The Effects of Political Knowledge on Voter Turnout: A Nonparametric Bounds Approach

Enrijeta Shino[‡] **Armand Kapllani**[‡‡]

Link to Most Current Version

Abstract

Empirical models find a positive effect of political knowledge on turnout, however this may not reflect the true causal effect. The identification of the average causal effect of political knowledge on turnout is hindered by a key identification issue: the endogenous acquisition of political knowledge. Using survey data from the 2016 American National Election Studies (ANES), we employ a nonparametric bounding method to overcome this identification challenge. This method relies on weak and credible assumptions to partially identify the average treatment effect (ATE) of political knowledge on turnout. Specifically, we provide informative bounds by exploring in a sequential fashion the identifying power of different assumptions. We find that the joint combination of the monotone treatment response, monotone treatment selection, and monotone instrumental variable, allows us to improve, in the sense that the estimated upper bound is lower than the point estimates reported in the literature.

Keywords: political knowledge, turnout, bounds, partial identification

[‡]Department of Political Science, University of North Florida, UNF 1, Jacksonville, FL 32611-7140. PO Box 117140. e-mail: e.shino@unf.edu.

[‡‡]Department of Economics, University of Florida, 322 MAT, Gainesville, Florida 32611-7140. PO Box 117140. Telephone: +1 (352) 392-0382, e-mail: akapllani@ufl.edu.

1 Introduction

The extant literature in voter turnout has shown that voting is conditioned by individual and institutional factors such as psychological and socio-economic, electoral rules, political competition, voter registration, and mobilization efforts (Wolfinger and Rosenstone 1980; Huckfeldt and Sprague 1992; Sniderman, Brody and Tetlock 1993; Rosenstone and Hansen 1993; Highton and Wolfinger 2001; Rolfe 2012; Leighley and Nagler 2013; Burden et al. 2014; Shino and Smith 2018). Who votes matters, as it has implications on the policy outcome and representation. A diverse electorate will induces democratic responsiveness from elected representatives, while "unequal participation spells unequal influence" (Lijphart 1997). Focusing on the individual-level characteristics, the extant empirical literature on turnout has documented that education and political knowledge are key determinants of turnout as they lower the cost of voting (Downs 1957; Wolfinger and Rosenstone 1980; Lassen 2005; Larcinese 2007).

Foundational studies of public opinion have pained a grim picture of political knowledge in the mass public (Lazarsfeld, Berelson and Gaudet 1944; Campbell, Gurin and Miller 1954; Converse 1964; Lau and Redlawsk 2001). Even though, there are different types of political knowledge (Zaller 1992; Barabas et al. 2014) and different scholars have used different ways to measure it (Luskin 1987; Mondak 1999, 2001; Prior and Lupia 2008; Gibson and Caldeira 2009; Pietryka and MacIntosh 2013), the most commonly used measure is the five-item knowledge battery recommended by Delli Carpini and Keeter (1996) (also see, (Hayes 2008; Mutz 2002; Nyhan and Reifler 2010)). Delli Carpini and Keeter (1996, p.10) define political knowledge as "the range of factual information about politics that is stored in long-term memory." This is the working definition of political knowledge used in this study.

Point identifying the average treatment effect of political knowledge on participation is a challenging task because voter's decision to acquire knowledge is not exogenous, as unobserved factors jointly drive both the voter's decision in acquiring knowledge and their political participation.[1] Hence, to identify the average treatment effect of political knowledge on turnout, an instrumental variable approach has been used in the empirical literature. To our knowledge, only two studies have empirically studied the causal effect of political knowledge on turnout. The first study by Lassen (2005), used survey data from a natural experiment decentralization referendum conducted in the city of Copenhagen to point identify the causal effect of being politically informed on turnout. Prior to the referendum, the city of Copenhagen carried out a pilot project of decentralization in four of its fifteen districts. The treatment group consisted of voters living in the districts where the pilot decentralization was implemented and the control group consisted of the voters living in districts where the pilot decentralization project was not implemented. Lassen defines a voter as informed if the voter had an opinion on the decentralization. Lassen corrects for endogeneity by instrumenting information using *assignment to the pilot city district*.[2] Lassen (2005) finds that the average treatment effect of being informed on voter turnout is 20%. However, this study relies on self-reported turnout data, which suffer from over-reporting.

In a different empirical study, Larcinese (2007) used data from British 1997 General Elections to test the causal effect of political knowledge on turnout. To address endogeneity, Larcinese used a control function approach to point identify the causal effect of political knowledge on turnout. Differing from Lassen (2005), Larcinese (2007) used validated turnout data. In addition, to address the key identification issue, endogeneity of political knowledge, Larcinese instruments for knowledge using proxies for the news supply on the British mass media.[3] Larcinese (2007) finds that a voter with the highest level of political knowledge is approximately 33% more likely to vote when compared to those with lowest level of political knowledge.

However, the instrumental variable approach relies on linear/nonlinear response mod-

[1] In Appendix B, following Larcinese (2007) we provide a simple conceptual framework to show why political knowledge knowledge is endogenous.

[2] Lassen uses an IV-Probit and Bivariate Probit to estimate his model and a nearest neighbor propensity matching technique to improve balance between treatment and control groups.

[3] Larcinese (2007) motivates the proposed instruments using Luskin (1990).

els coupled with the assumption that there exists at least one observed instrumental variable that induces variation in political knowledge but does not directly affect turnout. Also, the linear response model forces homogeneity on the linear response function, which in reality does not hold, i.e. it ignores heterogenous effects of political knowledge on turnout. Nonlinear models (e.g. logit or probit) impose strong distributional assumptions on the error component of the model. Hence, the credibility of the estimate depends on the functional form, distributional assumptions, and most importantly the credibility of the instrument(s). Given that political knowledge is endogenous, it confounds the inference of the average treatment effect of political knowledge on turnout. Therefore, we propose using a nonparametric bounds approach based on Manski and Pepper (2000), which explores in a sequential fashion weak and credible assumptions to partially identify the average treatment effect. Differing from the instrumental variable approach, the nonparametric bounds approach has two advantages. First, it depends on relatively weak but credible assumptions. Second, the constructed bounds neither depend on the linearity assumption nor on the instrumental variable assumption.

The bounds we provide are sharp, in the sense that they exhaust all the available information in the data and the assumptions imposed. The main finding of this study is that under a combination of three middle-ground assumptions: monotone treatment response, monotone treatment selection, and monotone instrumental variable, we are able to improve the upper bound of the average treatment effect of political knowledge on turnout, as the estimate we obtain is lower than the point estimates reported in the literature. This study is important as it contributes to the voting behavior literature by providing informative bounds on the effect of political knowledge on turnout, which present a more accurate estimation range of the effect.

The reminder of this study is organized as follows. Section 2 describes the data we use in the analysis. Section 3 introduces the empirical framework and each assumption sequentially. Section 4 discusses estimation and inference. Section 5 discusses the results of the nonparametric bound analysis and section 6 concludes.

2 Data

In this study, we use survey data from the 2016 American National Election Studies (ANES). The 2016 ANES is a national representative sample of the voting age population and data were collected using face-to-face (1,180 responses) and web surveys (3,090 responses). To avoid survey mode effects, in our analysis we use data collected via web only (Ansolabehere and Schaffner 2014; Homola, Jackson and Gill 2016; Malhotra and Krosnick 2007; Shino, Martinez and Binder 2021; Shino and Martinez 2021). Another reason for using the web portion of the survey is the sample size; web survey sample is larger, allowing for more variation and confidence in our estimates.

To create the political knowledge measure, we rely on the standard political questions suggested by Delli Carpini and Keeter (1996), measuring respondent's ability to recall the current offices held by John Roberts (Chief Justice), Joe Biden (Vic President), and Paul Ryan (Speaker), as well as which party held the majority in the House (Republican) and in the Senate (Republican) before the 2016 General Election. The political knowledge variable varies from 0, if respondents did not correctly respond to any or only one of the questions,[4] to 4 if they correctly responded to all five questions. The alpha value for the political knowledge scale is $\alpha = 0.61$.

We use the validate turnout question to make sure that our estimates are not affected by overreporting and social desirability bias. To do so, we merged the 2016 vote validation dataset with the 2016 ANES time series data using respondent ID as the unique identifier. In the final dataset, we include only those respondents who responded to all survey questions used in the analysis. After missing values were removed, our final sample includes 1,500 observations.[5]

[4] We merge these two categories because there were not enough cases for those respondents who did not correctly respond to any of the questions.

[5] In Table A1 Appendix A we provide a summary statistic of our sample.

3 Empirical Framework

Let $\mathcal{J}$ denote the population of voters and let each member $j \in \mathcal{J}$ be endowed by observable covariates $x_j \in \mathcal{X}$ and response function $y_j(\cdot) : \mathcal{T} \to \mathcal{Y}$. The response function maps mutually exclusive and exhaustive treatments $t \in \mathcal{T}$ into outcomes $y_j(t) \in \mathcal{Y}$. In our study the treatments correspond to the political knowledge level of the voter. We have 5 ordered possible treatment levels, from $t = 0$ (min) to $t = 4$ (max). The outcome of interest is turnout, a binary variable, taking unity if the voter cast a ballot and zero otherwise. Let $z_j \in \mathcal{T}$ denote the realized treatment of voter j and let $y_j \equiv y_j(z_j)$ denote the realized outcome. The latter is the actual turnout for the voter who actually received treatment z_j while the latent variable $y_j(t)$ for $t \neq z$ is turnout if the voter were to receive a different treatment. We can only observe the realized treatments (political knowledge) and the realized outcomes (voted or not in 2016 General Election), however, we are unable to observe the response functions.

Our interest is to derive informative bounds on the average treatment effect of political knowledge on turnout, i.e.

$$\Delta(s, t|x \in \mathcal{X}) \equiv \mathbb{P}[y(s) = 1|x \in \mathcal{X}] - \mathbb{P}[y(t) = 1|x \in \mathcal{X}] \quad \forall s > t, \tag{1}$$

and on the probability of turnout $\mathbb{P}[y(t) = 1|x \in \mathcal{X}]$. For ease of notation, when needed we suppress conditioning on observed explanatory variables. The average treatment effect represented in equation (1) has two components: the mean turnout we would observe if all voters have political knowledge level s and the mean turnout we would observe if all voters have political knowledge level t. However, we can only observe the turnout for the actual political knowledge obtained by the voter. To make this point clear we use the law of total probability to decompose $\mathbb{P}[y(t) = 1]$ as follows,

$$\mathbb{P}[y(t) = 1] = \mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z = t) + \mathbb{P}[y(t) = 1|z \neq t]\mathbb{P}(z \neq t). \tag{2}$$

From equation (2), using the sampling process we are able to point identify only $\mathbb{P}[y(t) =$

$1|z = t]$ and $\mathbb{P}(z = t)$. However, the sampling process does not reveal any information regarding the conditional latent probability $\mathbb{P}[y(t) = 1|z \neq t]$ since we are unable to observe the voting outcome $y(t)$ for $z \neq t$. But this does not impede us from deriving informative bounds on $\mathbb{P}[y(t) = 1]$ and also on the average treatment effect, $\Delta(s,t) = \mathbb{P}[y(s) = 1] - \mathbb{P}[y(t) = 1]$, for all $s > t$.

Our analysis proceeds as follows. First, we derive the worst-case bounds on the average treatment effect of political knowledge on turnout, for which we do not impose any assumptions. Then we explore in a sequential fashion weak but credible assumptions to tighten our bounds. All the bounds we estimate are sharp, in the sense that they exhaust all the information in the sample and assumptions imposed.

3.1 Worst-case Bounds

By construction the support of turnout is bounded between zero and one. By the definition of probability the minimum value that the latent conditional probability, $\mathbb{P}[y(t) = 1|z \neq t]$, in equation (2) can take is zero and the maximum value it can take is one. Manski (1989) shows that under a bounded support of the outcome, one can construct a lower and upper bound on $\mathbb{P}[y(t) = 1]$. In our case, using equation (2) and the fact that $\mathbb{P}[y(t) = 1|z \neq t]$ is bounded between zero and one, we derive the following identification region for $\mathbb{P}[y(t) = 1]$:

$$\mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z = t) \leqslant \mathbb{P}[y(t) = 1] \leqslant \mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z = t) + \mathbb{P}(z \neq t). \quad (3)$$

We will refer to the bounds on $P[y(t) = 1]$ as the worst-case bounds since we did not impose any assumptions. All the other bounds we derive, will fall within the range of the worst-case bounds. The latter are not very informative due to the fact that they are wide as we will show in our findings below. In a similar fashion, we derive the worst-case bounds for $\mathbb{P}[y(s) = 1]$ as follows:

$$\mathbb{P}[y(s) = 1|z = s]\mathbb{P}(z = s) \leqslant \mathbb{P}[y(s) = 1] \leqslant \mathbb{P}[y(s) = 1|z = s]\mathbb{P}(z = s) + \mathbb{P}(z \neq s). \quad (4)$$

Combining equations (3) and (4), we obtain the worst-case bounds for the average treatment effect of political knowledge on turnout, $\Delta(s,t)$, by subtracting the lower (upper) bound on $\mathbb{P}[y(t)=1]$ from the upper (lower) bound on $\mathbb{P}[y(s)=1]$ to obtain the upper (lower) bound. That is,

$$
\begin{aligned}
\mathbb{P}[y(s)=1|z=s]\mathbb{P}(z=s) - \mathbb{P}[y(t)=1|z=t]\mathbb{P}(z=t) - \mathbb{P}(z\neq t) \\
\leqslant \Delta(s,t) \leqslant \\
\mathbb{P}[y(s)=1|z=s]\mathbb{P}(z=s) + \mathbb{P}(z\neq s) - \mathbb{P}[y(t)=1|z=t]\mathbb{P}(z=t)
\end{aligned}
\tag{5}
$$

In our analysis, we tighten the worst-case bounds on $\mathbb{P}[y(t)=1]$ and on the average treatment effect, $\Delta(s,t)$, by sequentially exploring the identifying power of the monotone treatment response (hereon MTR), monotone treatment selection (hereon MTS), monotone instrumental variable (hereon MIV), joint MTR+MTS assumption, joint MTR+MIV assumption, and joint MTR+MTS+MIV assumption, in that order.

3.2 Monotone Treatment Response

The first assumption (shape restriction) we explore is the monotone treatment response (MTR), proposed by Manksi (1997). This assumption explores the monotonic relationship that exists between voter's level of political knowledge (treatment) and voting turnout (response function). Before introducing the assumption formally, let s and t be two different levels of political knowledge for voter j with $s>t$. We formally state the MTR assumption as follows:

$$s>t \implies y_j(s)\geqslant y_j(t) \quad \text{for all} \quad j\in\mathcal{J} \tag{6}$$

MTR says that the turnout function is a weakly increasing function of political knowledge. The implication of MTR is that the probability of voting turnout is non-decreasing as the political knowledge of the voter improves,[6] i.e., $\mathbb{P}[y_j(s) = 1|z = q] \geqslant \mathbb{P}[y_j(t) = 1|z = q]$ for any q and any $s > t$. It is worth noting that this implication does not rule out the case where the probability of turnout does not change at all as the political knowledge of the voter increases. However, it is difficult to think of a case where increasing one's level of political knowledge leads to a decrease in the likelihood of voting.

When compared to a linear probability model the MTR assumption is weaker, since in a linear probability model the turnout function is modeled as a linear parametric response function of political knowledge, i.e., $y_j(t) = \alpha + \beta t + u_j$. In the linear probability model, a linear response is imposed which a strong assumption and likely not supported in the population of voters and which can lead to non-credible point estimates.

To apply the MTR assumption, we first break down the population of voters into three subpopulations: (1) $z < t$, voters with political knowledge less than t; (2) $z = t$, voters with political knowledge equal to t; and (3) $z > t$, voters with political knowledge greater than t. For the subpopulation of voters with political knowledge less than t, the MTR assumption implies that the probability of voting is weakly lower than the probability of voting had their level of political knowledge been equal to t. And for the subpopulation of voters with political knowledge greater than t, by MTR, the probability of voting is weakly greater than the probability of voting had their level of political knowledge been equal to t. Hence, to further tighten the bounds on $P[y(t) = 1]$, we use the probability of voting for the subpopulation of voters with political knowledge less than t as a lower bound and with political knowledge greater than t as an upper bound. Formally,

$$\mathbb{P}[y(t) = 1|z \leqslant t]\mathbb{P}(z \leqslant t) \leqslant \mathbb{P}[y(t) = 1] \leqslant \mathbb{P}(z < t) + \mathbb{P}[y(t) = 1|z \geqslant t]\mathbb{P}(z \geqslant t) \tag{7}$$

While to obtain the bounds on the average treatment effect, we subtract the lower (upper) bound on $\mathbb{P}[(y(t) = 1)]$ from the upper (lower) bound on $\mathbb{P}[y(s) = 1]$ to obtain the upper

[6] The proof of this result is shown in Manksi (1997) proposition M1.

(lower) bound. That is,

$$0 \leqslant \Delta(s,t) \leqslant \mathbb{P}(z < s) + \mathbb{P}(y(s) = 1|z \geqslant s)\mathbb{P}(z \geqslant s) - \mathbb{P}(y(t) = 1|z \leqslant t)\mathbb{P}(z \leqslant t). \quad (8)$$

As we show in our findings, by imposing the MTR assumption, we are able to improve the bounds on the average treatment effect when compared to the bounds obtained in the worst-case scenario.

3.3 Exogenous Treatment Selection

Before introducing the second assumption, the monotone treatment selection, in our analysis we also consider the exogenous treatment selection assumption (ETS hereon). The ETS says that the assignment of political knowledge (treatment) to the population of voters is random. That is, for every political knowledge level $t \in \mathcal{T}$,

$$\mathbb{P}[y(t) = 1|z] = \mathbb{P}[y(t) = 1]. \quad (9)$$

In other words, this assumption says that sorting of voters into different levels of political knowledge is exogenous. However, this is a very strong assumption since voters who choose to acquire more political knowledge may be different from the ones acquiring less. For example, the voter who acquires a higher level of political knowledge probably faces a lower cost of acquiring such knowledge when compared to another voter who faces a higher cost. Hence, the ETS assumption is very strong, as it ignores sorting of voters into different levels of political knowledge.

3.4 Monotone Treatment Selection

In this section, we relax the ETS assumption, as not plausible, and introduce the monotone treatment selection assumption, proposed by Manski and Pepper (2000). Formally, the MTS assumption is specified as follows,

$$r > q \implies \mathbb{P}[y(t) = 1|z = r] > \mathbb{P}[y(t) = 1|z = q] \quad \text{for all} \quad t \in \mathcal{T}. \tag{10}$$

The MTS assumption says that voters who acquire higher levels of political knowledge have a higher probability to vote when compared to the ones who acquire less. This is because voters with higher level of political knowledge are potentially different from their counterparts. One can think of the application of the MTS assumption as first partitioning the population of voters intro three subpopulations; (1) $z < t$, (2) $z = t$, and (3) $z > t$, for each political knowledge level t. Then, by the MTS assumption, the probability of voting for the subpopulation of voters with political knowledge less than t is weakly lower than the probability of voting for the voters in the subpopulation with political knowledge equal to t. This implies that the latter (which is observed) serves as an upper bound for the probability of voting for voters with political knowledge level $z < t$ and a lower bound for the probability of voting for voters with political knowledge $z > t$. Then, the bounds on $\mathbb{P}[y(t) = 1]$ are as follows:

$$\mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z \geqslant t) \leqslant \mathbb{P}[y(t) = 1] \leqslant \mathbb{P}(z > t) + \mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z \leqslant t) \tag{11}$$

and the bounds on average treatment effect, $\Delta(s,t)$, are as follows

$$\begin{aligned} &\mathbb{P}[y(t) = 1|z = s]\mathbb{P}(z \geqslant s) - \mathbb{P}(z > t) + \mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z \leqslant t) \\ &\leqslant \Delta(s,t) \leqslant \\ &\mathbb{P}[y(t) = 1|z = s]\mathbb{P}(z \leqslant s) + \mathbb{P}(z > s) - \mathbb{P}[y(t) = 1|z = t]\mathbb{P}(z \geqslant t) \end{aligned} \tag{12}$$

To preview our findings below, by imposing the MTS assumption which is weaker when compared to the ETS assumption we further tighten the bounds on the average treatment effect when compared to the MTR assumption and the worst-case bounds.

3.5 Monotone Instrumental Variable

The instrumental variable approach, is a common approach used to deal with sorting of voters into different levels of political knowledge. Let v be an observable variable and for each value of v we partition the sample of voters into different subpopulations. For each subpopulation, we obtain a lower and upper bound on the effect of political knowledge on turnout. We can use this variation in the bounds across different subpopulations if the variable v satisfies the instrumental variable (hereon IV) assumption (Manski and Pepper 2000). We say that the variable v satisfies the IV assumption if for all treatments $t \in \mathcal{T}$ and for every $k \neq k'$:

$$\mathbb{P}[y(t) = 1|v = k] = \mathbb{P}[y(t) = 1|v = k']. \tag{13}$$

Assumption (13) says that the probability turnout function is independent across different subpopulations defined by v. Under this assumption, the upper bound on $\mathbb{P}[y(t) = 1]|v = k]$ will be the infimum upper bound over all subpopulations defined by $v = k$ and the lower bound will be the supremum lower bound over all subpopulations defined by $v = k$. Formally,

$$\sup_k \left\{ \text{lb}_{\mathbb{P}[y(t)=1]|v=k]} \right\} \leqslant \mathbb{P}[y(t) = 1] \leqslant \inf_k \left\{ \text{ub}_{\mathbb{P}[y(t)=1|v=k]} \right\} \tag{14}$$

where $\text{lb}_{\mathbb{P}[y(t)=1]|v=k]}$ and $\text{ub}_{\mathbb{P}[y(t)=1|v=k]}$ are defined in equation (18). However, finding a variable v that satisfies the IV assumption (13) is not an easy task. For example, Larcinese (2007) used the supply of news on the British media as an instrument for political knowledge. He argues that this is a credible instrument as voters are exogenously exposed to the supply of political information. However, we relax the IV assumption, which assumes mean independence and use a weaker and more credible assumption, the monotone instrumental variable (hereon MIV) proposed by Manski and Pepper (2000). The MIV says that for each level of political knowledge exists a weakly monotone relationship between our instrument

v and the probability of the turnout function. A key difference between the monotone instrumental variable and instrumental variable assumption, is that the former allows for a direct impact on turnout. We formally state the MIV assumption as follows,

$$k'' > k > k' \implies \mathbb{P}[y(t) = 1|v = k''] \geqslant \mathbb{P}[y(t) = 1|v = k] \geqslant \mathbb{P}[y(t) = 1|v = k'] \tag{15}$$

First, we partition the sample space using different values of the categorical variable v that satisfies equation (15). Then, for each subpopulation formed, we can obtain a lower and upper bound. It is from the MIV assumption that we can infer that the conditional probability $\mathbb{P}[y(t) = 1|v = k]$ cannot take a higher value than the upper bound on $\mathbb{P}[y(t) = 1|v = k'']$ and cannot take a lower value than the lower bound on $\mathbb{P}[y(t) = 1|v = k']$. To obtain a lower bound on $\mathbb{P}[y(t) = 1|v = k]$, we first obtain the lower bounds of all subpopulations with $k' \leqslant k$. Then, we take the supremum over all these lower bounds to obtain the lower bound. In a similar fashion we proceed to obtain an upper bound by first obtaining the upper bounds for all subpopulations with $k \leqslant k''$ and then take the infimum over all these upper bounds. Formally,

$$\sup_{k' \leqslant k} lb_{\mathbb{P}[y(t)|v=k]} \leqslant \mathbb{P}[y(t) = 1|v = k] \leqslant \inf_{k'' \geqslant k} ub_{\mathbb{P}[y(t)|v=k]} \tag{16}$$

Then, by taking $\sum_k P(v = k)$ on the three part inequality above, we can obtain the following sharp bounds on $\mathbb{P}[y(t) = 1]$:

$$\sum_k \mathbb{P}(v = k)\Big\{\sup_{k' \leqslant k} lb_{\mathbb{P}[y(t)|v=k]}\Big\} \leqslant \mathbb{P}[y(t) = 1] \leqslant \sum_k \mathbb{P}(v = k)\Big\{\inf_{k'' \geqslant k} ub_{\mathbb{P}[y(t)|v=k]}\Big\} \tag{17}$$

where

$$\begin{aligned}
\mathbb{P}[y(t)=1] &= \sum_k \mathbb{P}[y(t)=1|v=k]\mathbb{P}(v=k) \quad \text{[By the law of total probability]} \\
lb_{\mathbb{P}[y(t)=1|v=k]} &= \mathbb{P}[y(t)=1|v=k,z=t]\mathbb{P}(z=t|v=k) \\
ub_{\mathbb{P}[y(t)=1|v=k']} &= \mathbb{P}[y(t)=1|v=k',z=t]P(z=t|v=k')+\mathbb{P}(z\neq t|v=k')
\end{aligned} \tag{18}$$

The bounds on the average treatment effect (1) can easily be obtained by subtracting the lower (upper) bound on $P[y(t)=1]$ from the upper (lower) bound on $P[y(s)=1]$ to obtain the upper (lower) bound. As a monotone instrumental variable, we use spousal education, one of the family background characteristics. Also other family characteristics such as parental education have been utilized as instruments (Card 1999) in estimating returns to education. We code the latter in three categories; (1) high school or less, (2) some college, and (3) college or more. Under the monotone instrumental variable assumption, a voter whose spouse has a higher level of education is more likely to vote compared to a voter whose spouse has a lower level of education.

3.6 Monotone Treatment Response + Monotone Treatment Selection

To further tighten our bounds, we combine the monotone treatment response and monotone treatment selection. By imposing both assumptions, we assume that the turnout function is weakly increasing in political knowledge and that voters who acquire higher levels of political knowledge have a higher probability to vote. Then, the identification region[7] for

[7] For a detailed derivation and proof see Manski and Pepper (2000), Proposition 2, Corollary 2.

$\mathbb{P}[y(t)=1]$ is as follows:

$$\sum_{u<t}\mathbb{P}[y(u)=1|z=u]\mathbb{P}(z=t)+\mathbb{P}[y(t)=1|z=t]\mathbb{P}(z\geqslant t)$$
$$\leqslant \mathbb{P}[y(t)=1]\leqslant \tag{19}$$
$$\sum_{u>t}\mathbb{P}[y(u)=1|z=u]\mathbb{P}(z=u)+\mathbb{P}[y(t)=1|z=t]\mathbb{P}(z\leqslant t).$$

To obtain the bounds on the average treatment effect, $\Delta(s,t)$, we subtract the lower (upper) bound on $\mathbb{P}[y(t)=1]$ from the upper (lower) bound on $\mathbb{P}[y(s)=1]$ to obtain the upper (lower) bound. When compared to the worst-case bounds, the bounds under MTR+MTS are much tighter and they are informative as we will show in our findings.

3.7 Monotone Treatment Response + Monotone Instrumental Variable

Another combination of assumptions we consider is the joint MTR and MIV assumptions and we construct bounds on $\mathbb{P}[y(t)=1]$ [8] as follows:

$$\sum_{u}\mathbb{P}(v=u)\Big\{\sup_{k<u}\left[\mathbb{P}(y(t)=1|v=k,t\geqslant z)\mathbb{P}(t\geqslant z|v=k)\right]\Big\}$$
$$\leqslant \mathbb{P}[y(t)=1]\leqslant \tag{20}$$
$$\sum_{u}\mathbb{P}(v=u)\Big\{\inf_{k'\geqslant u}\left[\mathbb{P}(y(t)=1|v=k',t\leqslant z)\mathbb{P}(t\leqslant z|v=k')+\mathbb{P}(t>z|v=k')\right]\Big\}$$

We also derive the bounds on the average treatment effect $\Delta(s,t)$ for all $s>t$. As shown in the findings section, under the joint MTR and MIV we are able to improve when compared to the worst-case bounds.

[8] A detailed derivation is shown in Manski and Pepper (2000) Proposition 2, Corollary 1.

3.8 Monotone Treatment Response + Monotone Treatment Selection + Monotone Instrumental Variable

Finally, we consider jointly the identifying power of MTR, MTS, and MIV assumptions to further tighten our bounds. To construct the bounds using the joint combination of MTR+MTS+MIV combination we proceed in two steps. The *first step* consists of applying the MTR+MTS assumption to derive a lower and upper bound on $\mathbb{P}(y(t) = 1|v = u)$ for the subpopulation of voters where the instrument v takes a value u. While, the *second step* consists of applying the MIV bounds shown in equation (16) using the lower and upper bounds of $\mathbb{P}(y(t) = 1|v = u)$ obtained in the first step. Following this two-step process we derive the following identification region on $\mathbb{P}[y(t) = 1]$:

$$\begin{aligned}
\sum_u \mathbb{P}(v = u)\Big\{ \sup_{k<u} \big[&\mathbb{P}(y(t) = 1|v = k, z < t)\mathbb{P}(z < t|v = k) \\
&+ \mathbb{P}(y(t) = 1|v = k, z = t)\mathbb{P}(z \geqslant t|v = k)\big]\Big\} \\
&\leqslant \mathbb{P}[y(t) = 1] \leqslant \\
\sum_u \mathbb{P}(v = u)\Big\{ \inf_{k' \geqslant u} \big[&\mathbb{P}(y(t) = 1|v = k', z = t)\mathbb{P}(z \leqslant t|v = k') \\
&+ \mathbb{P}(y(t) = 1|z > t, v = k')\mathbb{P}(z > t|v = k')\big]\Big\}
\end{aligned} \tag{21}$$

To derive the bounds on the average treatment effect, $\Delta(s, t)$, we subtract the lower (upper) bound on $P[y(t) = 1]$ from the upper (lower) bound on $P[y(s) = 1]$ to obtain the upper (lower) bound. Exploring a combination of these three plausible assumptions, allows us to improve the upper bound on the average treatment effect when compared to the ETS estimates and the point estimates reported by Lassen (2005) and Larcinese (2007).

4 Estimation and Inference

We estimate the nonparametric bounds using their sample analogs by replacing each of the probabilities, $\mathbb{P}(\cdot)$, with their respective empirical probabilities, $\mathbb{P}_N(\cdot)$, and under standard regularity conditions the empirical probabilities are consistent estimators. For inference we construct 90% confidence intervals proposed by Imbens and Manski (2004) that asymptotically cover the estimates of interest (upper and lower bounds) with fixed probability. [9] The same approach can be used to construct confidence intervals for the ETS point estimates. We construct the Imbens-Manski 90% CI as follows:

$$\text{CI}_{90\%} = [\widehat{lb} - c_{IM} \times \widehat{se}_{lb}, \widehat{ub} + c_{IM} \times \widehat{se}_{ub}] \tag{22}$$

where $\hat{lb}$ and $\hat{ub}$ are the estimated lower and upper bounds, respectively, and $\hat{se}_{lb}$ and $\hat{se}_{ub}$ are their respective estimated standard errors which are obtained from 1,000 bootstrap replications. The parameter c_{IM} in equation (22) is obtained by solving the following nonlinear equation,

$$\Phi\left(c_{IM} + \frac{\widehat{ub} - \widehat{lb}}{\max\{\widehat{se}_{lb}, \widehat{se}_{ub}\}}\right) - \Phi(-c_{IM}) = 0.90 \tag{23}$$

where Φ is the cumulative distribution function of the standard normal distribution. We report the Imbens-Manski 90% confidence intervals for all the nonparametric bounds and the exogenous treatment selection (ETS) point estimates.

[9] Different studies by Horowitz and Manski (2000), Chernozhukov, Hong and Tamer (2007) and Beresteanu and Molinari (2008) propose confidence intervals that instead of covering the parameters of interest (lower and upper bounds) with fixed probability they cover the entire identification region with fixed probability.

5 Results

We present our results by initially comparing the estimated nonparametric bounds on the probability of voting with the point estimates obtained under the ETS assumption, for each level of political knowledge ($t = 0, 1, \ldots, 4$). The point estimates obtained under ETS are equivalent to the estimates obtained when using a linear probability model. The ETS assumption implies that sorting of voters into different levels of political knowledge is exogenous, and that results in point identification as shown in Table 1 Panel A. Focusing on Table 1 Panel A, we observe that the worst case bounds on the probability of voting for each level of political knowledge are very wide and hence not informative. Imposing the MTR assumption, the bounds improve by becoming tighter when compared to the worst-case bounds. Under MTR, the probability of voting for voters with the highest level of political knowledge ($t = 4$), lies between 77.07% and 94.87%, while those with the lowest level of political knowledge ($t = 0$), lies between 4.27% and 77.07%. Imposing the MTS assumption slightly improves the upper bounds on the probability of voting when compared to MTR but most of the improvement is on the lower bounds for political knowledge level $t = 0, 1, 2, 3$.

We now turn our attention to Panel B of Table 1. Under MIV, where spousal education is used as a monotone instrumental variable, the bounds on the probability of voting slightly improve compared to the worst-case bounds shown in Panel A, but they are still wide and not informative. Introducing the joint MTR and MTS, we observe that bounds improve drastically. We find that for voters with the highest level of political knowledge their probability of voting ranges from 77.07% to 84.35%, while for those with the lowest level, bounds lie between 59% and 77.07%. Bounds using the joint MTR and MIV improve relative to the worst-case bounds but they are not as informative as the bounds obtained under the joint MTR and MTS. By adding the MIV assumption to MTR and MTS, we further tighten the bounds. As shown in Table 1 Panel B, the bounds are informative and tighter when compared to all others. Hence, under a combination of MTR+MTS+MIV, we find that voters with the lowest level of political knowledge have a probability of voting that lies between 66.93% and 77.07%, while for the ones with the highest level of political knowledge lies between 77.07% and 83.6%.

Table 1: Nonparametric bounds on the probability of voting for different levels of political knowledge under different assumptions

Panel A

Assumptions	ETS	Imbens-Manski 90% Conf. Int.		Worst-case		Imbens-Manski 90% Conf. Int.		MTR		Imbens-Manski 90% Conf. Int.		MTS		Imbens-Manski 90% Conf. Int.	
		lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub
$\mathbb{P}[y(0)=1]$	0.5981	0.5963	0.6000	0.0427	0.9713	0.0425	0.9715	0.0427	0.7707	0.0425	0.7711	0.5981	0.9713	0.5961	0.9715
$\mathbb{P}[y(1)=1]$	0.6379	0.6364	0.6395	0.0740	0.9580	0.0737	0.9582	0.1167	0.7993	0.1163	0.7998	0.5924	0.9322	0.5911	0.9325
$\mathbb{P}[y(2)=1]$	0.7207	0.7196	0.7217	0.1393	0.9460	0.1390	0.9462	0.2560	0.8413	0.2555	0.8417	0.5857	0.8937	0.5848	0.8941
$\mathbb{P}[y(3)=1]$	0.8169	0.8162	0.8177	0.2380	0.9467	0.2376	0.9469	0.4940	0.8953	0.4935	0.8957	0.5060	0.8770	0.5053	0.8775
$\mathbb{P}[y(4)=1]$	0.8435	0.8428	0.8442	0.2767	0.9487	0.2762	0.9489	0.7707	0.9487	0.7702	0.9489	0.2767	0.8435	0.2762	0.8442

Panel B

Assumptions	MIV		Imbens-Manski 90% Conf. Int.		MTR + MTS		Imbens-Manski 90% Conf. Int.		MTR + MIV		Imbens-Manski 90% Conf. Int.		MTR + MTS + MIV		Imbens-Manski 90% Conf. Int.	
	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub
$\mathbb{P}[y(0)=1]$	0.0647	0.9588	0.0642	0.9590	0.5981	0.7707	0.5962	0.7711	0.0647	0.7707	0.0642	0.7711	0.6693	0.7707	0.6674	0.7711
$\mathbb{P}[y(1)=1]$	0.1382	0.9507	0.1366	0.9511	0.6351	0.7735	0.6337	0.7740	0.2029	0.7993	0.2001	0.8000	0.6754	0.7739	0.6736	0.7743
$\mathbb{P}[y(2)=1]$	0.1465	0.9398	0.1449	0.9403	0.7023	0.7890	0.7015	0.7895	0.3324	0.8413	0.3278	0.8425	0.7101	0.7847	0.7084	0.7852
$\mathbb{P}[y(3)=1]$	0.2380	0.9441	0.2354	0.9445	0.7620	0.8256	0.7614	0.8262	0.5294	0.8887	0.5223	0.8905	0.7618	0.8258	0.7598	0.8268
$\mathbb{P}[y(4)=1]$	0.2767	0.9455	0.2735	0.9459	0.7707	0.8435	0.7702	0.8442	0.7707	0.9470	0.7606	0.9495	0.7707	0.8360	0.7686	0.8372

Note: The Imbens-Manski 90% confidence intervals are shown in parenthesis and the standard errors used to construct the confidence intervals, equation (22), are obtained from 1000 bootstrapped replications and the critical values c_{IM} are obtained from solving the nonlinear equation (23) using the nleqslv function in R.

In Table 2, we present the findings for the most important estimator, the average treatment effect (ATE) of political knowledge on turnout. In other words, we want to estimate the bounds on the changes in probability of voting when political knowledge of the voters is increased from one level to another. We provide the estimated bounds on $\mathbb{P}[y(t)] - \mathbb{P}[y(t-1)]$ for $t = 1, 2, 3, 4$ and $\mathbb{P}[y(t)] - \mathbb{P}[y(0)]$ for $t = 2, 3, 4$ and their respective Imbens-Manski 90% confidence intervals. In Panel C, under the ETS assumption, if we move voters' with the lowest level of knowledge ($t = 0$) to the highest level ($t = 4$), then they are 24.54% more likely to vote. The worst-case bounds on the ATEs are not informative since they are wide. The worst-case bounds on $\Delta(4, 0)$ show that the latter could fall anywhere between -69.47% and 90.60%. By imposing the MTR assumption we improve the bounds on the average treatment effect when compared to the worst-case. For example, when voter's level of knowledge is increased from zero to one, $\Delta(1, 0)$, the average effect falls between 0 and 75.67%, while when increasing it from 3 to 4, $\Delta(4, 3)$, it can lie anywhere between 0 and 45.47%. When we impose the MTS assumption, the upper bounds significantly improve when compared both to the worst-case and MTR, however the lower bounds are not informative as they are negative. The estimate $\Delta(4, 0)$ can be anywhere between -69.47% and 24.54%.

In Panel D, under the MIV assumption, the bounds are slightly tighter when compared to the worst-case bounds, but still they are not informative due to their large width. When combining the identifying power MTR and MTS, the bounds on the ATEs are tighter and informative. For example, $\Delta(4, 0)$ can be anywhere between 0 and 24.54% and $\Delta(1, 0)$ lies between 0 and 17.54%. While, under a combination of MTR+MIV the bounds are informative but less tighter when compared to the ones derived under MTR+MTS. When using a combination of MTR+MTS+MIV, we obtain the tightest and the most informative bounds. Starting with moving voters from one level of political knowledge to the next, we obtain information on the variation of the change in probability of voting along different levels of political knowledge. Under MTR+MTS+MIV, we find that the upper bounds on $\Delta(4, 3)$, $\Delta(3, 2)$, $\Delta(2, 1)$, and $\Delta(1, 0)$ are 7.43%, 11.57%, 10.93%, and 10.45%, respectively. This shows that the maximum change is not very high and it does not have a high variation along the treatments.

The bounds on $\Delta(4,0)$, $\Delta(3,0)$, and $\Delta(2,0)$ are tighter when compared to all the other assumptions we imposed, but more importantly the upper bounds are lower than the ETS estimates (equivalent to OLS estimates). This shows that under MTR+MTS+MIV assumption we have significantly improved, as our upper bounds on $\Delta(4,0)$, $\Delta(3,0)$, and $\Delta(2,0)$ are all lower than the ETS estimates reported in Panel C. Starting with $\Delta(4,0)$, we find that the average effect of moving voters from the lowest to the highest level of political knowledge on turnout is at most 16.67%, approximately 1.5 times lower when compared to the ETS estimate of 24.54% in Panel C. Also, the upper bounds on $\Delta(3,0)$ is 15.65% which is approximately 1.4 times lower than the ETS estimate (21.88%) and the upper bound on $\Delta(2,0)$ is 11.54%, slightly lower than the ETS estimate (12.26%). These upper bounds are not only significantly lower when compared to our ETS estimates but they are also significantly lower than the estimates reported by the previous studies Lassen (2005) and Larcinese (2007). Hence, under the identifying power of middle-ground that impose credible restrictions on the relationship between political knowledge and turnout, we are able to provide informative bounds on the average treatment effect of political knowledge on turnout.

Table 2: Nonparametric bounds on the ATE of political knowledge on voting under different assumptions

Panel C

Assumptions	ETS	Imbens-Manski 90% Conf. Int.		Worst-case		Imbens-Manski 90% Conf. Int.		MTR		Imbens-Manski 90% Conf. Int.		MTS		Imbens-Manski 90% Conf. Int.	
		lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub
$\Delta(4,3)$	0.0266	(0.0255	0.0276)	-0.6700	0.7107	(-0.6705	0.7112)	0.0000	0.4547	(0.0000	0.4552)	-0.6003	0.3375	(-0.6010	0.3384)
$\Delta(3,2)$	0.0962	(0.0949	0.0976)	-0.7080	0.8073	(-0.7085	0.8078)	0.0000	0.6393	(0.0000	0.6399)	-0.3877	0.2913	(-0.3884	0.2923)
$\Delta(2,1)$	0.0828	(0.0810	0.0845)	-0.8187	0.8720	(-0.8191	0.8724)	0.0000	0.7247	(0.0000	0.7252)	-0.3465	0.3013	(-0.3474	0.3027)
$\Delta(1,0)$	0.0398	(0.0374	0.0422)	-0.8973	0.9153	(-0.8977	0.9157)	0.0000	0.7567	(0.0000	0.7571)	-0.3789	0.3340	(-0.3803	0.3361)
$\Delta(4,0)$	0.2454	(0.2433	0.2474)	-0.6947	0.9060	(-0.6951	0.9065)	0.0000	0.9060	(0.0000	0.9063)	-0.6947	0.2454	(-0.6952	0.2475)
$\Delta(3,0)$	0.2188	(0.2167	0.2209)	-0.7333	0.9040	(-0.7338	0.9045)	0.0000	0.8527	(0.0000	0.8531)	-0.4654	0.2788	(-0.4660	0.2809)
$\Delta(2,0)$	0.1226	(0.1204	0.1247)	-0.8320	0.9033	(-0.8324	0.9037)	0.0000	0.7987	(0.0000	0.7991)	-0.3857	0.2955	(-0.3866	0.2976)

Panel D

Assumptions	MIV		Imbens-Manski 90% Conf. Int.		MTR + MTS		Imbens-Manski 90% Conf. Int.		MTR + MIV		Imbens-Manski 90% Conf. Int.		MTR + MTS + MIV		Imbens-Manski 90% Conf. Int.	
	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub	lb	ub
$\Delta(4,3)$	-0.6674	0.7075	(-0.6680	0.7080)	0.0000	0.0815	(0.0000	0.0824)	0.0000	0.4176	(0.0000	0.4176)	0.0000	0.0743	(0.0000	0.0752)
$\Delta(3,2)$	-0.7018	0.7976	(-0.7027	0.7995)	0.0000	0.1233	(0.0000	0.1244)	0.0000	0.5564	(0.0000	0.5581)	0.0000	0.1157	(0.0000	0.1170)
$\Delta(2,1)$	-0.8042	0.8016	(-0.8066	0.8035)	0.0000	0.1539	(0.0000	0.1554)	0.0000	0.6384	(0.0000	0.6404)	0.0000	0.1093	(0.0000	0.1107)
$\Delta(1,0)$	-0.8205	0.8860	(-0.8233	0.8887)	0.0000	0.1754	(0.0000	0.1774)	0.0000	0.7346	(0.0000	0.7373)	0.0000	0.1045	(0.0000	0.1062)
$\Delta(4,0)$	-0.6821	0.8807	(-0.6826	0.8813)	0.0000	0.2454	(0.0000	0.2475)	0.0000	0.8823	(0.0000	0.8823)	0.0000	0.1667	(0.0000	0.1688)
$\Delta(3,0)$	-0.7208	0.8794	(-0.7217	0.8800)	0.0000	0.2275	(0.0000	0.2295)	0.0000	0.8240	(0.0000	0.8242)	0.0000	0.1565	(0.0000	0.1586)
$\Delta(2,0)$	-0.8123	0.8751	(-0.8145	0.8757)	0.0000	0.1909	(0.0000	0.1929)	0.0000	0.7766	(0.0000	0.7769)	0.0000	0.1154	(0.0000	0.1176)

Note: The Imbens-Manski 90% confidence intervals are shown in parenthesis and the standard errors used to construct the confidence intervals, equation (22), are obtained from 1000 bootstrapped replications and the critical values c_{IM} are obtained from solving the nonlinear equation (23) using the nleqslv function in R.

6 Conclusion

A positive association between political knowledge and turnout does not necessarily reflect a causal mechanism, as unobservables that affect the voter to acquire political knowledge also affect turnout, and as such confound inference. Hence, to address the endogeneity of political knowledge, different studies take an instrumental variable approach that primarily relies on the credibility of the instruments employed. In this paper, different from the current literature, we employ a nonparametric bounds approach and explore the identifying power of different assumptions to construct informative bounds on the average treatment effect of political knowledge on turnout. This approach relies on weaker and more credible assumptions when compared to the assumptions imposed by instrumental variable approach or ordinary least squares.

We show that under the joint combination of the monotone treatment response, monotone treatment selection, and monotone instrumental variable the bounds on the average treatment effect of political knowledge on turnout are informative, meaning that the derived upper bounds are lower than the point estimates obtained under the ETS assumption and the ones reported in the empirical literature under the instrumental variable approach. Under MTR+MTS+MIV, we find that the probability of turnout increases by 16.67% when voter's political knowledge is moved from min to max. The assumptions we explore serve as a middle-ground and allow us to overcome the key identification issue, endogeneity of political knowledge, without imposing linearity, homogeneous effects, or the IV restriction. Hence, the partial identification approach is well-suited in this setting to overcome the challenge of point identification when the assumptions imposed by other conventional methods fail to convince the researchers.

References

Ansolabehere, Stephen and Brian F. Schaffner. 2014. "Does Survey Mode Still Matter? Findings from a 2010 Multi-mode Comparison." *Political Analysis* 22(3):285–303.

Barabas, Jason, Jennifer Jerit, William Pollock and Carlisle Rainey. 2014. "The Question(s) of Political Knowledge." *American Political Science Review* 108(4):840–855.

Beresteanu, Arie and Francesca Molinari. 2008. "Asymptotic Properties for a Class of Partially Identified Models." *Econometrica* 76(4):763–814.

Burden, Barry C., David T. Canon, Kenneth R. Mayer and Donald P. Moynihan. 2014. "Election Laws, Mobilization, and Turnout: The Unanticipated Consequences of Election Reform." *American Journal of Political Science* 58(1):95–109.

Campbell, Angus, Gerald Gurin and Warren E. Miller. 1954. *The Voter Decides*. Row, Peterson, and Co.

Card, David. 1999. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics* 3:1801–1863.

Chernozhukov, Victor, Han Hong and Elie Tamer. 2007. "Estimation and Confidence Regions for Parameter Sets in Econometric Models 1." *Econometrica* 75(5):1243–1284.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics. In Ideology and Discontent, ed. David Apter. New York: Free Press.".

Delli Carpini, Michael X. Delli and Scott Keeter. 1996. *What Americans Know about Politics and Why it Matters*. Yale University Press.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. Harper & Row New York.

Gibson, James L. and Gregory A. Caldeira. 2009. "Knowing the Supreme Court? A Reconsideration of Public Ignorance of the High Court." *The Journal of Politics* 71(2):429–441.

Hayes, Danny. 2008. "Does the Messenger Matter? Candidate-Media Agenda Convergence and its Effects on Voter Issue Salience." *Political Research Quarterly* 61(1):134–146.

Highton, Benjamin and Raymond E. Wolfinger. 2001. "The Political Implications of Higher Turnout." *British Journal of Political Science* 31(1):179–223.

Homola, Jonathan, Natalie Jackson and Jeff Gill. 2016. "A Measure of Survey Mode Differences." *Electoral Studies* 44:255–274.

Horowitz, Joel L. and Charles F. Manski. 2000. "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data." *Journal of the American Statistical Association* 95(449):77–84.

Huckfeldt, Robert and John Sprague. 1992. "Political Parties and Electoral Mobilization: Political Structure, Social Structure, and the Party Canvass." *American Political Science Review* 86(1):70–86.

Imbens, Guido W. and Charles F. Manski. 2004. "Confidence Intervals for Partially Identified Parameters." *Econometrica* 72(6):1845–1857.

Larcinese, Valentino. 2007. "Does Political Knowledge Increase Turnout? Evidence from the 1997 British General Election." *Public Choice* 131(3):387–411.

Lassen, David Dreyer. 2005. "The Effect of Information on Voter Turnout: Evidence from a Natural Experiment." *American Journal of Political Science* 49(1):103–118.

Lau, Richard R. and David P. Redlawsk. 2001. "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* pp. 951–971.

Lazarsfeld, Paul Felix, Bernard Berelson and Hazel Gaudet. 1944. *The People's Choice*. Duell, Sloan & Pearce.

Leighley, Jan E. and Jonathan Nagler. 2013. *Who Votes Now?: Demographics, Issues, Inequality, and Turnout in the United States*. Princeton University Press.

Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma Presidential Address, American Political Science Association, 1996." *American Political Science Review* 91(1):1–14.

Luskin, Robert C. 1987. "Measuring Political Sophistication." *American Journal of Political Science* 31(4):856–899.

Luskin, Robert C. 1990. "Explaining Political Sophistication." *Political Behavior* 12(4):331–361.

Malhotra, Neil and Jon A. Krosnick. 2007. "The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to

Internet Surveys with Nonprobability Samples." *Political Analysis* 15(3):286–323.

Manksi, Charles F. 1997. "Monotone Treatment Response." *Econometrica* 65(6):1311–1334.

Manski, Charles F. 1989. "Anatomy of the Selection Problem." *Journal of Human resources* pp. 343–360.

Manski, Charles F. and J. Pepper. 2000. "Monotone Treatment Response, with an Application to the Returns to Schooling." *Econometrica* 68:997–1012.

Mondak, Jeffery J. 1999. "Reconsidering the Measurement of Political Knowledge." *Political Analysis* 8(1):57–82.

Mondak, Jeffrey J. 2001. "Developing Valid Knowledge Scales." *American Journal of Political Science* 45(1):224–238.

Mutz, Diana C. 2002. "Cross-Cutting Social Networks: Testing Democratic Theory in Practice." *American Political Science Review* 96(1):111–126.

Nyhan, Brendan and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior* 32(2):303–330.

Pietryka, Matthew T. and Randall C. MacIntosh. 2013. "An Analysis of ANES Items and Their Use in the Construction of Political Knowledge Scales." *Political Analysis* 21(4):407–429.

Prior, Markus and Arthur Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American Journal of Political Science* 52(1):169–183.

Riker, William H. and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62(1):25–42.

Rolfe, Meredith. 2012. *Voter Turnout: A Social Theory of Political Participation*. Cambridge University Press.

Rosenstone, Steven J. and John Hansen. 1993. *Mobilization, Participation, and Democracy in America*. Macmillan Publishing Company,.

Shino, Enrijeta and Daniel A. Smith. 2018. "Timing the Habit: Voter Registration and Turnout." *Electoral Studies* 51:72–82.

Shino, Enrijeta and Michael D. Martinez. 2021. "The Different Faces of Public Opinion: Is the American Voter Tinted by Mode?" *International Journal of Public Opinion Research* 33(4):756–778.

Shino, Enrijeta, Michael D. Martinez and Michael Binder. 2021. "Determined by Mode? Representation and Measurement Effects in a Dual-mode Statewide Survey." *Journal of Survey Statistics and Methodology* . https://doi.org/10.1093/jssam/smab012.

Sniderman, Paul M., Richard A. Brody and Phillip E. Tetlock. 1993. *Reasoning and Choice: Explorations in Political Psychology*. Cambridge University Press.

Wolfinger, Raymond E. and Steven J. Rosenstone. 1980. *Who Votes?* Yale University Press.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press.

Appendix

A Summary Statistics

Table A1: Summary Statistics

	Treatment levels				
	$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 4$
Average turnout	0.60	0.64	0.72	0.82	0.84
Standard deviation	0.49	0.48	0.45	0.39	0.36
Average spousal education	0.80	0.86	1.16	1.30	1.39
Standard deviation	0.77	0.79	0.79	0.76	0.74
Number of observations	107	174	290	437	492
% Total observations	7.1%	11.6%	19.3%	29.1%	32.8%

The treatment levels a

B Conceptual Framework

To illustrate conceptually why political knowledge is endogenous we follow a simple framework proposed by Larcinese (2007). Assume we have a single-dimensional policy space and there are only two candidates, A and B. Let $\alpha_0 \in \mathbb{R}_{++}$ denote a measure of the distance between the policy platforms p_A and p_B proposed by candidates A and B, respectively. In the classical model of Riker and Ordeshook (1968), the voter's decision of casting a ballot is characterized by the following inequality:

$$U \equiv \mathbb{P} \times \Delta\mathcal{B}(\alpha_0) + \mathcal{D} > \mathcal{C}. \tag{24}$$

The exogenous probability of casting a decisive vote is denoted by $\mathbb{P}$ and $\Delta\mathcal{B}(\alpha_0)$ measures the differential benefit, a function of the distance α_0 between the two proposed policy platforms, that the voter would receive if her preferred candidate wins. $\mathcal{D}$ is a measure of the psychological gains that a voter will obtain when casting a ballot and $\mathcal{C}$ measures the cost of voting.

The voter maximizes her utility (24) with respect to the voting decision and the indirect utility function is as follows:

$$V(\alpha_0) \equiv \max \Big(\mathbb{P} \times \Delta \mathcal{B}(\alpha_0) + \mathcal{D} - \mathcal{C} \Big), \tag{25}$$

where α_0 is unknown to the voter. Let $\overline{\alpha}$ and $\tilde{\alpha}$ be two estimates of α_0, where we will assume that the former is a better estimate compared to the latter. Let $f(\alpha_0)$ denote the probability density function of α_0 and let c be the cost that the voter will face if she acquires more political knowledge and b the non-instrumental benefit of acquiring political knowledge. Then, a voter would acquire more political knowledge if the following inequality holds:

$$\int_{\mathbb{R}_{++}} [V(\overline{\alpha};\alpha_0) - V(\tilde{\alpha};\alpha_0)] dF(\alpha_0) + b > c. \tag{26}$$

It is the presence of the non-instrumental benefits of acquiring political knowledge in (26) that leads to endogeneity of political knowledge, due to the fact that the non-instrumental benefits and psychological gains are both driven by the same forces.