

# Pre-analysis Plan for Individual and Social Voter Learning and Measuring Support for Democratic Institutions

Daron Acemoglu, Cevat Giray Aksoy, Ceren Baysan, and Carlos Molina

May 24, 2023

# 1 Introduction

According to Freedom House (2021), the global decline in democracy has accelerated in recent years. The share of countries designated Not Free (i.e., countries where basic political rights and civil liberties are absent) has reached its highest level since the deterioration of democracy began in 2006. Today, less than a fifth of the world’s population lives in fully democratic countries. These trends have negative implications for economic growth and civil liberties. Whether these trends can be reversed partly depends on voter understanding and sensitivity to the quality of democratic institutions. However, in authoritarian regimes, voters’ opinions are influenced by leaders who frequently claim that existing systems are democratic. In addition, since state centralization is high in these environments, voters are also not exposed to local variation in institutional quality. Voters also have limited knowledge of changes in democratic institutions because of state censorship.

Whether exposure to credible information on democratic institutions changes voter understanding and support for democratic institutions remains an open question. This study aims to answer this question using a large-scale field experiment in the context of high-stakes presidential and general Turkish elections. Turkiye is an ideal setting for this research question due to the following reasons: (i) over the past 20 years, Turkiye has experienced a sharp decline in democratic norms; (ii) the May 2023 presidential and general elections are considered critical for democratic institutions in Turkiye and all political parties are campaigning on democracy.

Studies have shown that democratization has a causal impact on economic growth (e.g., Acemoglu et al, 2019) and democratic norms, like independent media, reduce corruption (e.g., Ferraz and Finan, 2008), but whether voters appreciate and support democratization in authoritarian environments is unclear. The lack of local variation in democratic institutions and state performance makes evaluating voter understanding of institutional quality challenging. Several studies use online experiments to study public commitment to democratic values (Miller, 2021; Svolik 2019), but evidence shows that online experiments can be biased due to participants’ tendency to provide socially desirable responses (Levitt and List, 2007). We contribute to this literature using a large-scale field experiment in Turkiye in which voters are randomly assigned at the neighborhood-level to receive information, via door-to-door election campaigns, on the causal relationship between independent media and corruption and time-series figures using data from V-Dem. Neighborhoods are also randomly assigned to a second treatment arm in which voters are exposed to campaigns with only persuasive content on corruption. Such “slogans” are typically used in campaigns, and political scientists and social psychologists believe they effectively change voter behavior. Therefore, we will directly compare whether slogans have a different impact than information on changing voter outcomes. We will compare these two treatment arms to one another and to a control group in which neighborhoods are not visited at all.

Our study is also related to the broader literature on the impact of information on voter behavior in middle-income and lower-income countries (e.g., Baysan 2022; Ferraz and Finan 2008; Casey 2015; Cruz, Keefer, and Labonne 2021; Platas and Raffler 2021; Cruz et al., 2018; Kendall et al., 2015; Orkin, 2019). We extend this literature by focusing on a previously unexplored angle that is salient in authoritarian countries – knowledge of the quality of democratic institutions and their relationship to economic development. Another

advantage of our study is that we will be able to study both real (i.e., neighborhood-level vote share) and attitudinal (i.e., via a household survey) outcomes.

Our study also contributes to this literature by disentangling between the effects of the information source and information on voter outcomes and estimating the role of two types of spillovers on individual beliefs. First, we have a unique opportunity to randomize the information source of the same information campaigns. Following an inclusive and non-partisan approach, and as discussed in our IRB, we contacted each political party's general secretary to participate in our study. Randomizing the information source is important because the information source can, for example, affect voters' beliefs of the information's credibility (e.g., Bray and Kreps, 1987) and whether voters are subject to persuasion (e.g., DellaVigna and Kaplan, 2007) or beliefs of the information source's (political party) reputation (e.g., Gentzkow and Shapiro, 2006). Second, to estimate spillovers effects of the campaigns at the household and street-level, we embed the neighborhood level randomization with a residential-unit level randomization of 3,020 voters and their neighbors, and a street-level randomization. For example, this strategy will allow us to estimate "treatment spillovers" (and "control spillovers") in which we compare the impact of each campaign type on treated households in control streets to treated households in treated streets, While other studies have estimated "control spillovers," (e.g., Giné and Mansuri 2018), there is less research on treatment spillovers. This spillover allows us to measure the extent to which voters update their beliefs because of social multipliers through interactions or because they know others were exposed to the same message.

This document summarizes the main tests that we intend to conduct. We have submitted the document prior to the public release of the electoral data for the 2023 May elections. We do not rule out the possibility of running other tests and engaging in exploratory analysis. We will make clear in the paper which estimations are specified in this pre-analysis plan and which are not (Casey, Glennerster, and Miguel, 2012).

## 2 Experimental design

We use a randomized controlled trial (RCT) to evaluate our research questions. To initiate the research program and conduct impartial research, in December 2022, we invited all political parties with representation in parliament to participate in a study on voters' perspectives on democracy. We sent an initial and reminder email to the provincial leaders of each party in three major provinces (Istanbul, Ankara, and Izmir). The Republican People's Party (CHP) and the Good Party (Iyi) agreed to participate. One of the campaign managers from the CHP also asked to involve a group of non-partisan activists interested in conveying information on democracy to voters.

As part of the research project, the two parties and independent canvassers carried out a door-to-door information campaign containing the same content on democracy and the relationship between the strength of independent media and corruption. The information was conveyed orally and in a pamphlet. The pamphlets were designed by the parties.

The information campaigns were randomized at the neighborhood level. In addition, neighborhoods were randomly assigned to a second treatment arm in which voters are exposed to campaigns with only persuasive content ("slogans") on corruption. We cross-

randomized the information source of the two treatment campaigns as either the İyi Party, Non-Partisan, or the CHP. This was revealed to voters through a pamphlet and orally by the canvassers. In total, there are 252 treatment neighborhoods and 302 control neighborhoods.

Our sample selection took place in March 2023 and a baseline survey was conducted in April 2023. An endline survey started approximately May 18th, 2023, but has been paused and will tentatively continue after the run-off elections on May 28th. There were some issues with baseline data collection, and so there is uncertainty in whether we will be able to conduct the individual analyses using the survey data. We expect that the administrative electoral data will be available around May 25th, 2023.

## 2.1 Sample

The primary field experiment takes place in Izmir, Turkiye. We have scraped ballot-box level vote share and voter turnout data for the 2018 presidential and general elections from the Supreme Election Council's website (YSK). We have aggregated the ballot-box level data to the neighborhood-level since ballot-box locations change over time.<sup>1</sup> We use these data to randomly assign neighborhoods to the different treatment groups.

Prior to randomization, we selected the experimental sample from the full population of 1,294 neighborhoods in Izmir as follows:

- We dropped neighborhoods where the number of registered voters was greater than 6500 in 2018
- We dropped districts where the share of neighborhoods with less than or equal to 100 voters was greater than 11%
- We dropped neighborhoods where the number of registered voters was less than 150 in 2018
- We dropped neighborhoods where either of the two major electoral alliances (People's Alliance and National Alliance) had more than 70% of the vote share in 2018
- We dropped neighborhoods where the National Alliance's vote share in 2018 was less than .15 and the presidential candidate, Selahattin Demirtas, had more than 50% of the vote share.
- Finally, among the remaining districts, we dropped the ones where there were less than 9 neighborhoods remaining.

Randomization is stratified by six quantiles of the 2018 vote share for the National Alliance to estimate heterogeneous treatment effects and to increase statistical power. We selected a well-balanced draw from among 100000 re-randomizations to avoid other chance imbalances (following Banerjee, Chassang, and Snowberg 2017) for each treatment arm. Each treatment arm is listed below.

---

<sup>1</sup>We may be able to construct a ballot-box level panel if we are able to access ballot box locations from 2018.

- Control
- $T^{I,CHP}$ : Information by CHP Canvassers
- $T^{I,IYI}$ : Information by IYI Canvassers
- $T^{I,IND}$ : Information by Independent Canvassers
- $T^{S,CHP}$ : Slogan by CHP Canvassers
- $T^{S,IYI}$ : Slogan by IYI Canvassers
- $T^{S,IND}$ : Slogan by Independent Canvassers

## 2.2 Data sources

- Electoral data that are publicly available from the Supreme Election Council's website (YSK).<sup>2</sup>
- Baseline survey
  - Demographics and ideology; views about the causal relationship between democratic institutions and various outcomes like corruption; views about the quality of democratic institutions and general questions on democratic norms; affective polarization; previous voting behavior. The baseline survey was conducted in Izmir.
- Endline Survey
  - Views about the causal relationship between democratic institutions and various outcomes like corruption; views about the quality of democratic institutions and general questions on democratic norms; affective polarization; voting behavior in the May 14 and May 28 general and presidential elections; questions about interactions and views of neighbors; demographics and ideology among voters not surveyed at baseline.
- Online Experiment
  - We conducted an online experiment with an information treatment arm similar to the door-to-door information campaign. The online experiment included additional information treatment arms and a placebo treatment group for each information treatment. The survey accompanying the experiment has similar variables as the baseline and endline surveys. The online experiment included voters from different parts of Turkey.

---

<sup>2</sup>These data are publicly available at <https://sonuc.ysk.gov.tr/sorgu>

### 3 Experimental integrity

To establish experimental integrity we will compare each of the treatment groups to the control group on key baseline and arguably pre-determined outcome variables. In expectation, the mean and variance of these variables in each treatment group and the control group should not be distinguishable. We will compare balance across neighborhood-level characteristics and individual-level characteristics. These variables include 2018 electoral data, geographic units, and the number of households surveyed at baseline.

### 4 Regression specification

#### 4.1 Electoral data analysis

The publicly available electoral data allows us to analyze the effect of our treatments at the ballot box level and neighborhood level. Our primary regression is given by:

$$Y_{bn} = \alpha + T'_n \beta + X'_n \lambda + \epsilon_{bn} \quad (1)$$

where  $b$  and  $n$  stand for ballot-box and neighborhood respectively.  $Y_{bn}$  is the ballot box-level electoral outcomes (for different parties, candidates, or alliances) or voter turnout for the 2023 presidential and general elections.  $T'$  is a vector of indicator variables identifying the different treatment arrangements that we discussed above. In addition, we will use residential unit level information on treatment implementation intensity: whether a canvasser successfully gave a pamphlet or completed a conversation with a voter. We will explore instrumenting this information (aggregated to the ballot box or neighborhood level) with treatment assignments or using it in our heterogeneity analysis.  $\beta$  is a vector carrying our main coefficients of interest measuring the differential effects on electoral outcomes between neighborhoods assigned to these treatment arrangements and the control group (the omitted category).

In the regression, we will include a vector of control variables,  $X'_n$ . We will use 2018 electoral data as control variables to improve the precision of the estimates and to control for any imbalance of variables that are predictive of the outcome. In addition, we will include other variables that we balanced on and that improve precision. The additional variables include geographic fixed effects, the number of households surveyed at baseline, and strata fixed effects (six quantiles of the 2018 vote share of the opposition).

In addition to the aggregate specification, we also estimate versions of equation (1) by quantiles of the 2018 vote share for the opposition.

Given the nature of our treatment, we cluster standard errors at the neighborhood level. We will also run the regression collapsed at the neighborhood level with and without weighting observations by the number of registered voters.

## 4.2 Individual level analysis

For the individual-level analysis, the treatment assignment is at the residential unit level. Our primary specification, in this case, is given by:

$$Y_{iu} = \alpha + T'_u \beta + X'_{iu} \lambda + \epsilon_{iu} \quad (2)$$

where  $i$  and  $u$  represent an individual and a residential unit respectively.  $Y_{iu}$  is an outcome variable for the individual.  $T'$  is a vector of indicator variables identifying the different treatment arrangements discussed above and  $\beta$  is a vector carrying our main coefficients of interest. As before, we include control variables aiming to improve the precision of the estimates.

We cluster the standard errors in these regressions at the residential unit level, which is the level at which the treatment is assigned.

To estimate spillovers effects of the campaigns at the household and street-level, we embed the neighborhood level randomization with a residential-unit level randomization of voters in our baseline sample and their neighbors, and a street-level randomization. Therefore, we can estimate spillovers and peer effects if we can successfully collect endline data.

Finally, we will also incorporate the results of an individual-level online experiment into the paper. We conducted an online experiment with an information treatment arm similar to the door-to-door information campaign. The online experiment included additional information treatment arms, such as an information treatment on the impact of democracy on mitigating the negative impacts of natural disasters, and a placebo treatment group for each information treatment. The placebo treatments allow us to isolate the impact of information on self-reported outcomes. In analyzing the online experiment, all analysis will be at the individual level and we will not cluster the standard errors.

## 5 Hypotheses

### 5.1 Campaign Treatment Effects

Using equation (1), we plan to estimate various treatment effects. They include the average effect of our campaigns ( $T$  is an indicator variable equal to one if the neighborhood is assigned to any of the six treatment groups described above), the effect of the campaign type ( $T = [T^I \ T^S]$ ), the effect by party source ( $T = [T^{CHP} \ T^{IYI} \ T^{IND}]$ ) and the effect by type and party source.

These estimates will allow us to test the hypothesis of whether information on democracy affects voter behavior and whether the source of information (right-wing, left-wing, independent source) matters in shaping voters' voting behavior.

## 6 Heterogeneity Analysis

In analyzing the electoral data, we will estimate heterogeneous treatment effects by quantiles of the opposition vote share. We will explore instrumenting our measure of treatment implementation intensity (aggregated to the ballot box or neighborhood level) with treatment

assignments or using it in our heterogeneity analysis. For the individual level analysis, we will estimate heterogeneous treatment effects using various baseline data like voters' prior party support/ideology and priors on the quality of democratic institutions. We will also consider heterogeneous effects by neighborhood or ballot box voting data.

## References

- [1] Acemoglu, D., Naidu, S., Restrepo, P., Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1), 47-100.
- [2] Baysan, C. (2022). Persistent polarizing effects of persuasion: Experimental evidence from Turkey. *American Economic Review*, 112(11), 3528-3546.
- [3] Bray, M., Kreps, D. M. (1987). Rational learning and rational expectations. *Palgrave Macmillan UK*.
- [4] Casey, K. (2015). Crossing party lines: The effects of information on redistributive politics. *American Economic Review*, 105(8), 2410-2448.
- [5] Casey, K., Glennerster, R., Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *The Quarterly Journal of Economics*, 127(4), 1755-1812.
- [6] Banerjee, A., Chassang, S., Snowberg, E. (2017). Decision Theoretic Approaches to Experiment Design and External Validity. In *Handbook of Field Experiments*, Vol. 1, edited by Abhijit Banerjee and Esther Duflo, 141–74. Amsterdam: Elsevier.
- [7] Cruz, C., Keefer, P., Labonne, J. (2021). Buying informed voters: New effects of information on voters and candidates. *The Economic Journal*, 131(635), 1105-1134.
- [8] Cruz, C., Keefer, P., Labonne, J., Trebbi, F. (2018). Making policies matter: Voter responses to campaign promises (NBER Working Paper No. 24785). National Bureau of Economic Research.
- [9] DellaVigna, S., Kaplan, E. (2007). The Fox News effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3), 1187-1234.
- [10] Ferraz, C., Finan, F. (2008). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2), 703-745.
- [11] Freedom House. (2021). The global decline in democracy has accelerated. Technical report.
- [12] Gentzkow, M., Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280-316.
- [13] Giné, X., Mansuri, G. (2018). Together we will: Experimental evidence on female voting behavior in Pakistan. *American Economic Journal: Applied Economics*, 10(1), 207-235.
- [14] Kendall, C., Nannicini, T., Trebbi, F. (2015). How do voters respond to information? Evidence from a randomized campaign. *American Economic Review*, 105(1), 322-353.
- [15] Levitt, S. D., List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives*, 21(2), 153-174.

- [16] Miller, M. K. (2021). A republic, if you can keep it: Breakdown and erosion in modern democracies. *The Journal of Politics*, 83(1), 198-213.
- [17] Orkin, K., Kreft, B. (2019). Everybody loves a winner: A field experiment providing information on polls in South Africa. Working Paper.
- [18] Platas, M. R., Raffler, P. J. (2021). Closing the gap: Information and mass support in a dominant party regime. *The Journal of Politics*, 83(4), 1619-1634.
- [19] Svolik, M. W. (2019). Polarization versus democracy. *Journal of Democracy*, 30(3), 20-32.