

# Ballot initiatives and information processing, an experimental approach (Pre-Analysis Plan)\*

Angèle Delevoye<sup>†</sup> and Trevor Incerti<sup>‡</sup>

*Yale University*

June 11, 2019

## Abstract

X

---

\*

<sup>†</sup>angele.delevoye@yale.edu

<sup>‡</sup>trevor.incerti@yale.edu

# 1 Introduction

One of the ways by which citizens produce policy is through ballot initiatives. Ballot initiatives have been hailed as an efficient direct democracy mechanism, providing citizens with a much-needed voice in policy making and politics. Defenders of ballot initiatives claim that they provide a way for citizens to voice their opinions, and produce policies without as much interference from moneyed interests and even parties. On the other hand, a lot of scholarly work caution about recent excesses of direct democracy: F. Rosenbluth and I. Shapiro argue that decentralizing power to the grass roots has worsen the quality of policies outputs and decision making in general by weakening the parties (Responsible Parties). Doubts about voters' competency are also central to this question, and the need for knowledge might be even larger on a ballot initiative than it is in a traditional partisan election (less cues and shortcuts available, often technical topics). The central research question for this paper will be: how do voters process information when asked to vote on a non-partisan, technical policy proposal? Are they able to distinguish between different sources (qualities?) of information?

We propose a field experiment to test whether citizens acting as policy makers (in the context of a ballot initiative) give more credence to high quality research compared to no information or overtly partisan or biased information (from lobbies). We complement this field experiment with a survey experiment replicating the exact same treatment and outcomes measures than the field experiment. The goal is to see whether in the context of ballot initiatives, technical policy decisions, etc, suvey experiments are less far from the reality of field experiments than we are learning they are in other political settings (campaign messages and corruption information at least). We would register our priors that survey experiments might get us closer to real-world conditions in the context of ballot initiatives. If we are wrong, then we would add to the growing litterature establishing the limits of survey experiments. If we are right, then we would open the doors for more survey experiments in the

context of information provision and ballot initiative. Survey experiments offer options that field experiments will never do: providing fake information to participants as the treatment for example. But first we need to make sure they get us close enough to a real-world situation.

## 2 Theory

### *2.1 Science, research and democracy*

This project may speak to broader theoretical questions on the functioning of our current democracies, and the potential gaps between reality and the ideals of deliberative democracies. In this sense, we hope this project will contribute to ongoing discussions within philosophy of science, such as perceptions of research methods and consumption of scientific evidence outside of the academy.

[Tensions and relationships between science and democracy]

### *2.2 Information experiments*

Campaigns information (Kalla Brookman meta analysis) Other political info, corruption and else

### *2.3 Ballot initiatives*

## 3 Experimental design: field experiment

### *3.1 Treatment groups and randomization*

We would provide voters with different pieces of information leading to a vote on a ballot initiative. This could be different information sources (business vs advocacy group vs academic/scientific), or information vs not information. This information provision could be done through mail (a postcard) or canvassing and phone calls. Postcards: more power, and we would contribute to the literature (postcards don't work in general election campaigns

- Kalla Brookman etc, but could they work on a single, specific policy proposal). Canvassing: less power/geographic reach and probably costlier, but much higher expected treatment effect

### *3.2 Outcomes*

Vote on the ballot initiative (at the precinct or smaller level).

### *3.3 Choice of policy and election cycles*

Ideal policy: a scientific consensus exists, on a technical subject on which people have very weak priors. Health care, housing, etc.

Election cycle: off-year vs on-year election. On-year would probably give us access to a lot more voters without strong priors on the specific proposal.

## **4 Experimental design: survey experiment**

### *4.1 Why do a survey experiment?*

- Methods: field vs survey experiments. Contribute to the growing evidence that survey experiments don't give us what we want (campaign reach outs, corruption information, etc). Do it in a very control environment: we would be able to design the survey so that it exactly replicates the field. - Our priors: survey experiments might be closer to the reality of the real world/a field experiment than in other political contexts. Because technical information, specific proposal with weak priors, no partisan shortcuts, no emotions/feelings involved bc not voting for a person. - If right: open the door for more survey experiments, with more ambitious treatments (fake news?). If wrong: survey experiments don't even work in a context where we felt confident that they would. Let's keep building evidence on this.

## 4.2 Design

# 5 Potential natural experiment

The idea would be to use a real-life variation of information supply. Ex.: data on billboards, with their contents and where exactly they were put on, or a report sent to some precinct but not other. Info provision will probably not be random, so we would need to do a diff and diff (precinct with billboard ? precinct without billboard, but which other difference? Vote on another measure, vote in previous elections, opinion polls on the measure before billboards went up, early voting)?

# 6 Experimental analysis

## 6.1 Estimation of treatment effects

As noted above, we intend to use block random assignment in order to increase the precision of our treatment effect estimates as well as to facilitate (preregistered) estimation of heterogeneous treatment effects. Treatment effect estimates will therefore be calculated as the difference-in-means of the response rate from subjects in each of treatment groups and the response rate from subjects in the control group in each block, weighted by the number of subjects in each block relative to the total number of subjects. More formally, average treatment effects will be estimated as:

$$ATE = \sum_{j=1}^J \frac{N_j}{N} ATE_j$$

where  $J$  is the number of blocks, blocks are indexed by  $j$ , and  $\frac{N_j}{N}$  represents the share of subjects who belong to block  $j$ .

In practice, these differences-in-means will be calculated using a weighted least squares regression of response rate on treatment assignment, with weights corresponding to the inverse probability of treatment for each unit (IPW). All p-values will be calculated using randomization inference. As the reference/control group in the experiment is in effect the

“No information / lower tier” treatment group, all effects must be interpreted in relation to this treatment. In other words, treatment effects should be interpreted as the change in response rate relative to the “No information / lower tier” group.

## 6.2 *Heterogenous treatment effects*

In addition, we will test for heterogenous treatment effects on legislator party and educational attainment. Our test will take the form of regressing our outcome variable on treatments conditional upon the data representing the covariate of interest. In other words, we will split our sample by subject attributes (party and educational attainment), then estimate conditional average treatment effects (CATEs) separately for each of these attributes. For example, to test for heterogenous effects by party, we will regress outcomes on treatments for all legislators in each party. Following estimation of CATEs, we will use randomization inference to test the null hypothesis that CATEs in both groups (e.g. Democrat and Republican) equal to the ATE.

We recognize that the search for heterogenous treatment effects often suffers from the multiple comparisons problem. In a dataset with a large number of covariates, if a large enough number of subgroups is examined, it is highly likely that statistically significant interaction effects will emerge merely by chance. In other words, the more significance tests are performed, the higher the likelihood of falsely rejecting the null hypothesis at least once. We minimize this risk by preregistering our heterogenous effects of interests, as well as by performing multiple comparisons corrections using Bonferroni, Holm-Bonferroni, and Benjamin-Hochburg adjustment procedures.

## 7 Conclusion

Our next steps on this project include finalization of some of the research design decisions that remain open (i.e. federal or state level subjects, partnership with an organization, and choice of the policy). We hope to finalize these details over the summer and to preregister a final Pre-Analysis-Plan with EGAP by the end of 2019. We would then conduct the experiment

in 2020. We believe that this timeline would be ideal because of the political and electoral contexts expected in 2020. An election year increases the likelihood that parties will be seeking innovative policy ideas (and decreases the time spent on other legislative activities), and it is therefore possible that legislators will be particularly responsive to information on such policies.

## A Appendix