of the vote-buying candidate/party of constituency i covered by radio station j in 2014 and 2009 respectively. $y_{ij,2014}^2$ and $y_{ij,2009}^2$ are the voter turnouts of constituency i covered by radio station j in 2014 and 2009 respectively. T_k is the treatment group indicator of radio station j . B_5 is a vector of 5 block indicators. ϵ_{ij}^1 and ϵ_{ij}^2 are the idiosyncratic error terms.

Hypothesis tests:

1. The treatment decreases the vote-share of vote-buying candidates (one-tailed test: $\beta_2^1 < 0$). We will identify the vote-buying candidates based on the response of journalists of who they thought was doing the highest amount of secret campaign spending.
2. The treatment changes turnout in constituencies with most vote-buying (two-tailed test: $\beta_2^2 \neq 0$). We will identify the constituencies with the highest levels of vote-buying in three different ways: i) Whether the constituency has an above median number of “vulnerable” polling stations as identified by the ECI, ii) Whether the constituency has an above median number of reported election rules violations, and iii) Whether the constituency has an above median number/quantity/value of seizures of cash, liquor, drugs and other articles used for vote-buying. The treatment could decrease turnout if the reason many of the voters vote is because they received bribes and the treatment is effective in urging voters to not reciprocate. On the other hand the treatment could mobilize voters and increase turnout. Thus the net effect of the treatment on turnout could be positive or negative.

Exploratory Analyses

Since vote-buying is difficult to measure and since uncertainty surrounds which variables are good proxies for identifying vote-buying, assessing which candidates and constituencies are prone to vote-buying remains an open question. To address the robustness of our measurement approach, we will carry out exploratory analyses to test the effect of the treatment using the below described proxies. Note that testing some of proxies such as violations and seizures is subject to availability of the data from the ECI.

Proxies for identifying vote-buying candidates:

1. Whether the candidates has held the highest number of public events.
2. Whether the candidates has the highest number of election workers.
3. Whether the candidate has the highest number of violations reported.
4. Whether the candidate is perceived to be doing the most campaign spending.
5. Whether the candidate has the highest total audited campaign expenditure.
6. Whether the candidate is accused of crimes.
7. Whether the candidate has declared the highest assets.
8. Whether the candidate has the most number of election rules violations reported.
9. Whether the candidate is the sitting representative or belongs to same party as the sitting representative.
10. Whether the candidate belongs to the same party as state's chief minister.
11. Whether the candidate belongs to the same party as the prime minister.
12. Whether candidates is expected to win.

Proxies for identifying constituencies with high levels of aggregate vote-buying:

1. Whether the constituency has any candidates accused of crimes.
2. Whether the constituency has above median total audited campaign expenditures.
3. Whether the constituency has above median proportion of rural voters.
4. Whether the constituency has above median proportion of illiterate voters.
5. Whether the constituency has above median proportion of Scheduled Caste (SC) or Scheduled Tribe (ST) voters.