SIERRA LEONE 2012 ELECTIONS PROJECT

PRE-ANALYSIS PLAN: INDIVIDUAL LEVEL INTERVENTIONS

PIs: Kelly Bidwell (IPA), Katherine Casey (Stanford GSB) and Rachel Glennerster (JPAL MIT)

THIS DRAFT: 15 August 2013

This study examines the impact of providing citizens with information about Parliamentary candidates via structured inter-party debates in the lead up to the Sierra Leone November 2012 Elections. Randomization and treatments were conducted on multiple levels: constituency, polling center and individual (details on sampling and randomization are available in the project's AEA trial registry). This pre-analysis plan governs the analysis of the individual level treatments only. It was written and registered before analysis of the individual treatments data. It incorporates learning from analysis of the before/after screening data within the PC-level treatment sites.

1.   Background

Our NGO partner, Search for Common Ground (SFCG), hosted and filmed debates between Parliamentary candidates in 14 constituencies. In order to unpack the distinct types of information revealed by these debates, and the extent to which voters differentially react to such information, we implemented four separate treatment arms and two control arms in a crossed randomized design at the individual level. The arms aim to disentangle four different components of the debate that might impact voting behavior: i) objective information such as candidates' policy platforms and past experience; ii) less tangible indicators of candidate capacity such as candidates' interpersonal and persuasion skills, which signal greater productive ability on the job as an MP; iii) superficial information about candidates such as good looks and wealth; and iv) the social mobilization effect of the community gathering together to discuss politics.

The four treatment arms were as follows: i) the exact same debate screened in the polling-center level treatment but now delivered to an individual; ii) a "getting to know you" video of the same candidates speaking informally about their hobbies and interests; iii) a recording of an independent moderator or journalist summarizing the main policy positions articulated by the candidates during the debates; and iv) a "lab" experiment where respondents make quick inferences after being exposed to pairs of isolated images, short (20 second) video recordings, and a list of candidate names, all collected from MP races outside the respondents' district. The first two treatments were screened to individuals via tablet, the third via audio device and the fourth via a combination of pictures and video recordings on a tablet. At the time of treatment, individuals were surveyed both before and immediately after exposure to the treatment. A fifth control group was also surveyed. Respondents in all five arms were then surveyed again at Election time in the exit polls. We are interested in both short run effects, which are captured in the "After" survey data, as well as longer run effects captured the exit poll data. A sixth control group was not contacted at the time of treatment but was included in the exit poll. This second "pure" control group serves as a robustness check on the concern that the information conveyed by debates does not in fact matter but that instead the surveys and research activity simply increase the salience of the election and encourage greater deliberation in voting choice. We will test for this by comparing the first group of controls, who may be primed to pay attention to the election through the experience of being surveyed, to

these pure controls who are exposed to the same potential spillover effects but who first encounter the research teams in the exit polls after they have voted.

Within the 14 constituencies selected for participation in the debates, we first allocated polling centers to the polling-center level intervention and control group. This PC-level sample drew in VRCs that were sufficiently small (fewer than ~900 registered voters) and far apart from their nearest neighbor (at least ~ one mile). In constituencies where there were a sufficient number of polling centers left over after this process, we allocated the remaining larger and closer together VRCs into the individual-level treatment and control group, stratifying by number of registered voters and distance to nearest neighboring VRC. Within each of the treatment VRCs, households were divided into those with only female registered voters, only male, and both male and female registered voters (based on an earlier household listing exercise). We randomly assigned the six arms to households within each of these bins, and randomly selected respondents within each household to receive the individual-level treatments and/or survey(s).

2.   Econometric Specifications

We are interested in the absolute treatment effect of each of the three treatment arms (debate, get to know you and radio report) compared to the control group, as well as the net or relative effect of each treatment arm compared to the other treatments. Analysis of the lab experiment treatment arm will be conducted separately. Analysis will take the following form:

$$Y_{ihtpc} = \beta_0 + \delta_t T_{htpc} + X'_{hpc}\Pi + Z'_{pc}\Gamma + W'_{ihtpc}\Psi + c_p + \varepsilon_{ihtpc} \qquad (1)$$

where outcome Y (i.e. vote choice) is measured for individual i living in household h assigned to treatment arm t registered in polling center p located in Parliamentary constituency c. T is a dummy variable indicating assignment to treatment arm t; X is a vector of indicator variables that denote the stratification bin from which the household was drawn (where the bins were determined by the gender composition of registered voters); Z is a vector of indicator variables that denote the stratification bin from which the polling center was drawn (where the bins were constructed by number of registered voters and distance to nearest neighboring center); W is a set of additional control variables that will be determined from analysis of the control group data and will vary by hypothesis with an eye toward identifying individual characteristics that do not vary with treatment and that help explain variation in a particular outcome (i.e. education and radio ownership are likely positively correlated with general political knowledge); c is a set of constituency-specific fixed effects (the level of debate and candidates); and ε is an idiosyncratic error term clustered at the polling center level. Our main specification includes the full set of controls (X, Z and W); we will also show results for the sparser specification that includes only the stratification variables as controls (X and Z only) as a robustness check. For each treatment arm, the coefficient of interest is δ_t, the average treatment effect for treatment t compared to the control group. The control group is defined as only the respondents in the first control arm for immediate outcomes in the After survey; and respondents in both the first and "pure" control arms for longer run outcomes in the exit poll survey. Participants in the Lab experiment arm, who were not exposed to any media concerning candidates from their own constituency, will also serve as controls. We will further conduct tests pooling multiple treatment arms together, for example $\delta_{t\in D,R,G} \geq 0$. We will further test a series of hypotheses about the relative effects of the different treatment arms that take the form $\delta_t \geq \delta_{\sim t}$ explained below. For each arm and combination of arms, we will examine immediate effects captured in the After survey as

well as sustainable effects captured in the exit poll survey. Tests will be one-sided in the direction stated in the hypothesis, unless specified as two-sided. We will report treatment effects of all individual outcomes of interest as well as mean effects indices by hypothesis and sub-hypothesis as organized below.

Heterogeneous Effects

We will test for heterogeneous treatment effects at the level of constituency, candidate and voter, adjusting for multiple inference within each level (i.e. grouping together the tests for all of the voter-level sub-groups when adjusting standard errors). Specifically we will test for differential effects along the following dimensions:

- **Competitiveness of constituency**: the impact of debates on voting choices is expected to be increasing in the competitiveness of the race, as measured by the (decreasing) margin between vote shares for first and second place MP candidates in the previous 2007 election. We will also test (two-sided) whether competitiveness differentially effects measures of policy alignment and likeability.
- **Candidate performance:** better debate performance (as measured by the expert panel) is expected to enhance the effects of debates on the ability of voters to correctly locate candidate's policy positions, and divergence in performance is expected to enhance treatment effects on votes for the debate winner.
- **Lesser known candidates (secondary):** voter response to strong (weak) performance by less well known candidates (including PMDC, female and non-incumbents) may be stronger than that for other better known candidates, as voters may have greater scope for updating their beliefs
- **Subgroup analysis**: the voting literature suggests that the impact of debates could vary by gender, age, and level of political informedness / naïvete. These tests will be two-sided. We further predict weaker effects for people who do not speak Krio well and may have had trouble understanding the debate.

Descriptive Analysis

- **Dissipation of effects:** For each arm and hypothesis we will test for dissipation of effects by comparing the size of TE measured in the After survey to its equivalent in the Exit poll.

3. Hypotheses and Outcomes

This part of our research agenda seeks to answer two broad questions: first, whether debates, get to know you and radio reports impact voting behavior; and second, if they do have an impact, what drives the effects of debates - information about policy or persona or both. Part of this second question concerns the speed with which voters can infer information about individual candidates.

Note that the set of outcomes available differs between the after survey and the exit poll, where in general the exit poll is a subset of the After survey (however a few items are only available in the exit poll). The list below is comprehensive of both sets, but tests of short- and longer-run effects will necessarily be limited to outcomes available in the respective datasets.

Notation: Individual level treatment arms are denoted by D = Debate; G = Getting to know you; R = Radio report; and L = Lab.

- **Vote choice outcomes**
  - Tests to conduct: $\delta_t \geq 0$ *for* $t \in D, R, G$; $\delta_{t \in D,R,G} \geq 0$; $\delta_t \neq \delta_{\sim t}$ *for* $t \in D, R, G$
  a. Hypothesis 1: Exposure to debates increases **vote shares** for the candidate that performed the best in the debates
      i. TE measured by vote choice
      ii. Debate winner / loser measured by audience ratings and expert assessment
  b. Hypothesis 2: Exposure to debates increases the willingness to **vote across party lines**
      i. TE measured by vote choice and ethnicity (limited to members of affiliated tribes)
      ii. Three additional (also primary) measures expand the concept to incorporate non-affiliated tribes: voting for a different party for MP in 2012 than in 2007; splitting ticket for MP (i.e. party MP different than party Pres or party LC); and voting for a party other than the self-reported party affiliation from before survey.
  c. *Mechanisms of impact*
      i. *Policy stance and professional qualifications matter* for vote choice
          1. As the radio report isolates the impact of policy stances and professional qualifications information conveyed by debates from the persona aspects that are also revealed by debates, $\delta_R \geq 0$ provides evidence in support of the role of these characteristics in voting decisions (Note that an absence of these effects for R does not on its own rule out the value of policy and professional background for vote choice.)
          2. For the debate treatment, we expect to see the impacts of debates on votes for the winning candidate to be increasing in the divergence between candidates in expert assessments of competence (professional qualification and connectedness).
      ii. *Candidate persona matters* for vote choice
          1. As the Get to know you video removes the policy content, $\delta_G \geq 0$ provides evidence in support of the role that candidate persona and likeability play in vote choice

- **Voting knowledge and behavior outcomes**
  a. Tests to conduct: $\delta_t \geq 0$ *for* $t \in D, R$; $\delta_{t \in D,R} \geq 0$; $\delta_t \neq \delta_{\sim t}$ *for* $t \in D, R$. Note that G only relevant for subset of outcomes indicated below.
  b. Hypothesis 3: Exposure to debates increases **political knowledge** and leads to more informed voting
      i. TE measured for general political knowledge includes ability to name MP roles, CFF amount, healthcare entitlement, gender equity percentage, the meaning of MP, who approves the government budget, and exact date of election. Note we do not expect heterogeneous effects on these measures across candidates or constituencies.

        ii. TE measured for individual candidate attributes includes ability to name candidates, distinguish better educated, greater public office experience and older candidates, incumbency and more likely to report personal characteristic as primary determinant of voting choice. Note distinguishing between candidates only applies where the characteristic diverges across candidates in the constituency.

            1. Include $t = G$ for naming candidates and distinguishing better educated, more experienced and older candidates.

        iii. TE measured for candidate policy stances by ability to correctly place candidate view on Gender equity, first priority issue, free health care implementation, and CFF transparency; plus local policy issues (where available)

c. *Mechanism of impact*

        i. ***Comprehension and attention*** may vary by mode of information delivery. A finding that $\delta_D > \delta_R$ for general political knowledge questions (H3) suggests that debates may better engage the audience than radio summaries. Check for waning attention by placement of knowledge questions in the program (i.e. MP roles at the beginning, date of election at the end)

        ii. For D, the impact on correctly locating candidate positions should be increasing in the performance of the candidates in answering policy questions as assessed by the expert panel.

d. Hypothesis 4: Exposure to debates increases **policy alignment**

        i. TE measured on the empirical match between voter's position expressed in survey and reported stance of his/her selected candidate expressed in the debate on gender equity, priority issues, and free health care implementation. This alignment can result from persuasion and/or voters choosing aligned candidates, and we aim to unpack the two mechanisms using the before and after data.

e. *Mechanism of impact*:

        i. ***Persuasion*** reflects voters adopting their preferred candidate's policy stances (i.e. because they are convinced by the candidate's arguments, learn something about the policy from the debates, and/or are reducing cognitive dissonance between their candidate and policy preferences). These measures will be tailored by constituency to reflect candidate positions expressed in the debate, but in most cases we expect APC-(SLPP-)leaning voters to express a more positive (negative) view of FHC implementation compared to their counterparts in control areas, and for treated voters to report greater support of the GEB. A before/after test limits the sample to voters whose candidate choice matches their pre-stated party loyalty (i.e. holding party choice constant) and checks for improved alignment in the after survey. Limiting the sample to Independent voters and those of historically unaffiliated tribes is another test, both for policy alignment and more generally for changing voter positions on issues irrespective of vote choice. Power for these tests, however, will likely be low given small sample sizes.

        ii. ***Choosing aligned candidates*** reflects moving to a candidate whose policy positions match the voters *ex ante* policy preferences. These measures will be

tailored by constituency to reflect the actual divergence amongst candidates. A before/after test focuses on the voters policy preference stated in the before survey and looks for differential improvements in alignment in constituencies where the candidates diverge on a particular issue compared to constituencies where all candidates agree.

      iii. Finding that $\delta_D > \delta_R$ for measures of policy alignment provides evidence for the impact of ***productive intangibles*** related to MP job performance not easily captured by policy positions and professional background. Finding $\delta_G \geq 0$ for measures of candidate education and professional experience (H3) suggests that these attributes may be conveyed by pattern of speech and presentation.

  f. Hypothesis 5: Exposure to debates enhances **voter openness** to other parties
      i. TE measured by higher likeability ratings for all candidates (i.e. own party, rival party, and third party where applicable) in exit polls
         1. Include $t = G$

  g. *Mechanism of impact*:
      i. Finding that $\delta_D > \delta_R$ and $\delta_G > 0$ for measures of candidate **likeability** suggests that voters value the general affability of candidates

- **Thin slice outcomes**
  a. Hypothesis 10: T**hin slices of information** enable voters to make accurate inferences about individual candidate attributes
        i. Benchmark test 1: Inferences regarding ***facial competence from photos***: test whether voter inferences about leadership ability and their voting preference after viewing photos predict the winners of MP races with an accuracy greater than chance
        ii. Benchmark test 2: Inferences regarding ***productive intangibles from public speaking***: test whether voter inferences about leadership ability and their voting preference after viewing video clips predict the winners of debates and MP races with an accuracy greater than chance

  b. *Mechanism of impact*:
        i. ***Productive intangibles are more important than purely superficial cues*** for voting preference: accuracy of benchmark test 2 > benchmark test 1
        ii. If accurate, explore which inferences are most correlated with predictive power, i.e. judge to be better leader vs better looking?

- **Secondary outcomes** See details in PC-level PAP
  a. Hypothesis 6: Exposure to debates mobilizes the public and leads to greater **turnout**
        i. TE measured by turnout question in exit polls
        ii. Note we are "demoting" this hypothesis to secondary since turn out in the actual election was extremely high (87.3%) and we thus expect to have insufficient power to detect treatment effects on this outcome, but are retaining it in the analysis as we expect it to be of general interest
  b. Hypothesis 7: Exposure to D, G and R increases the perceived **legitimacy** of elections
  c. Hypothesis 8: Exposure to D, G and R increases **interest in politics**

Additional Secondary Questions

- Hypothesis 12: **Social mobilization**, the community gathering aspect of the polling center-level treatment intensifies the effect of the debate on voter behavior
    a. TE measured by comparing the TE of the I-level D to the PC-level D. If comprehension and other effects are similar, but other effects are lower for I-level, conclude that the community gathering aspect had a separate effect from the content.
- Hypothesis 13: Effect of treatments is not just **priming** / getting people to focus on the election
    a. TE measured as no differences between pure controls and surveyed controls in Individual T implementation areas (note: both subject to the same potential spillovers from speaking with neighbors in other T arms, but only the surveyed one is primed)
- Hypothesis 14: Debates do not **increase ethnic voting**
    a. Check whether exposure to D and G increases the ability of voters to correctly guess individual candidate ethnicity
    b. In L, check that ability to guess individual ethnicity based on names (which are printed on ballot) and photos (which are posted outside polling stations) is at least as strong as that based on video clips (which are additional information revealed by debates)