Voter Turnout and Preference Aggregation*

Kei Kawai†
University of California at Berkeley

Yuta Toyama‡
Northwestern University

Yasutora Watanabe§
Hong Kong University of Science and Technology

September 2016

Abstract

We study how voter turnout affects the aggregation of voter preferences in elections. Given that voting is costly, election outcomes disproportionately aggregate the preferences of voters with low voting cost and high preference intensity. We show that the correlation structure among preferences, costs, and perceptions of voting efficacy is identified, and explore how the correlation affects preference aggregation. Using 2004 U.S. Presidential election data, we find that young, low-income, less-educated, and minority voters are underrepresented. All of these groups tend to prefer Democrats except for the less-educated. Democrats would have won 9 more states if all eligible voters turned out.

keyword: *voter turnout, preference aggregation, election*

---

*We thank Heski Bar-Isaac, Alessandra Casella, Matias Iaryczower, Alessandro Lizzeri, Antonio Merlo, Kris Ramsay, Raul Sanchez de la Sierra, Laura Silver, Francesco Trebbi, and Chamna Yoon for helpful comments. Yasutora Watanabe acknowledges the financial support by the Research Grants Council of Hong Kong Government under the General Research Fund Projects 16506015 and 16528116.

†Department of Economics, University of California at Berkeley 530 Evans Hall #3880 Berkeley, CA, 94720-3880. Email: kei@berkeley.edu

‡Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. Email: yuta-toyama@u.northwestern.edu

§Department of Economics, HKUST Business School, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: yasutorawatanabe@gmail.com

1 Introduction

Democracies rely on elections to aggregate the preferences of their citizens. Elections, however, aggregate the preferences of only those that participate. The importance of participation on preference aggregation is documented by studies of suffrage expansion in various contexts such as the abolition of apartheid in South Africa (Kroth et al., 2013), the passage of the Voting Rights Act of 1965 (Husted and Kenny, 1997; Cascio and Washington, 2013), and the passage of women's suffrage laws (Miller, 2008). Less dramatic measures that have reduced the voting costs of certain groups of voters have also been found to affect policy in important ways (Fujiwara, 2015).

While most democracies now enjoy universal suffrage, participation in elections is far from perfect given the voluntary nature of voting. To the extent that the preferences of those that turn out are systematically different from those that do not, election outcomes may poorly aggregate the preferences of all citizens. Thus, how well elections aggregate the preferences of citizens and whose preferences are underrepresented are open questions, even in mature democracies.

The issues of preference aggregation and underrepresentation are also relevant from a policy perspective. The concern that the preferences of certain groups of voters are underrepresented has led some to argue for compulsory voting (see, e.g., Lijphart, 1997). More moderate policy proposals, such as introducing Internet voting, relaxing registration requirements, and making an election day a holiday are also motivated by similar concerns. Understanding how voter turnout affects preference aggregation can thus provide a basis for more informed discussions on these policy proposals.

In this study, we explore the extent to which preferences are aggregated in elections, which hinges on how the preferences, voting costs, and perceptions of voting efficacy are correlated. We show that the joint distribution of these three terms is identified, and estimate it using county-level voting data from the 2004 U.S. Presidential election. We then simulate the counterfactual election outcome under compulsory voting. The difference between the simulated and actual outcomes allows us to quantify the degree to which preferences are aggregated.

We find that young, low-income, and minority citizens, who tend to prefer Democrats, are underrepresented. For example, we find that an electorate that is one year younger than the actual electorate would turn out less by 1.38 percentage points and would vote more for the Democrats by 0.79 percentage points in terms of the two-party vote share. Less-educated and religious citizens, who tend to prefer Republicans, are also underrepresented. In particular, an electorate that has one year less schooling than the actual electorate would turn out less by 6.14 percentage points and vote more for the Republicans by 4.43 percentage points. In the counterfactual experiment in which we let all voters turn out, we find that the Democrats would have won the presidential election by gaining 73 more electors, overturning the election results in 9 states.

The key challenge for studying the effect of turnout on preference aggregation is to identify the correlation between preferences and voting costs in the population. In particular, we need to identify how voter characteristics such as race and income simultaneously determine preferences and costs. However, this is not a straightforward task because a high level of turnout among a particular set of voters may be due to low voting cost or high preference intensity.

To illustrate, consider a plurality rule election in which voters have private values and choose to vote for candidate A, candidate B, or not turn out. Applying a discrete choice framework to the voter's decision, let $u_A(\mathbf{x})$ and $u_B(\mathbf{x})$ denote the utility of voting for candidates A and B, respectively, and $c(\mathbf{x})$ denote the cost of voting (relative to not voting), where $\mathbf{x}$ is a vector of voter characteristics. Then, the voter's mean utilities are as follows:

$$\begin{aligned} V_A(\mathbf{x}) &= u_A(\mathbf{x}) - c(\mathbf{x}), \\ V_B(\mathbf{x}) &= u_B(\mathbf{x}) - c(\mathbf{x}), \text{ and} \\ V_0(\mathbf{x}) &= 0, \end{aligned}$$

where V_0 represents the mean utility of not voting. While one can identify $V_A(\mathbf{x}) = u_A(\mathbf{x}) - c(\mathbf{x})$ and $V_B(\mathbf{x}) = u_B(\mathbf{x}) - c(\mathbf{x})$ by using vote share and turnout data (see Berry (1994) and Hotz and Miller (1993)), $u_A(\cdot)$, $u_B(\cdot)$, and $c(\cdot)$ are not separately identified without further restrictions. This is because making a voter care more about the election outcome (say, by adding an arbitrary function, $g(\mathbf{x})$, to both $u_A(\mathbf{x})$ and $u_B(\mathbf{x})$) is observationally equivalent to lowering the voting cost (by subtracting $g(\mathbf{x})$ from $c(\mathbf{x})$). Even if there are exogenous cost shifters, $\mathbf{z}$ (e.g., rainfall), they do not help separately identify $u_A(\cdot)$, $u_B(\cdot)$, and $c(\cdot)$.[1] Thus, most existing studies impose

[1] Suppose that the cost function is separated into two parts as $c = c_{\mathbf{x}}(\mathbf{x}) + c_{\mathbf{z}}(\mathbf{z})$, where $\mathbf{z}$ is a vector

ad-hoc exclusion restrictions on the way $\mathbf{x}$ enters $u_A(\cdot)$, $u_B(\cdot)$, and $c(\cdot)$, assuming that $\mathbf{x}$ is excluded from either $u_k(\cdot)$ or $c(\cdot)$. Imposing such exclusion restrictions assumes away the correlation structure among these terms and precludes the possibility that the preferences of those who vote are different from those who do not. Note that this identification challenge exists regardless of whether the data are available at the individual level or at the aggregate level.

In this paper we uncover the correlation structure between preferences and costs in a setting in which $\mathbf{x}$ is allowed to enter both $u_k(\cdot)$ and $c(\cdot)$. Our identification is based on the simple observation that, unlike consumer choice problems where choosing not to buy results in the outcome of not obtaining the good, choosing not to turn out still results in either A or B winning the election. This observation implies that the voter's preference depends only on the utility difference between the two election outcomes rather than the levels of utility associated with each outcome.[2] Barkume (1976) first used this observation to separately identify $u_k(\cdot)$ and $c(\cdot)$ in the context of property tax referenda for school districts.

To see how this observation leads to the identification of $u_k(\cdot)$ and $c(\cdot)$, consider the calculus of voting models of Downs (1957) and Riker and Ordeshook (1968). In these models, the utility of voting for candidate k can be expressed as $u_k = pb_k$, where p is the voter's beliefs that she is pivotal, b_A is the utility difference between having candidate A in office and candidate B in office, and b_B is defined similarly.[3] Hence, we have $b_A = -b_B$. The mean utilities can now be expressed as

$$\begin{aligned} V_A(\mathbf{x}) &= pb_A(\mathbf{x}) - c(\mathbf{x}), \\ V_B(\mathbf{x}) &= -pb_A(\mathbf{x}) - c(\mathbf{x}), \text{ and} \\ V_0(\mathbf{x}) &= 0. \end{aligned}$$

The property $b_A = -b_B$ allows us to separately identify preferences and costs. By

of cost shifters that is exluded from $u_A(\cdot)$ and $u_B(\cdot)$. Then, $u_A(\cdot) - c_{\mathbf{x}}(\cdot)$, $u_B(\cdot) - c_{\mathbf{x}}(\cdot)$ and $c_{\mathbf{z}}(\cdot)$ are all separately identified. However, $u_A(\cdot)$, $u_B(\cdot)$ and $c_{\mathbf{x}}(\cdot)$ are not separately identified. See the subsection titled Exogenous Cost Shifters towards the end of Section 4 for details.

[2]This implication holds as long as voters care about the ultimate outcome of the election. However, it may not hold for models in which voters gain utility from the act of voting for a candidate, e.g., models of expressive voting.

[3]More precisely, the utility of voting for candidate k *relative to not turning out* can be expressed as $u_k = pb_k$, by normalizing the utility of not turning out to be zero. See footnote 13 for details.

adding the first two expressions above, we have $V_A(\mathbf{x})+ V_B(\mathbf{x}) = -2c(\mathbf{x})$ because $pb_A(\mathbf{x})$ cancels out. Given that $V_A(\mathbf{x})$ and $V_B(\mathbf{x})$ are both identified from the vote share and turnout data, $c(\cdot)$ is identified. Similarly, we can identify $pb_A(\cdot)$ because we have $V_A(\mathbf{x})- V_B(\mathbf{x}) = 2pb_A(\mathbf{x})$ and $V_A(\mathbf{x})- V_B(\mathbf{x})$ is identified. Although this may appear mechanical, there is a straightforward intuition behind this result. $V_A(\mathbf{x})+ V_B(\mathbf{x})$ is primarily identified by voter turnout and $V_A(\mathbf{x})- V_B(\mathbf{x})$ is primarily identified by the vote share margin. Hence, voter turnout pins down $c(\cdot)$, while the difference in the two-party vote share pins down $pb_A(\cdot)$.

In this paper, we retain the basic structure of the calculus of voting model, but do not place additional restrictions on p such as rational expectations, in which p equals the actual pivot probability. In our model, we interpret p more broadly as the voter's perception of voting efficacy, which is allowed to differ across individuals and is allowed to be correlated with the true pivot probability in a general manner. In particular, we let p be a function of individual characteristics and the state in which the voter lives as $p = p_s \times \widetilde{p}(\mathbf{x})$, where p_s is a state fixed effect that we estimate and $\widetilde{p}(\cdot)$ is a function of voter characteristics, $\mathbf{x}$. By letting p depend on each state, we can take into account the nature of the electoral college system.[4] We show that the ratios of the state-specific components of efficacy, $p_s/p_{s'}$ $(\forall s, s')$, are identified in this model in addition to $\widetilde{p}(\cdot)$, $b_A(\cdot)$ and $c(\cdot)$ (up to a scalar normalization).[5] The ratio $p_s/p_{s'}$ is directly identified from the data without using equilibrium restrictions on p, such as rational expectations. Therefore, our identification and estimation results are agnostic about how voters formulate p.

Given the debate over how to model voter turnout, we briefly review the literature in order to relate our model to the various models of voter turnout.[6] The model that we estimate in this paper is based on the decision theoretic model of voter turnout introduced by Downs (1957) and Riker and Ordeshook (1968). In their models, a

[4]Under electoral college system, perception of voting efficacy may differ significantly across states. For example, electoral outcomes in battleground states such as Ohio were predicted to be much closer than outcomes in party strongholds such as Texas. Hence, we need to allow for the possibility that p is higher for voters in Ohio than for voters in Texas.

[5]More precisely, we can identify $p(\cdot)b_A(\cdot)$ state by state given that we have many counties within each state. Assuming that $\widetilde{p}(\cdot)$ and $b_A(\cdot)$ are common across states, we can identify $p_s/p_{s'}$. We also show that $\widetilde{p}(\cdot)$ and $b_A(\cdot)$ are separately identified up to a scalar multiple in our full specification with county level shocks to preferences and costs.

[6]For a survey of the literature, see, e.g., Dhillon and Peralta (2002), Feddersen (2004), and Merlo (2006).

voter turns out and votes for the preferred candidate if $pb - c + d > 0$, where p is the voter's beliefs over the pivot probability, b is the utility difference from having one's preferred candidate in office relative to the other, c is the physical and psychological costs of voting, and d is the benefit from fulfilling one's civic duty of voting. While the original studies do not endogenize any of these terms, the decision theoretic model has provided a basic conceptual framework for much of the subsequent work on voting and turnout.

Subsequent studies to Riker and Ordeshook (1968) have endogenized or micro-founded each of the terms in the calculus of voting model in various ways. Ledyard (1984) and Palfrey and Rosenthal (1983, 1985) introduce the pivotal voter model in which the pivot probability p is endogenized in a rational expectations equilibrium. They show that there exists an equilibrium with positive turnout in which voters have consistent beliefs about the pivot probability. Coate et al. (2008), however, points out that the rational expectations pivotal voter model has difficulties matching the data on either the level of turnout or the winning margin.[7] Moreover, using laboratory experiments, Duffy and Tavits (2008) finds that voters' subjective pivot probabilities are much higher than the actual pivot probability, which is at odds with the rational expectations assumption.

More recently, there are attempts at endogenizing p in ways other than rational expectations. For example, Minozzi (2013) proposes a model based on cognitive dissonance in the spirit of Akerlof and Dickens (1982) and Brunnermeier and Parker (2005). In his model, voters jointly choose p and whether or not to turn out in order to maximize subjective expected utility. Kanazawa (1998) introduces a model of reinforcement learning in which boundedly rational voters, who cannot compute the equilibrium pivot probabilities, form expectations about p from the correlation between their own past voting behavior and past election outcomes (see also Bendor et al. (2003) and Esponda and Pouzo (2016) for similar approaches). While these models are based on the basic calculus of voting model, the p term in them no longer carries the interpretation of the actual pivot probability.

[7] Note, however, that with aggregate uncertainty Myatt (2012) shows that the level of turnout can still be high with rational expectations. Levine and Palfrey (2007) also shows that combining the quantal response equilibrium with the pivotal voter model can generate high turnout, and finds that the results of laboratory experiments are consistent with the model prediction.

Another strand of the literature endogenizes the c and d terms. Harsanyi (1980) and Feddersen and Sandroni (2006) endogenize the d term by proposing a rule-utilitarian model in which voters receive a warm glow payoff from voting ethically. Based on their approach, Coate and Conlin (2004) estimates a group-utilitarian model of turnout. Shachar and Nalebuff (1999) also endogenizes the d term by considering a follow-the-leader model in which elites persuade voters to turn out. In a paper studying split-ticket voting and selective abstention in multiple elections, Degan and Merlo (2011) considers a model that endogenizes c to reflect the voter's mental cost of making mistakes.

In our paper we bring the calculus of voting model to the data without taking a particular stance on how the p, c, or d terms are endogenized. Specifically, our identification and estimation do not use the restriction that p is equal to the actual pivot probability as in the rational expectations model. Instead, the p term that we recover can be broadly interpreted as the voter's perception of voting efficacy, and can be consistent with a wide class of models including Palfrey and Rosenthal (1983, 1985), Minozzi (2013), and Kanazawa (1998). We purposely aim to be agnostic about the different ways of modeling voter turnout so that the estimated terms can be interpreted as a reduced form of a diverse set of models that endogenize the p, c, or d terms in specific ways. Instead of imposing equilibrium restrictions of a particular model a priori, we let the data directly identify the p, b, and $c - d$ terms.

Relatedly, our study does not impose a priori restrictions on how the covariates should enter the p, b, or $c - d$ terms, allowing, instead, the same set of covariates to affect all three terms. This is important because the way in which covariates enter the p, b, and $c - d$ terms determines the correlation structure among them, which, in turn, determines how well preferences are aggregated. In most existing studies, the sets of covariates that enter the p, b and $c - d$ terms are disjoint, precluding the possibility that preferences and costs are correlated. For example, Coate and Conlin (2004) and Coate et al. (2008) include demographic characteristics only in the b term,[8] while Shachar and Nalebuff (1999) include them in the $c - d$ term. In contrast, we let each demographic characteristic enter all three terms, allowing us to study the effects of turnout on preference aggregation.[9]

[8]To be more precise, Coate and Conlin (2004) and Coate et al. (2008) use demographic characteristics as covariates for the fraction of the population supporting one side.

[9]One possible exception is Degan and Merlo (2011). They consider a model based on the theories of regret in which the cost term is endogenized in a way that captures voters' preferences over candidates.

We use county-level data on voting outcomes from the 2004 U.S. Presidential election to estimate the model.[10] Our data on turnout and vote share incorporate the number of non-citizens and felons to account for the difference between the voting eligible population and the voting age population (McDonald and Popkin, 2001). We also construct the joint distribution of demographic characteristics within each county from the 5% Public Use Microdata Sample of the Census. A benefit of using actual voting data over survey data is that we can avoid serious misreporting issues often associated with survey data such as the overreporting of turnout and reporting bias in vote choice (see, e.g., Atkeson, 1999; DellaVigna et al., 2015).

We find that African Americans, Hispanics, and other minorities have high voting costs as do young, less-educated, low-income, and religious voters. Moreover, young voters have low perception of voting efficacy, which further depresses turnout among this group. Overall, young, less-educated, and low-income voters are particularly underrepresented. In terms of preferences, minority, young, highly educated, low-income, and non-religious voters are more likely to prefer Democrats.

Our results show that, overall, there is a positive correlation between voting cost and preference for Democrats that can be accounted for through observable characteristics. Except for two voter characteristics, namely years of schooling and being religious, we find that demographic characteristics that are associated with higher cost of voting are also associated with preferring Democrats. We also find that unobservable cost shocks are positively correlated with unobservable preference shocks for Democrats. These correlations result in fewer Democratic votes relative to the preferences of the underlying population. We estimate that turnout is significantly lower among the electorate who prefer Democrats to Republicans, at 55.7%, compared with turnout among those who prefer Republicans to Democrats, at 64.5%.

Our findings using actual voting data are broadly consistent with findings based on survey data that document the differences in preferences between voters and non-voters (see, e.g., Citrin et al. (2003), Brunell and DiNardo (2004), Martinez and Gill (2005),

They include the same set of covariates in the c and d terms. In one of their counterfactual analyses, they consider the effect of increasing voter turnout, but with a focus on split-ticket voting and selective abstention across Presidential and Congressional elections.

[10]Although we use aggregate data, we account for the issue of ecological fallacy by computing the behavior of individual voters and aggregating them at the county level.

and Leighley and Nagler (2013)).[11] Moreover, our paper provides an understanding of the mechanism that generates these differences between voters and non-voters through the correlation among preferences, perceptions of efficacy, and voting costs. Our results also shed light on how preference *intensity* affects preference aggregation (see Campbell (1999), Casella (2005) and Lalley and Weyl (2015)).

Regarding our results on the perception of voting efficacy, we find substantial across-state variation in our estimates of p_s, the state fixed effect in the voting efficacy. Furthermore, the estimates are correlated with the ex-post closeness of the election: Battleground states such as Ohio and Wisconsin tend to have high estimates of p_s, while party strongholds such as New Jersey and California have low estimates, which is consistent with the comparative statics of the pivotal voter model with rational expectations. However, the magnitude of the estimated ratio of p_s is at most three for any pair of states. This is in contrast to a much larger variation in the ratio implied by the pivotal voter model.[12] Our results are more consistent with models of turnout in which voters' perception of efficacy are only weakly correlated with the actual pivot probabilities.

In our counterfactual experiment, we simulate the voting outcome under compulsory voting. We find that the vote share of the Democrats increases in all states under compulsory voting. Overall, the increase in the Democrats' two-party vote share is about 3.7%. We also find that the increase in the Democratic vote share would overturn the election results in 9 states including key states such as Florida and Ohio, resulting in the Democrats winning a plurality of the electors.

[11] See also DeNardo (1980) and Tucker and DeNardo (1986) for early works that study the correlation between turnout and the Democratic vote share using aggregate data. For more recent work, see Hansford and Gomez (2010) that uses rainfall as an instrument for turnout.

[12] The pivotal voter model with rational expectations predicts high variation in the ratio of pivot probabilities across states given the winner-take-all nature of the electoral college system. Voters in only a handful of swing states have a reasonable probability of being pivotal (see, e.g., Shachar and Nalebuff, 1999).

2 Model

2.1 Model Setup

Anticipating the empirical application of the paper, we tailor our model to the U.S. Presidential election. Let $s \in \{1, ..., S\}$ denote a U.S. state, and $m \in \{1, ..., M_s\}$ denote a county in state s.

Preference of Voters We consider a model of voting with two candidates, D and R. Each voter chooses to vote for one of the two candidates or not to vote. We let b_{nk} denote voter n's utility from having candidate $k \in \{D, R\}$ in office, p_n denote her perception of voting efficacy, and c_n denote her cost of voting. Given that there are only two possible outcomes (either D wins or R wins the election), the utility of voting for candidate k, U_{nk}, only depends on $b_{nD} - b_{nR}$ rather than b_{nD} and b_{nR} individually:

$$
\begin{aligned}
U_{nD} &= p_n(b_{nD} - b_{nR}) - c_n, && (1)\\
U_{nR} &= p_n(b_{nR} - b_{nD}) - c_n, && (2)\\
U_{n0} &= 0,
\end{aligned}
$$

where U_{n0} is the utility of not turning out, which we normalize to zero.[13] When p_n is the actual pivot probability, our model is the same as the equilibrium of the pivotal voter model of Palfrey and Rosenthal (1983, 1985). However, we interpret p_n broadly as the voter's subjective perception of the voting efficacy as we discuss below. The cost of voting, c_n, includes both the physical and the psychological costs as well as the possible benefits of fulfilling one's civic duty. Hence, c_n can be either positive or negative. When c_n is negative, the voter turns out regardless of the value of p_n and $b_{nD} - b_{nR}$.

We let the preferences of voter n in county m of state s depend on her demographic characteristics, $\mathbf{x}_n$, as

$$b_{nk} = b_k(\mathbf{x}_n) + \lambda_{sk} + \xi_{mk} + \varepsilon_{nk}, \text{ for } k \in \{D, R\},$$

[13] Note that expressions (1) and (2) take the familiar form $pb - c$. This results from normalizing the utility of not turning out to be zero and normalizing the b term by a factor of 2. See pages 29 and 30 of Riker and Ordeshook (1968) for derivation.

where λ_{sk} is a state-specific preference intercept that captures state-level heterogeneity in voter preferences net of the effect of demographic characteristics, $b_k(\mathbf{x}_n)$.[14] ξ_{mk} and ε_{nk} are unobserved random preference shocks at the county level and at the individual level, respectively. ξ_{mk} captures the unobserved factors that affect preferences at the county level, such as the benefits that the voters in county m receive from policies supported by candidate k. Then, the expression for the utility difference is as follows:

$$b_{nD} - b_{nR} = b(\mathbf{x}_n) + \lambda_s + \xi_m + \varepsilon_n,$$

where $b(\mathbf{x}_n) = b_D(\mathbf{x}_n) - b_R(\mathbf{x}_n)$, $\lambda_s = \lambda_{sD} - \lambda_{sR}$, $\xi_m = \xi_{mD} - \xi_{mR}$, and $\varepsilon_n = \varepsilon_{nD} - \varepsilon_{nR}$. We assume that ε_n follows the standard normal distribution.

We also let voting cost c_n be a function of voter n's characteristics as

$$c_n = c(\mathbf{x}_n) + \eta_m,$$

where η_m is a county-level shock on the cost of voting. While our specification does not include state specific constant terms that capture the presence of other election such as gubernatorial and senatorial elections, previous studies (e.g., Smith, 2001) find that neither the presence nor the closeness of those elections affect turnout in presidential elections.[15] We assume that ξ_m and η_m are both independent of $\mathbf{x}_n$, but we allow ξ_m and η_m to be correlated with each other.

We let the voting efficacy term, p_n, depend on both the demographic characteristics of voter n as well as the state in which she votes as follows:

$$p_n = p_s(\mathbf{x}_n) = p_s \times \tilde{p}(\mathbf{x}_n),$$

where p_s is a state specific coefficient that we estimate. It is important to let p_n depend on the state in which the voter votes because of the winner-take-all nature of the electoral votes in each state.[16] For example, in the 2004 Presidential election, a vote in key

[14]While we treat $b_k(\cdot)$ as a primitive, it is possible to put more stucture on voter preferences. One example is a model in which the voter's bliss point is a function of $\mathbf{x}_n$ and the voter's preference depends on the distance between the voter's bliss point and the candidate's ideological position. Our model incorporates such a structure on voter preferences.

[15]The identification of the model does not depend on whether or not we include state specific constant terms in c_n.

[16]In U.S. Presidential elections, the winner is determined by the Electoral College. Each U.S. state

states such as Ohio was predicted to matter considerably more than a vote elsewhere. Our specification also allows for the possibility that p_n depends on voter's characteristics, $\mathbf{x}_n$. Previous work has shown that voter's social and economic status affects her general sense of political efficacy (see, e.g., Karp and Banducci, 2008).

Note that our specification corresponds to the equilibrium of the pivotal voter model with rational expectations if we set p_s equal to the actual pivot probability in state s and set $\tilde{p}(\mathbf{x}_n)$ equal to 1. In this sense, our specification nests the pivotal voter model as a special case. However, instead of imposing the pivotal voter model (and hence placing equilibrium restrictions on p_n), we estimate p_s and $\tilde{p}(\cdot)$ directly from the data. This approach allows us to interpret p_n consistently with models of turnout that endogenize p_n in various ways.

Substituting the expressions for $b_{nD} - b_{nR}$, c_n, and p_n into equations (1) and (2), the utility from choosing each of the alternatives can be expressed as follows:

$$\begin{aligned} U_{nD}(\mathbf{x}_n) &= p_s(\mathbf{x}_n)\left[b_s(\mathbf{x}_n) + \xi_m + \varepsilon_n\right] - c(\mathbf{x}_n) - \eta_m, \\ U_{nR}(\mathbf{x}_n) &= p_s(\mathbf{x}_n)\left[-b_s(\mathbf{x}_n) - \xi_m - \varepsilon_n\right] - c(\mathbf{x}_n) - \eta_m, \\ U_{n0}(\mathbf{x}_n) &= 0, \end{aligned}$$

where $b_s(\mathbf{x}_n)$ denotes $b(\mathbf{x}_n) + \lambda_s$.

A Voter's Decision Voter n's problem is to choose alternative $k \in \{D, R, 0\}$ that provides her with the highest utility:

$$k = \underset{\kappa \in \{D,R,0\}}{\arg\max}\, U_{n\kappa}(\mathbf{x}_n). \tag{3}$$

---

is allocated a number of electoral votes, roughly in proportion to the state's population. The electors of each state are awarded on a winner-takes-all basis in all states except for Maine and Nebraska. The Presidential candidate who wins the plurality of electors' votes becomes the winner of the election.

We can write the probability that voter n votes for candidate D as

$$\begin{aligned}
&\Pr\left(D = \arg\max_{\kappa\in\{D,R,0\}} U_{n\kappa}\right)\\
&\quad = \Pr\left(U_{nD} > U_{nR} \text{ and } U_{nD} > 0\right)\\
&\quad = \Pr\left(\varepsilon_n > -b_s(\mathbf{x}_n) - \xi_m \text{ and } \varepsilon_n > -b_s(\mathbf{x}_n) - \xi_m + \frac{c(\mathbf{x}_n)+\eta_m}{p_s(\mathbf{x}_n)}\right)\\
&\quad = 1 - \Phi\left(\max\left\{-b_s(\mathbf{x}_n) - \xi_m, -b_s(\mathbf{x}_n) - \xi_m + \frac{c(\mathbf{x}_n)+\eta_m}{p_s(\mathbf{x}_n)}\right\}\right),
\end{aligned}$$

where Φ is the CDF of the standard normal. We can derive a similar expression for candidate R.

Figure 1 depicts the behavior of a voter as a function of ε_n. There are two cases to consider: one in which the cost of voting is positive (Case 1) and the other in which the cost of voting is negative (Case 2). In Case 1, a voter with a strong preference for one of the candidates (which corresponds to a large positive realization or a large negative realization of ε_n) votes for her preferred candidate, while a voter who is relatively indifferent between the two candidates does not turn out. That is, a voter with high preference intensity relative to cost turns out, while a voter with low preference intensity does not. In Case 2, a voter always votes regardless of her preference intensity, as the cost of voting is negative.

Vote Share and Voter Turnout We can express the vote share for candidate k in county m, $v_{k,m}$, and the fraction of voters who do not turn out, $v_{0,m}$, as follows:

$$v_{D,m} \equiv \int 1 - \Phi\left(\max\left\{-b_s(\mathbf{x}_n) - \xi_m, -b_s(\mathbf{x}_n) - \xi_m + \frac{c(\mathbf{x}_n)+\eta_m}{p_s(\mathbf{x}_n)}\right\}\right) dF_{\mathbf{x},m}(\mathbf{x}_n), \tag{4}$$

$$v_{R,m} \equiv \int \Phi\left(\min\left\{-b_s(\mathbf{x}_n) - \xi_m, -b_s(\mathbf{x}_n) - \xi_m - \frac{c(\mathbf{x}_n)+\eta_m}{p_s(\mathbf{x}_n)}\right\}\right) dF_{\mathbf{x},m}(\mathbf{x}_n), \tag{5}$$

$$v_{0,m} \equiv 1 - v_{D,m} - v_{R,m} \tag{6}$$

where $F_{\mathbf{x},m}$ denotes the distribution of $\mathbf{x}$ in county m. Denoting the number of eligible voters in county m by N_m and the number of counties in state s as M_s, the vote share for candidate k in state s can be expressed as $\sum_{m=1}^{M_s} N_m v_{k,m} \Big/ \sum_{m=1}^{M_s} N_m$. The candi-

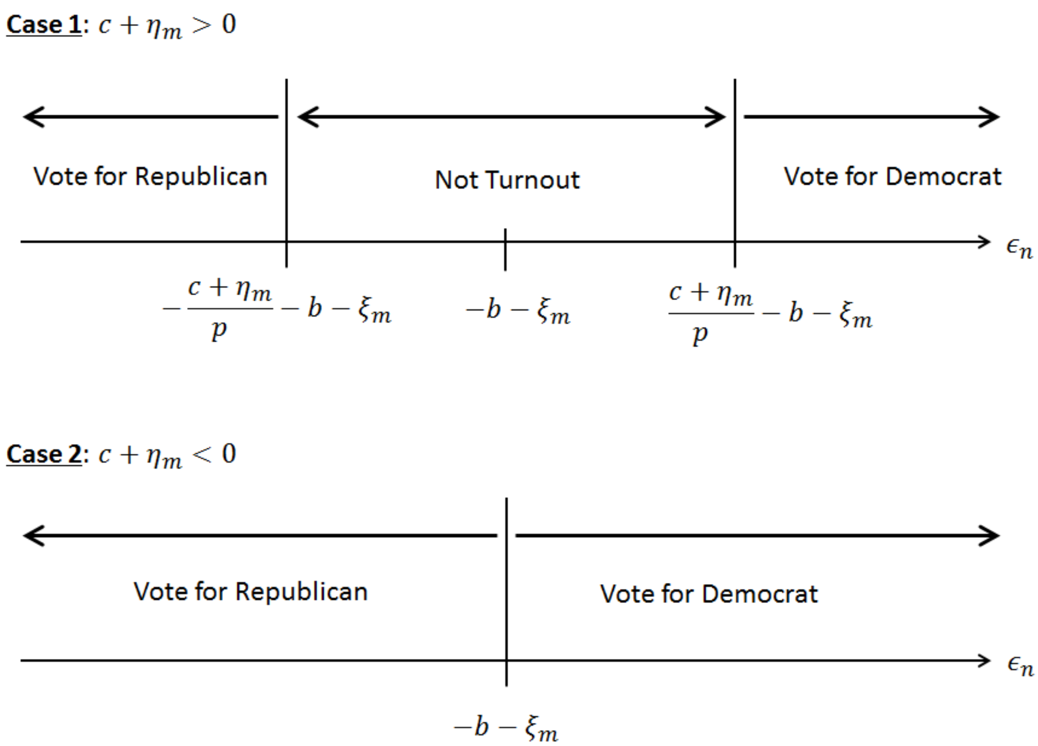

Figure 1: Voter's Decision as a Function of ε_n. The top panel corresponds to the case in which a voter has positive costs of voting. The bottom panel corresponds to the case in which a voter has negative costs of voting.

date with the highest vote share in state s is allocated all of the electors of that state.[17] The candidate who wins the plurality of the electors becomes the overall winner of the Presidential election.

Advertising and Campaign Visits An important feature of Presidential elections not explicitly modeled thus far is the campaign activities of candidates. Candidates target key states with advertisements and campaign visits during the election. These campaign activities are endogenous and depend on the expected closeness of the race in each state (see, e.g., Strömberg, 2008; Gordon and Hartmann, 2013).

While we do not have a specific model of political campaigns, the model accounts for the effect of campaigns on voters through the state fixed effect in the voter's utility,

[17] Maine and Nebraska use a different allocation method. Hence, we drop these two states from our sample. See also footnote 16.

λ_s. Because we treat λ_s as parameters to be estimated, λ_s may be arbitrarily correlated with the closeness of the race in the state. Hence, our estimates of voter preferences are consistent even when campaign activities are endogenous, and our results based on the estimates, such as our discussion underpresentation and preference intensity are not affected. We note, however, that the results of our counterfactual experiment take the level of campaigning as given.

Discussion on Voter's Information One of the factors that we do not specifically model is information. One way to explcitly model information is by endogenizing the voters' information acquisition (see, e.g., Matsusaka 1995, include citation, Degan and Merlo). In these models, the voters decide the amount of information they acquire about the candidates—which helps them form preferences over candidates—by paying the cost of information acquistion. Our specification can be thought of as the indirect utility of these models to the extent that information acquisition costs are functions of voter demographics. In fact, in our estimation results, we find that income and education are associated with low voting costs, which suggests that information acquisition cost may comprise an important part of voting cost (as the opportunity cost of these voters tend to be higher).

Another way to model information is to consider a common value environment in which voters obtain signals about the quality of the candidates (see, e.g., Feddersen and Pesendorfer 1997, 1999). In these models, voter's utility partly consists of the expected quality of the candidates which are computed by conditioning on the event that the voter is pivotal. To the extent that the prior beliefs and the signal distribution depend on the voter's demographic characteristics, the common value component is also a function of these characteristics. Hence, our specification of the utility can be interpreted as capturing a combination of private value and common value components in a reduced-form way.

For these reasons, our estimates and discussions in Section 6 are not affected by abstracting away from modeling voters' information about the candidates. Similar to the remark we made above on advertising and campaigns, however, the results of our counterfactual experiment take the endogenous level of information acquisition or the equilibrium belifes on candidate quality as fixed.

Discussion on p The modeling in our paper is purposely agnostic about how p is endogenized: We do not impose a particular model of p such as rational expectations (Palfrey and Rosenthal, 1983, 1985), overconfidence (Duffy and Tavits, 2008), or cognitive dissonance (Minozzi, 2013). Similarly, our estimation approach avoids using restrictions specific to a particular way of modeling voter beliefs. The important point for our purpose is that there exists an equilibrium p that corresponds to the data generating process regardless of the way in which p is endogenized. Our approach is to identify and estimate both the model primitives and the equilibrium p directly from the data with as little structure as possible. This empirical strategy is similar in spirit to that in the estimation of incomplete models in which some primitives are estimated from the data without fully specifying a model. For example, Haile and Tamer (2003) recovers bidder values without fully specifying a model of the English auction, using only the restriction that the winning bid lies between the valuations of the losers and the winner. Given that their estimation procedure also avoids using restrictions specific to a particular model of the English auction, the estimates are consistent under a variety of models.

In section 4, we show that the key primitives of the model are identified without fully specifying how voters form p. We show that the equilibrium p is also identified directly from the data.[18] The strength of our approach is that we impose few restrictions on beliefs and hence our estimates are consistent under a variety of behavioral assumptions regarding how p is formed. On the other hand, this approach limits the types of counterfactual experiments that we can conduct since we do not specify a particular model regarding p.

3 Data

In this section, we describe our data and provide summary statistics. We use the county-level voting data obtained from David Leip's Atlas of U.S. Presidential Elections. This dataset is a compilation of election data from official sources such as state boards of elections. We merge this dataset with county-level demographics data from the 2000 U.S. Census and county-level population data from the 2004 Annual Esti-

[18] More precisely, $p_s/p_{s'}$ is identified for any s and s', and $\tilde{p}(\cdot)$ is identified up to a scalar normalization. See Section 4 for details.

mates of the Resident Population from the Census Bureau. We construct the data on eligible voters for each county by combining the population estimates from the 2004 Annual Estimates and age and citizenship information from the 2000 Census. We then adjust for the number of felons at the state level using the data from McDonald (2016). Hence, our data account for the difference between the voting age population and voting eligible population (see McDonald and Popkin, 2001).

We construct the joint distribution of voters' demographic characteristics and citizenship at the county level by combining the county-level marginal distribution of each demographic variable and the 5% Public Use Microdata Sample (see Appendix A for details). We augment the Census data with county-level information on religion using the Religious Congregations and Membership Study 2000. In particular, we define the variable Religious using adherence to either "Evangelical Denominations" or "Church of Jesus Christ of Latter-day Saints."

Our data consist of a total of 2,909 counties from 40 states. Because we need a large number of counties within each state to identify the state-specific parameters, p_s and λ_s, we drop states that have fewer than 15 counties. These states are Alaska, Connecticut, District of Columbia, Delaware, Hawaii, Massachusetts, New Hampshire, Rhode Island, and Vermont. In addition, we drop Maine and Nebraska because these two states do not adopt the winner-takes-all rule to allocate electors. We also drop counties with a population below 1,000 because the vote shares and turnout rates for these counties can be extreme due to small population size.[19] Table 1 presents the summary statistics of the county-level vote share, turnout, and demographic characteristics. Note that a Hispanic person may be of any race according to the definition used in the Census.

In order to illustrate the degree to which turnout and expected closeness are related, Figure 2 plots the relationship between the (ex-post) winning margin and voter turnout at the state level. The two variables are negatively correlated, although the fitted line is relatively flat. The slope of the fitted line implies that a decrease in the (ex-post) winning margin of 10 percentage points is associated with an increase in turnout of only about 1.6 percentage points. While the negative correlation may be capturing some of the forces of the rational-expectations pivotal voter model, the flatness of the

[19]In addition, we drop one county, Chattahoochee, GA, as the turnout rate is extremely low (18.8%) relative to all other counties. The turnout rate for the next lowest county is 33%.

Voting Data					
	Obs	Mean	Std. Dev.	Min	Max
Vote Share: Democrat	2,909	0.24	0.08	0.04	0.57
Vote Share: Republican	2,909	0.37	0.09	0.07	0.70
Turnout Rate	2,909	0.60	0.09	0.33	0.89

County Demographics					
	Obs	Mean	Std. Dev.	Min	Max
% Hispanic	2,909	0.05	0.12	0.00	0.97
% Black/African American	2,909	0.09	0.15	0.00	0.87
% Neither Black nor White	2,909	0.05	0.08	0.00	0.95
Mean Age	2,909	46.75	2.64	35.94	56.07
Mean Income (USD 1,000)	2,909	42.71	9.23	23.33	93.40
Mean Years of Schooling	2,909	12.86	0.60	10.84	15.18
% Religious	2,909	0.26	0.17	0.00	1.00

Table 1: Summary Statistics. Voting Outcome and Demographic Characteristics of Eligible Voters. For Age, Income, and Years of Schooling, the table reports the mean, standard deviation, minimum, and maximum of the county mean. "% Religious" is the share of the population with adherence to either "Evangelical Denomination" or "Church of Jesus Christ of Latter-day Saints."

slope suggests that turnout is unlikely to be fully accounted for by the pivotal voter model.

4 Identification

In this section, we discuss the identification of our model as the number of counties *within each state* becomes large ($M_s \to \infty$, $\forall s$). Given that we have state specific parameters for $p_s(\cdot)$ and $b_s(\cdot)$, we require the number of observations for each state to be large. Much of the discussion in this section is based on the idea initially proposed by Barkume (1976) in the context of property tax referenda for school districts.

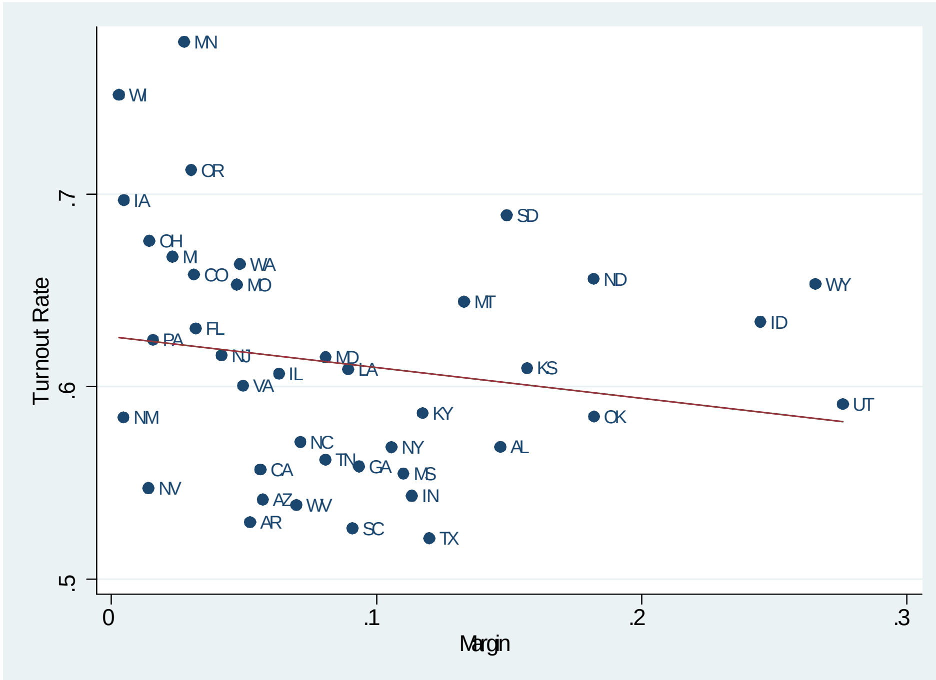

Figure 2: Relationship between the Ex-Post Winning Margin and Voter Turnout. The slope coefficient is -0.16 and not statistically significant.

Recall that the observed vote shares are expressed as:

$$
\begin{aligned}
v_{D,m} &\equiv \int 1 - \Phi\left(\max\left\{-b_s(\mathbf{x}_n) - \xi_m, -b_s(\mathbf{x}_n) - \xi_m + \frac{c(\mathbf{x}_n) + \eta_m}{p_s(\mathbf{x}_n)}\right\}\right) dF_{\mathbf{x},m}(\mathbf{x}_n), \\
v_{R,m} &\equiv \int \Phi\left(\min\left\{-b_s(\mathbf{x}_n) - \xi_m, -b_s(\mathbf{x}_n) - \xi_m - \frac{c(\mathbf{x}_n) + \eta_m}{p_s(\mathbf{x}_n)}\right\}\right) dF_{\mathbf{x},m}(\mathbf{x}_n), \\
v_{0,m} &\equiv 1 - v_{D,m} - v_{R,m}.
\end{aligned}
$$

For exposition, consider the simple case in which there is no heterogeneity in voters' observable characteristics, so that $\mathbf{x}_n = \overline{\mathbf{x}}_m$ for all n in county m.[20] In this case, the

[20]Note that we are well aware of the issues of ecological fallacy. In what follows, we consider a simplified setup with $\mathbf{x}_n = \overline{\mathbf{x}}_m$ for all n in county m, just for expositional purposes. In our empirical exercise, we fully address the fact that each county has a distribution of $\mathbf{x}$ by integrating the vote share for each $\mathbf{x}$ with respect to $F_{\mathbf{x},m}(\mathbf{x})$.

above expressions simplify as follows:

$$\begin{aligned}
v_{D,m} &\equiv 1-\Phi\left(\max\left\{-b_s(\overline{\mathbf{x}}_m)-\xi_m, -b_s(\overline{\mathbf{x}}_m)-\xi_m+\frac{c(\overline{\mathbf{x}}_m)+\eta_m}{p_s(\overline{\mathbf{x}}_m)}\right\}\right), &(7)\\
v_{R,m} &\equiv \Phi\left(\min\left\{-b_s(\overline{\mathbf{x}}_m)-\xi_m, -b_s(\overline{\mathbf{x}}_m)-\xi_m-\frac{c(\overline{\mathbf{x}}_m)+\eta_m}{p_s(\overline{\mathbf{x}}_m)}\right\}\right), &(8)\\
v_{0,m} &\equiv 1-v_{D,m}-v_{R,m}. &(9)
\end{aligned}$$

We now show that the primitives of the model are identified from expressions (7), (8), and (9).

Using the fact that Φ is a strictly increasing function, we can rewrite expressions (7) and (8) as follows:

$$\begin{aligned}
\Phi^{-1}(1-v_{D,m}) &= \max\left\{-b_s(\overline{\mathbf{x}}_m)-\xi_m,\ -b_s(\overline{\mathbf{x}}_m)-\xi_m+\frac{c(\overline{\mathbf{x}}_m)+\eta_m}{p_s(\overline{\mathbf{x}}_m)}\right\},\\
\Phi^{-1}(v_{R,m}) &= \min\left\{-b_s(\overline{\mathbf{x}}_m)-\xi_m,\ -b_s(\overline{\mathbf{x}}_m)-\xi_m-\frac{c(\overline{\mathbf{x}}_m)+\eta_m}{p_s(\overline{\mathbf{x}}_m)}\right\}.
\end{aligned}$$

Rearranging these two equations, we obtain the following expressions:

$$\begin{aligned}
\frac{\Phi^{-1}(1-v_{D,m})+\Phi^{-1}(v_{R,m})}{-2} &= b_s(\overline{\mathbf{x}}_m)+\xi_m, \text{ and} &(10)\\
\frac{\Phi^{-1}(1-v_{D,m})-\Phi^{-1}(v_{R,m})}{2} &= \max\left\{0, \frac{c(\overline{\mathbf{x}}_m)}{p_s(\overline{\mathbf{x}}_m)}+\frac{\eta_m}{p_s(\overline{\mathbf{x}}_m)}\right\}. &(11)
\end{aligned}$$

Note that the left-hand side of (10) closely reflects the difference in vote share, and the left-hand side of (11) reflects the turnout rate. This is because, if we ignore the nonlinearity of $\Phi^{-1}(\cdot)$ and the denominator, the left-hand side of (10) reduces to $1-v_{D,m}+v_{R,m}$ and the left-hand side of (11) to $1-v_{D,m}-v_{R,m}$. The left-hand side of expressions (10) and (11) can be directly computed using data on vote shares, $v_{D,m}$ and $v_{R,m}$.

We first consider the identification of $b_s(\cdot)$ and the distribution of ξ, $F_\xi(\cdot)$. Taking the expectation of (10) conditional on $\overline{\mathbf{x}}_m$, we have

$$\mathbf{E}\left[\left.\frac{\Phi^{-1}(1-v_{D,m})+\Phi^{-1}(v_{R,m})}{-2}\right|\overline{\mathbf{x}}_m\right]=b_s(\overline{\mathbf{x}}_m), \tag{12}$$

because $\mathbf{E}[\xi_m|\overline{\mathbf{x}}_m] = 0$. As the left-hand side of the above expression is identified, $b_s(\cdot)$ is (nonparametrically) identified for each s (note that the asymptotics is with respect to the number of counties within each state). The left-hand side of (12) is simply the mean of the left-hand side of (10), which reflects the vote share difference between Democrats and Republicans, as discussed above. Given that $b_s(\cdot)$ is the utility difference between Democrats and Republicans, it is intuitive that $b_s(\cdot)$ is identified by expression (12).

Now, consider the identification of $F_\xi(\cdot)$. Given that $b_s(\cdot)$ is identified and the left-hand side of (10) is observable, each realization of ξ_m can be recovered from (10). Hence, $F_\xi(\cdot)$ is also identified. Note that if $b_s(\cdot)$ is linear in $\overline{\mathbf{x}}_m$ (i.e., $b_s(\overline{\mathbf{x}}_m) = \boldsymbol{\beta}\overline{\mathbf{x}}_m$), one can simply regress the left-hand side of expression (10) on $\overline{\mathbf{x}}_m$ by OLS to obtain $\boldsymbol{\beta}$ as coefficients and ξ_m as residuals.

We now discuss the identification of $p_s(\cdot)$, $c(\cdot)$, and $F_\eta(\cdot)$. For simplicity, consider the case in which the second term inside the max operator of expression (11) is positive with probability 1, i.e.,

$$\frac{\Phi^{-1}(1 - v_{D,m}) - \Phi^{-1}(v_{R,m})}{2} = \frac{c(\overline{\mathbf{x}}_m)}{p_s(\overline{\mathbf{x}}_m)} + \frac{\eta_m}{p_s(\overline{\mathbf{x}}_m)}. \tag{13}$$

This corresponds to the case in which the turnout rate is always less than 100%. We show in Appendix B that $p_s(\cdot)$, $c(\cdot)$, and $F_\eta(\cdot)$ are identified also for the case in which the turnout rate is allowed to be 100%.

Taking the conditional moments of (13), we have

$$\mathbf{E}\left[\left.\frac{\Phi^{-1}(1 - v_{D,m}) - \Phi^{-1}(v_{R,m})}{2}\right|\overline{\mathbf{x}}_m\right] = \frac{c(\overline{\mathbf{x}}_m)}{p_s(\overline{\mathbf{x}}_m)}, \text{ and} \tag{14}$$

$$\mathbf{Var}\left[\left.\frac{\Phi^{-1}(1 - v_{D,m}) - \Phi^{-1}(v_{R,m})}{2}\right|\overline{\mathbf{x}}_m\right] = \frac{\mathbf{Var}(\eta_m)}{(p_s(\overline{\mathbf{x}}_m))^2}. \tag{15}$$

First, the ratio of $c(\cdot)/p_s(\cdot)$ is identified using (14) because the left-hand side of (14) is identified. Given that the ratio $c(\cdot)/p_s(\cdot)$ is identified for all s and that $p_s(\cdot) = p_s \times \tilde{p}(\cdot)$, we can identify $p_{s'}/p_{s''}$ from the ratio of $c(\overline{\mathbf{x}}_m)/p_{s'}(\overline{\mathbf{x}}_m)$ and $c(\overline{\mathbf{x}}_m)/p_{s''}(\overline{\mathbf{x}}_m)$. Intuitively, the ratio $c(\cdot)/p_s(\cdot)$ is identified in each state by the mean turnout because the left-hand side of (14) is simply the average of the left-hand side of (11), which reflects the voter turnout rate as discussed above. The ratio $p_{s'}/p_{s''}$ is identified by the ratio of mean turnout between two counties across states s' and s'' with the same demographics.

Second, $p_s(\cdot)$, $c(\cdot)$, and $F_\eta(\cdot)$ are identified up to a scalar normalization (up to $\mathbf{Var}(\eta_m)$). Because the left-hand side of (15) is identified, $p_s(\cdot)$ is identified up to $\mathbf{Var}(\eta_m)$. Then, $c(\cdot)$ is identified (also up to $\mathbf{Var}(\eta_m)$) from (14). Given that $p_s(\cdot)$ and $c(\cdot)$ are identified, we can recover the distribution of η_m, $F_\eta(\cdot)$, from (13).

The discussion has thus far been based on the simplified case in which all voters in county m have the same demographic characteristics, i.e., $\mathbf{x}_n = \overline{\mathbf{x}}_m$ for all n in m. As long as there is sufficient variation in $F_{\mathbf{x},m}(\mathbf{x})$, we can recover the vote shares conditional on each $\mathbf{x}$ and apply the identification discussion above.

Correlation between Unobserved Cost and Preference Shocks Our identification makes no assumptions regarding the correlation between the unobservables ξ_m and η_m. As ξ_m and η_m enter separately in (10) and (11), $\xi_m \perp \overline{\mathbf{x}}_m$ and $\eta_m \perp \overline{\mathbf{x}}_m$ are sufficient to identify the unknown primitives on the right-hand side in each equation. Hence, we do not require any restrictions on the joint distribution of ξ_m and η_m. In fact, we can nonparametrically identify the joint distribution of ξ_m and η_m from the joint distribution of the residuals in each equation. In our estimation, we specify the joint distribution of ξ_m and η_m as a bivariate Normal with correlation coefficient ρ.

Exogenous Cost Shifters Lastly, we discuss identification when there exist instruments (e.g., rainfall) that shift the cost of voting but not the preferences of the voters. The point we wish to make is that the existence of exogenous cost shifters are neither necessary nor sufficient for identification.

To illustrate this point, consider the following discrete choice setup with instruments $\mathbf{z}_n$,

$$\begin{aligned} V_A &= u_A(\mathbf{x}_n) - c_{\mathbf{x}}(\mathbf{x}_n) - c_{\mathbf{z}}(\mathbf{z}_n) \\ V_B &= u_B(\mathbf{x}_n) - c_{\mathbf{x}}(\mathbf{x}_n) - c_{\mathbf{z}}(\mathbf{z}_n) \\ V_0 &= 0. \end{aligned}$$

where V_k denotes the mean utility of choosing $k \in \{A, B, 0\}$. Here, $u_A(\mathbf{x}_n)$ is not necessarily equal to $-u_B(\mathbf{x}_n)$ and the cost function is separated into two components,

$c_{\mathbf{x}}(\mathbf{x}_n)$ and $c_{\mathbf{z}}(\mathbf{z}_n)$, where $\mathbf{z}_n$ is a vector of cost shifters excluded from $u_k(\mathbf{x}_n)$. For any arbitrary function $g(\mathbf{x}_n)$, consider an alternative model with $\tilde{u}_k(\mathbf{x}_n) = u_k(\mathbf{x}_n) + g(\mathbf{x}_n)$ ($k \in \{A, B\}$) and $\tilde{c}_{\mathbf{x}}(\mathbf{x}_n) = c_{\mathbf{x}}(\mathbf{x}_n) + g(\mathbf{x}_n)$ as follows:

$$\begin{aligned} V_A &= \tilde{u}_A(\mathbf{x}_n) - \tilde{c}_{\mathbf{x}}(\mathbf{x}_n) - c_{\mathbf{z}}(\mathbf{z}_n) \\ V_B &= \tilde{u}_B(\mathbf{x}_n) - \tilde{c}_{\mathbf{x}}(\mathbf{x}_n) - c_{\mathbf{z}}(\mathbf{z}_n) \\ V_0 &= 0. \end{aligned}$$

Because $\tilde{u}_k(\mathbf{x}_n) - \tilde{c}_{\mathbf{x}}(\mathbf{x}_n) = u_k(\mathbf{x}_n) - c_{\mathbf{x}}(\mathbf{x}_n)$, the two models are observationally equivalent, and thus, $u_k(\cdot)$ and $c_{\mathbf{x}}(\cdot)$ are not separately identified. In particular, the correlation between preferences and costs cannot be identified because this model cannot rule out $c_{\mathbf{x}}(\mathbf{x}_n) = 0$, $\forall \mathbf{x}_n$. This is true even if $\mathbf{z}_n$ has a rich support. Hence, it is *not* the availability of instruments, but rather the observation that we can express $u_A(\mathbf{x}_n) = -u_B(\mathbf{x}_n)$ that identifies the primitives of the model.

5 Specification and Estimation

5.1 Specification

We now specify $b_s(\cdot)$, $c(\cdot)$, $p_s(\cdot)$ and the joint distribution of ξ_m and η_m for our estimation. The function $b_s(\cdot)$, which is the utility difference from having candidates D and R in office, is specified as a function of a state-level preference shock, λ_s, and demographic characteristics, $\mathbf{x}_n$, consisting of age, race, income, religion, and years of schooling:

$$b_s(\mathbf{x}_n) = \lambda_s + \boldsymbol{\beta}_b' \mathbf{x}_n.$$

The intercept, λ_s, is a parameter that we estimate for each state. It captures the state-level preference shock for the Democrats that is unaccounted for by demographic characteristics. Note that the linear specification can be derived from a spatial voting model in which a voter's bliss point is linear in $\mathbf{x}_n$.[21]

[21] To illustrate this point, consider a unidimensional spatial voting model in which candidate D's ideological position is 0, candidate R's position is 1, and a voter's bliss point is $\alpha_n = \beta^{bliss} \mathbf{x}_n$. Under the quadratic loss function, the utility from electing candidate D and R are $-\alpha_n^2$ and $-(1-\alpha_n)^2$,

Voting cost is also specified as a linear function of $\mathbf{x}_n$ as

$$c(\mathbf{x}_n) = \boldsymbol{\beta}_c[1, \mathbf{x}_n']'.$$

We do not specifically model the presence of other elections such as gubernatorial and senatorial elections because previous studies (e.g., Smith, 2001) find that neither the presence nor the closeness of other elections affects turnout in Presidential elections. We also do not include weather-related variables in $c(\cdot)$ because there is insufficient variation in precipitation or temperature on the day of the 2004 Presidential election to affect turnout in a significant way.[22]

We specify the voter's perception of efficacy as $p_s \times \tilde{p}(\mathbf{x}_n)$, where $\tilde{p}(\cdot)$ is a function of her age, income, and years of schooling as follows:[23]

$$\tilde{p}(\mathbf{x}_n) = \exp(\boldsymbol{\beta}_p' \mathbf{x}_n).$$

We normalize $p_s = 1$ for Alabama and normalize $\tilde{p}(\cdot)$ such that $\tilde{p}(\bar{\mathbf{x}}) = 1$ where $\bar{\mathbf{x}}$ is the national average of $\mathbf{x}_n$.[24]

We specify the joint distribution of county-level preference shock ξ and cost shock η as a bivariate normal, $N(0, \Sigma)$, where Σ is the variance covariance matrix with diagonal elements equal to σ_ξ^2, σ_η^2 and off-diagonal elements $\rho\sigma_\xi\sigma_\eta$.

respectively, and the utility difference, $b_s(\mathbf{x}_n)$, is written as $-(1-\alpha_n)^2 + \alpha_n^2 = 2\beta^{bliss}\mathbf{x}_n - 1$. Thus, $b_s(\mathbf{x}_n)$ is linear in $\mathbf{x}_n$ in such a model.

[22] We included weather variables in the simple model that assumes $\mathbf{x}_n = \bar{\mathbf{x}}_m$ (i.e., the demographic characteristics of voters in each county are assumed to be the same within county) and found the coefficients on the weather variables to be small and insignificant.

[23] The set of variables included in $p_s(\mathbf{x}_n)$ is a subset of $\mathbf{x}_n$ that takes continuous values. Here, we do not include dummy variables such as race, and religion. The variation in $c(\cdot)$ changes the utility level additively, while the variation in $p_s(\cdot)$ changes it multiplicatively as $p_s \times b_s$. As dummy variables take only 0 and 1, it is difficult in practice to distinguish whether the effects of those variables are additive or multiplicative. Thus, estimating the model with dummy variables in both cost and efficacy is difficult, and we include only continuous variables in $p_s(\cdot)$.

[24] Note that we need two normalizations. Because we express $p_s(\mathbf{x}_n)$ as $p_s \times \tilde{p}(\mathbf{x}_n)$, we need a scalar normalization on either p_s or $\tilde{p}(\mathbf{x}_n)$. We normalize $p_s = 1$ for Alabama. We also need an additional normalization because $p_s(\cdot)$ is identified only up to the variance of η, i.e., the level of p_s is not identified in our model. Assuming that $\tilde{p}(\bar{\mathbf{x}}) = 1$ eliminates this degree of freedom.

5.2 Estimation

We use the method of moments to estimate the model parameters.[25] Recall that the vote shares (as a fraction of eligible voters) and turnout in county m are given by expressions (4), (5) and (6), where $F_{\mathbf{x},m}$ is the distribution of $\mathbf{x}_n$ in county m. For a fixed vector of the model parameters, $\boldsymbol{\theta} = (\boldsymbol{\beta}_b, \{\lambda_s\}, \boldsymbol{\beta}_c, \{p_s\}, \boldsymbol{\beta}_p, \sigma_\xi, \sigma_\eta, \rho)$, we can compute the moments of expressions (4), (5) and (6) by integrating over ξ and η. Our estimation is based on matching the moments generated by the model with the corresponding sample moments.

Specifically, we define the first and second order moments implied by the model as follows,

$$\begin{aligned}
\hat{v}_{k,m}(\boldsymbol{\theta}) &= \mathbf{E}_{\xi,\eta}[v_{k,m}(\xi,\eta;\boldsymbol{\theta})], \forall k \in \{D,R\}, \\
\hat{v}_{k,m}^{squared}(\boldsymbol{\theta}) &= \mathbf{E}_{\xi,\eta}[v_{k,m}(\xi,\eta;\boldsymbol{\theta})^2], \forall k \in \{D,R\}, \\
\hat{v}_{m}^{cross}(\boldsymbol{\theta}) &= \mathbf{E}_{\xi,\eta}[v_{D,m}(\xi,\eta;\boldsymbol{\theta}) v_{R,m}(\xi,\eta;\boldsymbol{\theta})],
\end{aligned}$$

where $v_{k,m}(\xi,\eta;\boldsymbol{\theta})$ is the vote share of candidate k given a realization of (ξ,η) and parameter $\boldsymbol{\theta}$.[26,27] Denoting the observed vote share of candidate k in county m as $\mathsf{v}_{k,m}$, our objective function, $J(\theta)$, is given by,

$$J(\boldsymbol{\theta}) = \sum_{k=\{D,R\}} \left(\frac{J_{1,k}(\boldsymbol{\theta})}{\widehat{\mathbf{Var}}(\mathsf{v}_{k,m})} + \frac{J_{2,k}(\boldsymbol{\theta})}{\widehat{\mathbf{Var}}(\mathsf{v}_{k,m}^2)} \right) + \frac{J_3(\boldsymbol{\theta})}{\widehat{\mathbf{Var}}(\mathsf{v}_{D,m}\mathsf{v}_{R,m})},$$

[25] In contrast to maximum likelihood estimation, which requires us to solve for the fixed point for ξ and η for each parameter value, the method of moments only requires integration with respect to ξ and η by simulation. The latter is substantially less computationally intensive.

[26] Computing $\hat{v}_{k,m}(\boldsymbol{\theta})$, $\hat{v}_{k,m}^{squared}(\boldsymbol{\theta})$, and $\hat{v}_{m}^{cross}(\boldsymbol{\theta})$ requires integration over (ξ,η). For integration, we use a quadrature with $[5 \times 5]$ nodes and pruning (see Jäckel, 2005) with a total of 21 nodes.

[27] Note that we do not need to know the value of ρ for computing $v_{k,m}(\xi,\eta;\boldsymbol{\theta})$. Thus, for each $\boldsymbol{\theta}\backslash\{\rho\}$ (i.e., the parameters except for ρ), we can compute $v_{k,m}(\xi,\eta;\boldsymbol{\theta})$ and solve for the implied realizations (ξ_m,η_m) that give the observed vote shares in county m. Berry (1994) guarantees that there exists a unique pair of (ξ_m,η_m) that equates $v_{k,m}(\xi,\eta;\boldsymbol{\theta})$ to the observed shares. By computing the correlation between the implied realizations of ξ_m and η_m, we can obtain the value of ρ that is consistent with the observed data. We then impose this value of ρ to compute $\hat{v}_{k,m}(\boldsymbol{\theta}), \hat{v}_{k,m}^{squared}(\boldsymbol{\theta})$, and $\hat{v}_{m}^{cross}(\boldsymbol{\theta})$. Our estmation procedure can be thought of as minimizing the objective function with respect to $\boldsymbol{\theta}\backslash\{\rho\}$ and the estimate of ρ is obtained by computing the correlation between the implied realizations of ξ_m and η_m given the estimate of $\boldsymbol{\theta}\backslash\{\rho\}$.

where

$$
\begin{aligned}
J_{1,k}(\boldsymbol{\theta}) &= \frac{1}{M}\sum_{s=1}^{S}\sum_{m=1}^{M_s}(\hat{v}_{k,m}(\boldsymbol{\theta}) - \mathsf{v}_{k,m})^2, \forall k \in \{D, R\}, \\
J_{2,k}(\boldsymbol{\theta}) &= \frac{1}{M}\sum_{s=1}^{S}\sum_{m=1}^{M_s}(\hat{v}_{k,m}^{squared}(\boldsymbol{\theta}) - \mathsf{v}_{k,m}^2)^2, \forall k \in \{D, R\}, \\
J_3(\boldsymbol{\theta}) &= \frac{1}{M}\sum_{s=1}^{S}\sum_{m=1}^{M_s}\left(\hat{v}_{m}^{cross}(\boldsymbol{\theta}) - \mathsf{v}_{D,m}\mathsf{v}_{R,m}\right)^2.
\end{aligned}
$$

$J_{1,k}$ is the sum of the squared differences between the expectation of the predicted vote share ($\hat{v}_{k,m}(\boldsymbol{\theta})$) and the actual vote share ($\mathsf{v}_{k,m}$).[28] $J_{2,k}$ is the sum of the squared differences between $\hat{v}_{k,m}^{squared}(\boldsymbol{\theta})$ and the squared vote share, $\mathsf{v}_{k,m}^2$. J_3 is the sum of the squared differences between the predicted and actual cross terms. M is the the total number of counties across all states, $\sum_{s=1}^{S} M_s$, and $\widehat{\mathbf{Var}}(z)$ denotes the sample variance of z.

The construction of our objective function closely follows our identification argument. The first moment, $J_{1,k}$, matches the conditional expectation of the vote shares from the model with that from the data. Intuitively, $J_{1,k}$ corresponds to (12) and (14), and pins down $\boldsymbol{\beta}_b$, $\{\lambda_s\}$, $\boldsymbol{\beta}_c/p_s$, $\boldsymbol{\beta}_c/\boldsymbol{\beta}_p$ and $p_s/p_{s'}$, following our identification discussion. The second and third moments, $J_{2,k}$ and J_3, correspond to (15). These moments pin down $\boldsymbol{\beta}_p$,σ_ξ,σ_η and ρ.

6 Results

The set of parameters that we estimate include those common across all states ($\boldsymbol{\beta}_b$,σ_ξ, $\boldsymbol{\beta}_c$,σ_η,$\boldsymbol{\beta}_p$,ρ) and those that are specific to each state ($\{\lambda_s\}$,$\{p_s\}$). Table 2 reports the estimates of the former set, while Tables 6 and 7 collect the parameter estimates of the latter set.

Estimates of $\boldsymbol{\beta}_b$, σ_ξ, $\boldsymbol{\beta}_c$, σ_η, $\boldsymbol{\beta}_p$, and ρ The first column of Table 2 reports the estimates of the preference parameters. The estimate of the constant term in the first

[28] We only use moments based on (4) and (5). The moment based on (6) is redundant because (4), (5) and (6) sum up to one.

column corresponds to the preference of the voter who has $\mathbf{x}_n$ equal to the national average and has λ_s equal to Alabama. We find that Age and Income enter the utility difference, $b_D - b_R$, negatively, implying that young and low income voters are more likely to prefer Democrats. Years of Schooling enters the utility difference positively, thus less educated voters are more likely to prefer Republicans. We also find that Hispanics, African Americans, and Other Races prefer Democrats relative to non-Hispanics and Whites (excluded categories). The Religion variable carries a negative coefficient implying that religious voters prefer Republicans more than do non-religious voters.

	Preference		Cost		Efficacy	
	Estimate	SE	Estimate	SE	Estimate	SE
Age	-0.0185	(0.0036)	-0.0086	(0.0091)	0.0648	(0.0157)
Years of Schooling	0.0788	(0.0160)	-0.2470	(0.0655)	-0.1694	(0.0623)
log(income)	-0.3747	(0.0496)	-0.1996	(0.0825)	-0.0534	(0.1728)
Hispanic	0.2474	(0.0843)	0.0750	(0.0726)		
African American	1.3392	(0.0592)	0.2799	(0.0807)		
Other Races	0.7380	(0.0745)	0.2502	(0.0859)		
Religious	-0.6616	(0.0577)	0.1442	(0.0440)		
Constant	-0.1034	(0.0374)	0.4778	(0.0341)		
Sigma	0.1978	(0.0086)	0.0097	(0.9045)		
Rho			0.0840 (0.0521)			

Table 2: Parameter Estimates. The table reports the parameter estimates of voters' preferences, costs, and perception of voting efficacy. The estimate of the constant term in the first column corresponds to the preference of the voter who has $\mathbf{x_n}$ equal to the national average and has λ_s equal to Alabama. The estimate of the constant term in the second column corresponds to the cost of a voter whose characteristics are set to the national mean. The variable log(income) is the log of income divided by 1000. Standard errors are computed by deriving analytically the asymptotic variance covariance matrix. The standard errors are re ported in parentheses.

In the second column of Table 2, we report the estimates of the cost parameters. Our estimates of the cost correspond to the difference between the cost of voting and the benefit of fulfilling civic duty. Moreover, our cost estimates include not only physical cost but also psychological cost, such as information acquisition cost. The estimate of the constant term in the second column corresponds to the cost of a voter whose characteristics are set to the national mean. We find that Age, Years of Schooling, and Income enter voting cost negatively. This implies that older, more educated, and higher income voters have lower cost of voting. Given that high income and more educated voters tend to have high opportunity cost, the negative coefficients on income and education suggest that information acquisition cost can be an important part of the voting cost. Hispanics, African Americans, and Other Races have higher cost of voting relative to non-Hispanics and Whites. Religious voters also have higher cost of voting compared to non-religious voters.

The third column of Table 2 reports the estimates of the efficacy parameters. We find that Age enters the perception of efficacy positively, while Years of Schooling and Income enter negatively. This implies that older, less-educated, and lower-income citizens tend to have higher perception of efficacy. Given that older voters have lower voting cost as well, they are more likely to be overrepresented compared with young voters. Regarding Years of Schooling and Income, the overall effect on participation depends on the relative magnitudes of the cost and efficacy coefficients. We discuss the net effect, reported in Figure 3, in the next subsection.

The last row of Table 2 reports the estimate of ρ, which is the correlation between unobservable shocks ξ and η. The estimate is positive (0.084), implying that the correlation in the unobservable shocks tends to suppress turnout among voters who prefer Democrats.

Representation and Preference Aggregation To understand how representation and preference vary across demographic groups, panels (a) to (f) in Figure 3 plot a measure of representation (left axis) and a measure of preference for the Democrats (right axis). We construct the representation measure of a group based on Wolfinger and Rosenstone (1980) as follows:

$$\frac{\text{share of the group among those who turn out}}{\text{share of the group among the overall electorate}}$$

Thus, overrepresented groups have representation measures larger than one, while underrepresented groups have measures less than one. The preference measure is the two-party vote share of the Democrats unconditional on turnout.

The panels of Figure 3 show that there is significant differences in representation among demographic groups. For example, panel (a) show that the XXXX year old voters are represented 1.XXXXX times more than overall voters, while 30 year old voters are represented only 0.XXXX times.

In terms of preference aggregation, Panel (a) of Figure 3 shows that young voters are underrepresented *and* they tend to prefer Democrats. Regarding education, panel (b) presents that less educated voters are are underrpresented and prefer Republicans. For example, voters with 12 years of schooling are represented only 0.75 times compared to the overall voters and prefer to vote for Republicans with 58.3% vote share, while those with 16 years of schooling are represented 1.42 times more and votes 54.2% for Democrats. Similarly, voters with low-income, hispanic, African American, and other race are underrepresented and tend to prefer Democrats. To the contrary, religious voters are underreposented and prefer Republicans. These results show that there is a systematic selection in the preferences of those who turn out.

Estimates of State Specific Effects, λ_s and p_s Figure YYY presents the result of the state fixed effects in the voter's utility relative to the fixed effect of Alabama, which we normalize to 0. Larger values imply that the voters in the corresponding state prefer Republicans, net of the effect of demographic characteristics. These state fixed effects may include the inherent preferences of voters and/or the effect of campaign activities by candidates. The figure show that states such as Wyoming, Kansas, and Louisiana have strong preferences for Republicans, while states such as Arkansas and Wisconsin have strong preferences for Democrats. Note also that Democratic strongholds such as New York and California tend to have high estimated values of λ_s, while "red" states such as Georgia and Texas tend to have low estimates. Table XXXX in Online Appendix reports the point estimates and the standard errors.

Figure ZZZZ plots the state specific component of the perception of voting efficacy, p_s, with normalization $p_{Alabama} = 1$. High values of p_s correspond to high perception of voting efficacy after controlling for demographics. The perception of voting efficacy may vary across states, partly reflecting the fact that the electors are determined at the state level. We find that battleground states such as Minnesota, Wisconsin, Ohio and Iowa have some of the highest estimated values of voting efficacy. We also find that states considered as party strongholds such as California and Texas have low estimated values. Table XXXX in Online Appendix reports the point estimates and standard errors. These results suggest a weak positive relationship between perception of voting

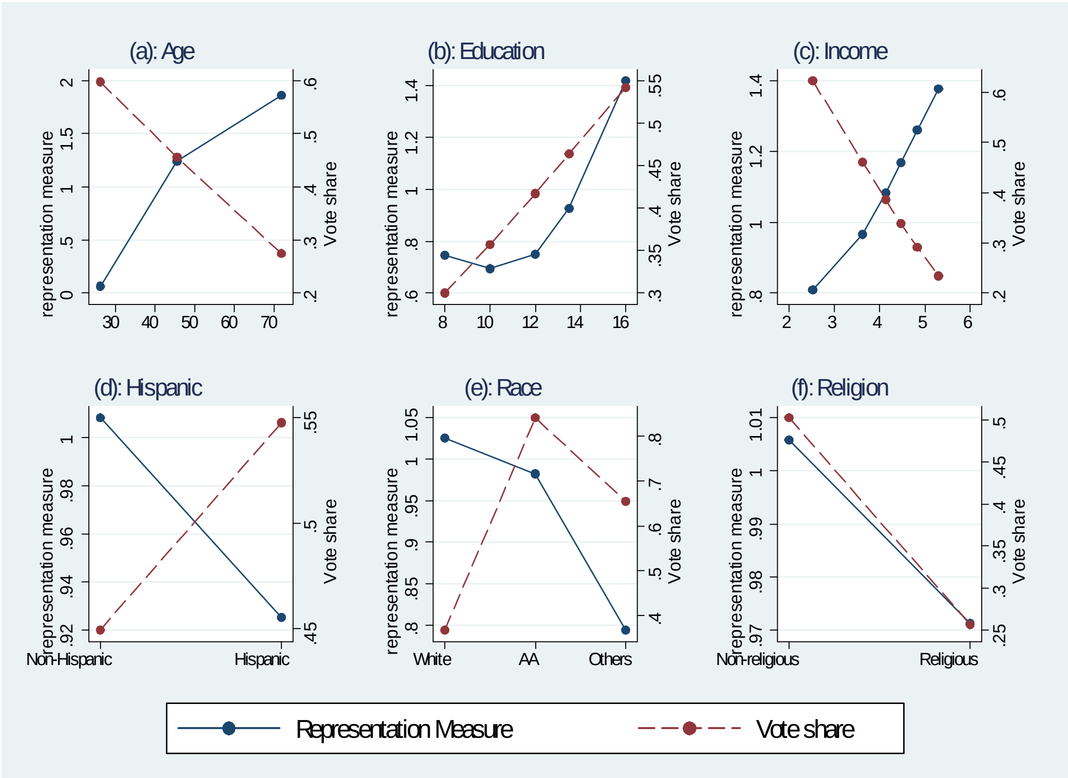

Figure 3: Representation and Preference by Demographic Characteristics. The horizontal axis represents the level and the category of the demographic variable, and the vertical axis corresponds to the representation measure (left axis) and preference of the group in terms of the two-party vote share (right axis). The horizontal axis in Panel (b) is Years of Schooling, and in Panel (c) is log of income. AA in Panel (e) denotes African Americans.

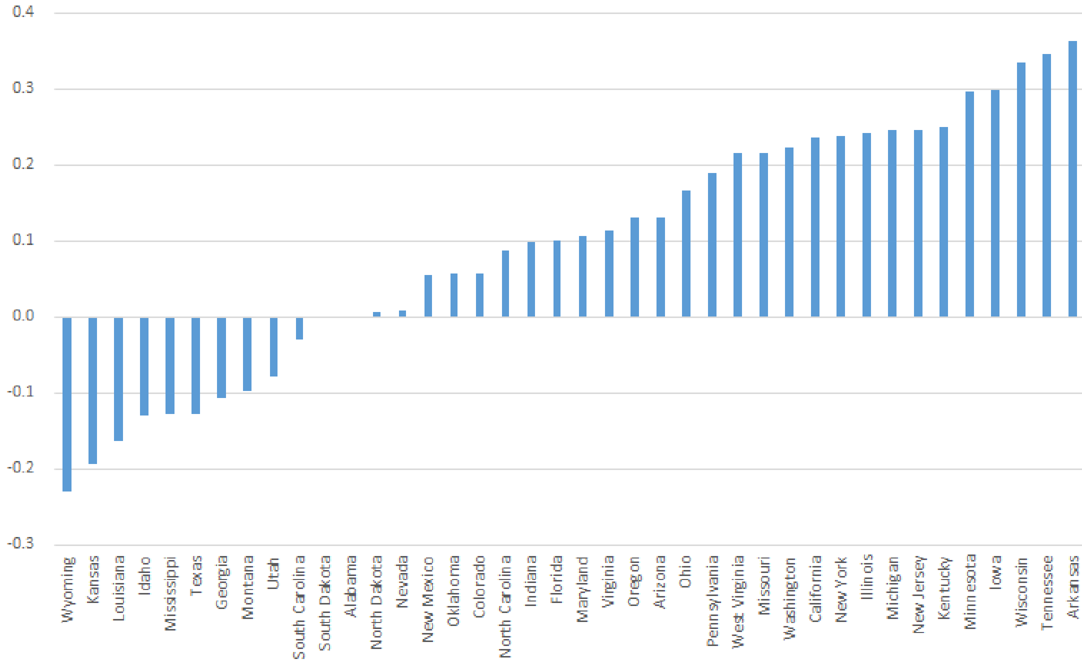

Figure 4: State FE: ADD CPATION!!!

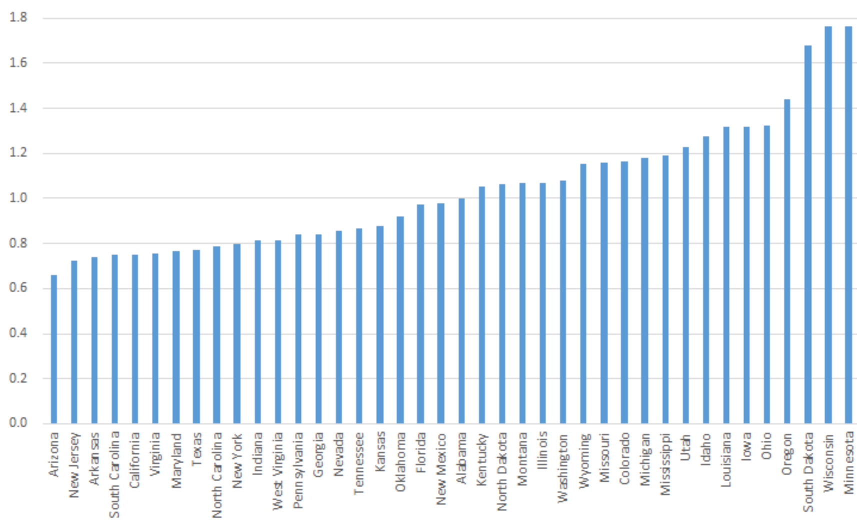

Figure 5: Pivot. ADD CAPTION!!!

efficacy and pivot probability. To illustrate this relationship, Figure 6 plots the estimates of the state-specific coefficient of efficacy, p_s, and the winning margin. The estimated perception of voting efficacy and margin have a negative relationship with a slope of -0.15.

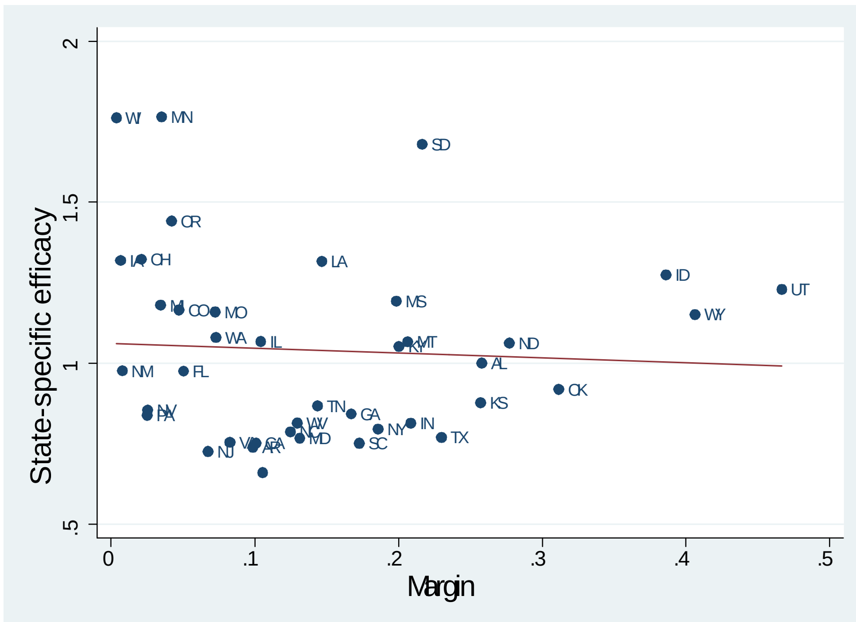

Figure 6: Margin and State-specific Efficacy. The horizontal axis is the winning margin and the vertical axis is the estimate of the state-specific component in voting efficacy. The fitted line has a slope of -0.15.

While some of the forces of the rational expectations pivotal voter model seem to be at play, the estimated values of p_s suggest that the pivotal voter model is unlikely to explain voting behavior very well. Models of voting based on rational expectations would require p_s in battleground states to be of orders of magnitude greater than those in party strongholds (see Shachar and Nalebuff (1999)). However, our estimates of p_s fall within a narrow range: the ratio of the estimated state-level efficacy parameters, $p_s/p_{s'}$, is at most three. This is unlikely under the rational-expectations pivotal voter model. Our results highlight the importance of relaxing the assumption of rational expectations on the pivot probability.

Fit To assess the fit of our model, Figure 7 plots the county-level vote share, voter turnout, and vote share margin predicted from the model against the data. The

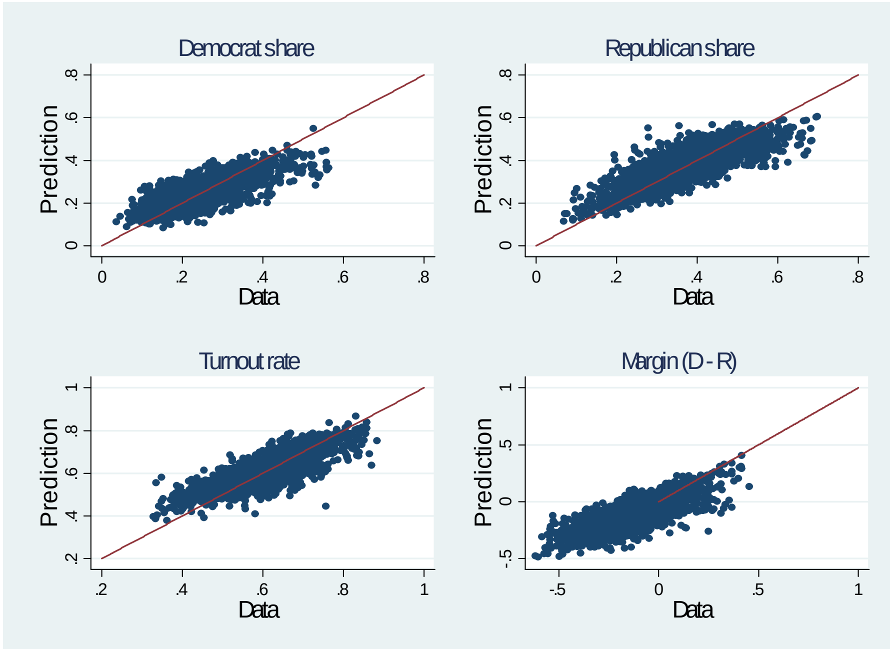

Figure 7: Fit. The figure plots the county-level predicted vote share, turnout, and vote share margin against the data.

predicted vote share, turnout, and vote share margin are computed by integrating out ξ and η. The plots line up around the 45-degree line, which suggests that the model fits the data well. In Online Appendix XXXX, we provide a further discussion of fit.

In previous work, Coate et al. (2008) discusses the difficulty of fitting the winning margin using the rational expectations pivotal voter model. In our paper, we do not impose the rational expectations assumption, and the model fits the winning margin in the data well.

Turnout and Preference Intensity We now discuss how preference intensity is distributed and how it affects preference aggregation. Our discussion is related to the literature that studies how preference intensity affects preference aggregation (see, e.g., Campbell (1999), Casella (2005) and Lalley and Weyl (2015)). In our discussion, we interpret the absolute value of the utility difference, $|b(\mathbf{x}_n)| = |b_s(\mathbf{x}_n)+\varepsilon_n|$, as intensity of preference. Note that our discussion on intensity in this subsection depends on the distributional assumption of idiosyncratic preference error ε_n.

Figure 8 plots the histogram of the utility difference, $b(\mathbf{x}_n)$ (top panel), and the proportion of voters who turn out for given levels of $b(\mathbf{x}_n)$ (bottom panel). Note that we construct the histogram in the top panel by weighting $b(\mathbf{x}_n)$ by the distribution of $\mathbf{x}_n$ in the whole country (and adjust appropriately for the differences in λ_s across states) so that it reflects the preference of all eligible citizens in the U.S. The top panel shows that the distribution of the utility difference is roughly centered around zero, and has a slightly fatter tail on the Democrats' side. The bottom panel shows that turnout is higher among Republican supporters than Democratic supporters at the same level of preference intensity. Overall, we estimate that turnout is significantly lower among the electorate who prefer Democrats over Republicans, at 55.7%, than turnout among those who prefer Republicans over Democrats, at 64.5%.

The panel also shows that there is high turnout among voters with high preference intensity for either party. For example, voters with preferences in the bin [2,2.5) are almost twice as likely to turn out as those with preferences in the bin [0,0.5) (70.3% compared to 36.5%). This implies that voters with intense preferences effectively have "more votes" than those who are indifferent, as pointed out by Campbell (1999) (see also Casella (2005) and Lalley and Weyl (2015)).

7 Counterfactual Experiment

In our counterfactual experiment, we consider what the outcome of the election would be if the preferences of all eligible voters are aggregated. The counterfactual result can be thought of as the election outcome under compulsory voting. In the results section above, we show that preference, voting cost, and perception of efficacy are correlated, and that the preferences of those who turn out do not necessarily reflect the preferences of the general population. In this counterfactual, we quantify the degree to which this discrepancy affects preference aggregation.

We compute the election outcome by setting voting cost to zero in this counterfactual. Individuals vote for Democrats or Republicans depending on the sign of $\widehat{b}_s(\mathbf{x}_n) + \widehat{\xi}_m + \varepsilon_n$, where $\widehat{b}_s(\cdot)$ and $\widehat{\xi}_m$ are the estimates of the net utility difference and county-level preference shock.[29] Hence, we calculate the counterfactual county-

[29]Note that there is a unique value of $(\hat{\xi}_m, \hat{\eta}_m)$ that rationalizes the actual vote outcome given our estimates of preference, cost and perception of efficacy as discussed in footnote 27. We use these values of $(\hat{\xi}_m, \hat{\eta}_m)$ to compute our counterfactual outcome.

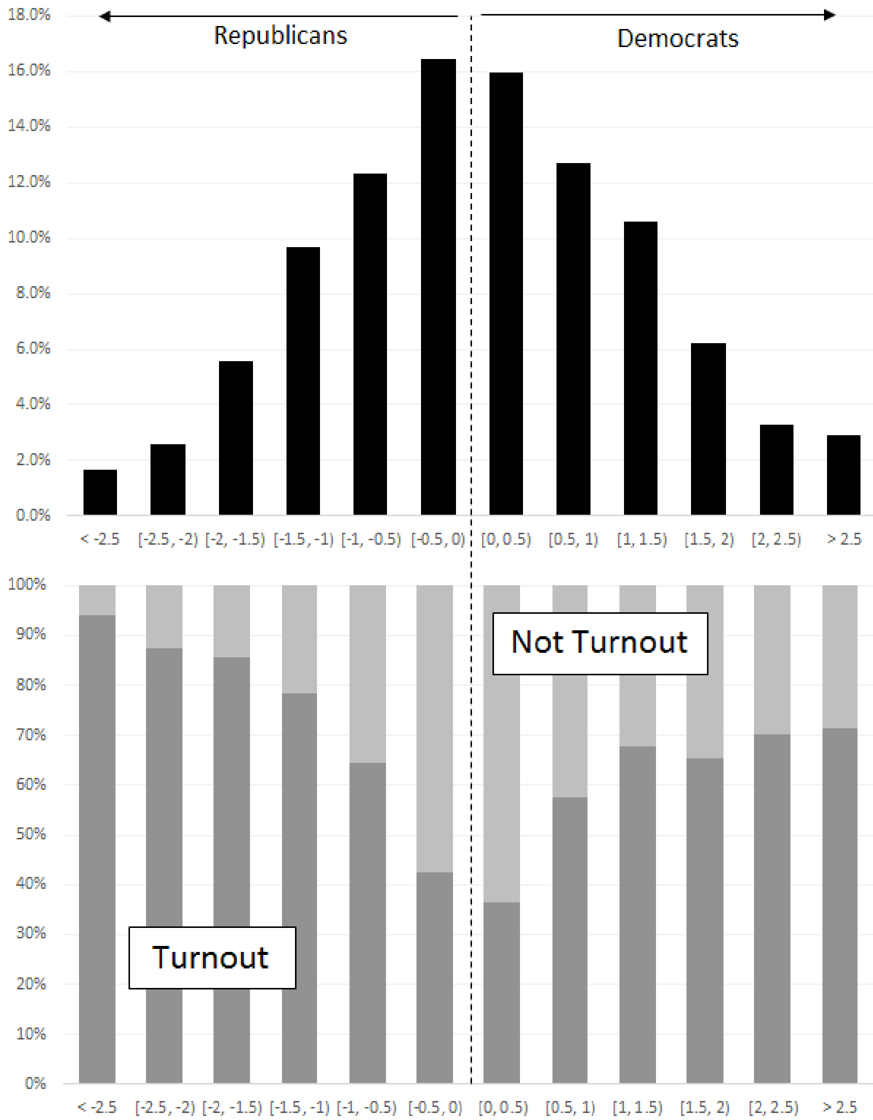

Figure 8: Histogram of Preference Intensity and Fraction of Turnout by Preference Intensity. The top panel plots the histogram of the utility difference, $b(\mathbf{x}_n)$. The bottom panel plots the proportion of those who turn out for given levels of preference intensity.

level vote shares, $\tilde{v}_{D,m}$ and $\tilde{v}_{R,m}$, as

$$\begin{aligned}\tilde{v}_{D,m} &\equiv \int \Phi\left(\widehat{b}_s(\mathbf{x}_n) + \widehat{\xi}_m\right) dF_{\mathbf{x},m}(\mathbf{x}_n),\\ \tilde{v}_{R,m} &\equiv 1 - \tilde{v}_{D,m}.\end{aligned}$$

Note that our counterfactual results are robust to equilibrium adjustments to voters' perception of efficacy because a voter's decision depends only on the sign of the utility difference and perception of efficacy does not affect $\tilde{v}_{D,m}$ and $\tilde{v}_{R,m}$.

	Two-Party Vote Share			# of Electors	
	Democrats	Republicans	Turnout Rate	Democrats	Republicans
Actual	48.2%	51.8%	60.1%	208	278
Counterfactual	51.9%	48.1%	100.0%	310	176
	(1.1%)	(1.1%)	n.a.	(30.5)	(30.5)

Table 3: Counterfactual Outcome Under Full Turnout. The table compares the actual outcome with the counterfactual outcome in which all voters turn out. The reported outcomes do not include the results for the 11 states that we drop from the sample. Standard errors are reported in parentheses.

Table 3 compares the actual outcomes with the counterfactual outcomes for the 40 states in our sample. The first row of Table 3 reports the actual vote share, turnout rate, and number of electors for the two parties. We report our counterfactual results in the second row. We find that the Democratic vote share in the counterfactual increases from 48.2% to 51.9%, reflecting our earlier finding that the preference for Democrats and voting costs are positively correlated.

In terms of electors, we find that the Democrats increases the number of electors by 102 from 208 to 310. Although 10 states and the District of Columbia (D.C.) are not included in our sample, the number of 310 electors is larger than that needed to win the election (270) even if those excluded states and D.C. vote for the Republicans electors.[30] Hence, our estimates suggest that the Democrats would likely have won the 2004 Presidential election if the preferences of all voters had been aggregated.

[30]There are a total of 538 electors including the states that are excluded from our sample. A candidate needs 270 electors to win.

	Two-Party Vote Share of Democrats			Turnout Rate	# of Electors for Democrats			Prob(Democrats win in counterfactual)
	Actual	Counterfactual		Actual	Actual	Counterfactual		
Alabama	37.1%	43.0%	(1.2%)	56.9%	0	0	(0.0)	0.0%
Arizona	44.7%	50.0%	(1.4%)	54.1%	0	10	(5.0)	48.8%
Arkansas	45.1%	50.2%	(1.4%)	53.0%	0	6	(3.0)	52.9%
California	55.0%	58.1%	(1.4%)	55.7%	55	55	(0.0)	100.0%
Colorado	47.6%	50.5%	(0.9%)	65.8%	0	9	(4.3)	65.3%
Florida	47.5%	51.2%	(1.1%)	63.0%	0	27	(8.6)	88.6%
Georgia	41.6%	46.7%	(1.2%)	55.9%	0	0	(0.9)	0.3%
Idaho	30.7%	36.1%	(1.2%)	63.4%	0	0	(0.0)	0.0%
Illinois	55.2%	57.9%	(1.2%)	60.7%	21	21	(0.0)	100.0%
Indiana	39.6%	45.5%	(1.3%)	54.3%	0	0	(0.0)	0.0%
Iowa	49.7%	51.9%	(0.9%)	69.7%	0	7	(0.4)	99.7%
Kansas	37.1%	42.5%	(1.2%)	61.0%	0	0	(0.0)	0.0%
Kentucky	40.0%	44.9%	(1.2%)	58.6%	0	0	(0.0)	0.0%
Louisiana	42.7%	47.3%	(1.0%)	60.9%	0	0	(0.9)	1.0%
Maryland	56.6%	59.7%	(1.1%)	61.5%	10	10	(0.0)	100.0%
Michigan	51.7%	54.4%	(1.0%)	66.7%	17	17	(0.0)	100.0%
Minnesota	51.8%	52.8%	(0.7%)	77.9%	10	10	(0.0)	100.0%
Mississippi	40.1%	45.9%	(1.2%)	55.5%	0	0	(0.0)	0.0%
Missouri	46.4%	49.5%	(0.9%)	65.3%	0	0	(4.7)	24.2%
Montana	39.7%	44.2%	(1.0%)	64.4%	0	0	(0.0)	0.0%
Nevada	48.7%	53.2%	(1.3%)	54.7%	0	5	(0.3)	99.7%
New Jersey	53.4%	56.5%	(1.2%)	61.6%	15	15	(0.0)	100.0%
New Mexico	49.6%	53.2%	(1.2%)	58.4%	0	5	(0.3)	99.7%
New York	59.3%	62.5%	(1.3%)	56.9%	31	31	(0.0)	100.0%
North Carolina	43.8%	48.3%	(1.2%)	57.1%	0	0	(3.8)	6.7%
North Dakota	36.1%	41.0%	(1.0%)	65.6%	0	0	(0.0)	0.0%
Ohio	48.9%	51.5%	(0.9%)	67.6%	0	20	(3.4)	97.0%
Oklahoma	34.4%	40.6%	(1.3%)	58.4%	0	0	(0.0)	0.0%
Oregon	52.1%	53.6%	(0.8%)	71.3%	7	7	(0.0)	100.0%
Pennsylvania	51.3%	54.0%	(1.1%)	62.4%	21	21	(0.0)	100.0%
South Carolina	41.4%	47.2%	(1.3%)	52.6%	0	0	(0.8)	1.0%
South Dakota	39.2%	43.1%	(0.8%)	68.9%	0	0	(0.0)	0.0%
Tennessee	42.8%	47.7%	(1.2%)	56.2%	0	0	(1.8)	2.7%
Texas	38.5%	45.1%	(1.5%)	52.1%	0	0	(0.0)	0.0%
Utah	26.7%	32.6%	(1.5%)	59.1%	0	0	(0.0)	0.0%
Virginia	45.9%	50.1%	(1.1%)	60.1%	0	13	(6.5)	48.5%
Washington	53.6%	55.6%	(1.0%)	66.4%	11	11	(0.0)	100.0%
West Virginia	43.5%	49.2%	(1.4%)	53.8%	0	0	(2.2)	25.3%
Wisconsin	50.2%	51.7%	(0.7%)	75.2%	10	10	(0.0)	100.0%
Wyoming	29.7%	35.2%	(1.1%)	65.3%	0	0	(0.0)	0.0%

Table 4: State-level Simulation Results under Full Turnout. The shaded rows correspond to the states in which the winning party in the counterfactual differs from that of the actual data. The total number of electors is 538 (the number of electors for the states included in our data is 486), and 270 electors are needed to win the election.

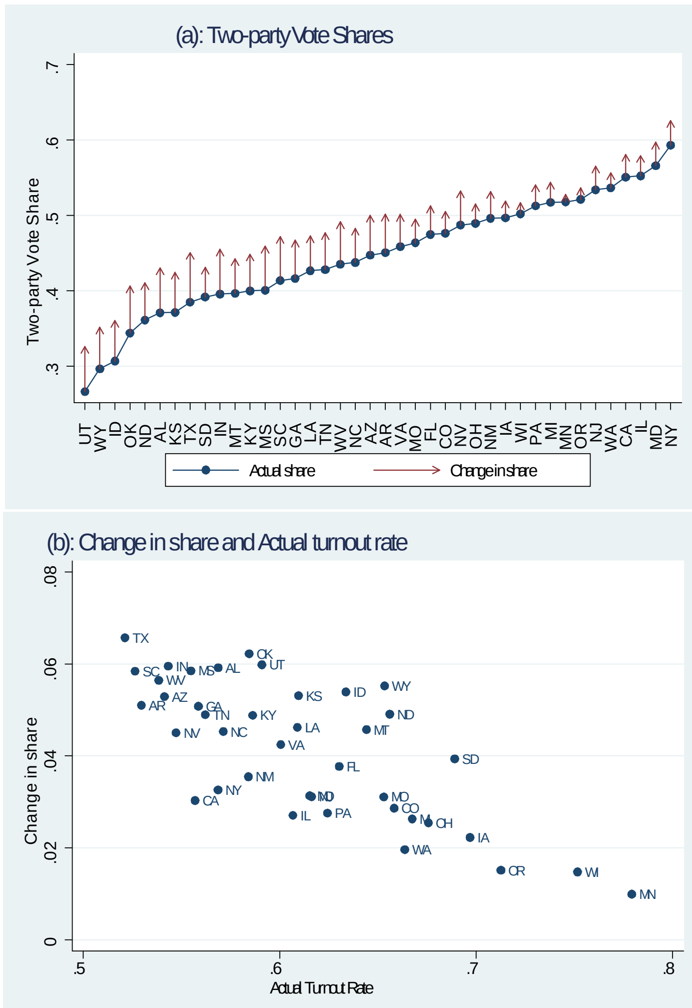

Figure 9: Comparison of Actual and Counterfactual Results by State. Panel (a) plots the two-party vote share of the Democrats in the data and in the counterfactual. Each arrow represents the change in the vote share from the actual to the counterfactual. Panel (b) plots the change in the vote share against the level of actual turnout rate..

The standard errors in our parameter estimates translate to an 87.7% probability that the number of electors for the Democrats exceeds 270. Note that this probability is a lower bound on the probability that the Democrats win the election because it assumes that all states excluded from our sample vote for the Republicans. If we instead assume that all states excluded from our sample vote in the same way as they did in the actual election, the probability that the Democrats win is 96.0%.

Table 4 and Figure 9 present the state-level breakdown of the counterfactual results for the 40 states in our sample. We find that the two-party vote share of the Democrats increases in the counterfactual in all states, and that the results are overturned in 9 states (shaded in the table) in the counterfactual. The table also shows that there is considerable heterogeneity in the magnitude of the change across states. For example, in Texas, we find that the change in the two-party vote share for the Democrats is more than 5 percentage points (from 38.5% to 45.1%), while, in Minnesota, the change is only 1.0 percentage point.

To examine the heterogeneity across states, we plot the state-level change in two-party vote share against the two-party vote share in the actual data (Panel (a)) and the turnout (Panel (b)) in Figure 9. Panel (a) shows that the Democrats would increase its vote share more in the states where they currently have low vote share. Panel (b) of the figure shows that the change in vote share in state with high voter turnout tend to be small, while state with low voter turnout tend to have large difference in two-party vote share between the actual data and the counterfactual. Note that the correlation between actual two-party vote share and turnout ratio is low at 0.22, while the correlation for Panel (a) is -0.79 and for Panel (b) -0.74. KOKOKARA: WHY IS THIS THE CASE??

In our second counterfactual, we examine how the outcomes change as we vary the level of turnout. We compute the outcomes corresponding to different levels of turnout (from 10% to 100% in increments of 10%) by adding (or subtracting) a constant to the cost of voting for all voters. Table ZZZZZ reports the results. We find a monotonic increase for both the vote share and the number of electors in favor of Democrats as the turnout increases, except for the change in the number of electors between 10% and 20% turnout. We also find that the election results would be overturned at a voter turnout level between 70% and 80%, at least for the subset of the states in our sample. Unlike the counterfactual of the compulsory voting discussed above, however, this

Turnout		Vote Share Democrats	Vote Share Republicans	Two-Party Democrat Share	Electors Democrats	Electors Republicans
	100.0%	51.9%	48.1%	51.9%	310	176
	90.0%	46.1%	43.9%	51.2%	281	205
	80.0%	40.2%	39.8%	50.3%	245	241
	70.0%	34.5%	35.5%	49.3%	220	266
(Actual)	60.1%	29.0%	31.1%	48.2%	208	278
	60.0%	28.9%	31.1%	48.2%	208	278
	50.0%	23.3%	26.7%	46.6%	198	288
	40.0%	17.9%	22.1%	44.7%	143	343
	30.0%	12.7%	17.3%	42.2%	62	424
	20.0%	7.7%	12.3%	38.5%	31	455
	10.0%	3.2%	6.8%	32.4%	0	486

Table 5: Election Outcomes at Various Levels of Turnout. We change the constant term in the voter's cost function to compute the election outcome under various levels of turnout.

counterfactual exercise (other than the 100% turnout case) has a limitation that we cannot account for possible endogenous changes in the perception of efficacy.

References

Akerlof, George A and William T Dickens, "The Economic Consequences of Cognitive Dissonance," *American Economic Review*, 1982, pp. 307–319.

Atkeson, Lonna Rae, ""Sure, I Voted for the Winner!" Overreport of the Primary Vote for the Party Nominee in the National Election Studies," *Political Behavior*, 1999, *21* (3), 197–215.

Barkume, Anthony J, "Identification of Preference for Election Outcomes from Aggregate Voting Data," *Public Choice*, 1976, *27* (1), 41–58.

Bendor, Jonathan, Daniel Diermeier, and Michael Ting, "A Behavioral Model of Turnout," *American Political Science Review*, 2003, *97* (02), 261–280.

Berry, Steven T, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, *25* (2), 242–262.

Brunell, Thomas L and John DiNardo, "A Propensity Score Reweighting Approach to Estimating the Partisan Effects of Full Turnout in American Presidential Elections," *Political Analysis*, 2004, *12* (1), 28–45.

Brunnermeier, Markus K and Jonathan A Parker, "Optimal Expectations," *American Economic Review*, 2005, *95*, 4.

Campbell, Colin M, "Large Electorates and Decisive Minorities," *Journal of Political Economy*, 1999, *107* (6), 1199–1217.

Cascio, Elizabeth U and Ebonya Washington, "Valuing the Vote: The Redistribution of Voting Rights and State Funds Following the Voting Rights Act of 1965," *Quarterly Journal of Economics*, 2013.

Casella, Alessandra, "Storable Votes," *Games and Economic Behavior*, 2005, *51* (2), 391–419.

Citrin, Jack, Eric Schickler, and John Sides, "What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections," *American Journal of Political Science*, 2003, *47* (1), 75–90.

Coate, Stephen and Michael Conlin, "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 2004, pp. 1476–1504.

_ , _ , and Andrea Moro, "The Performance of Pivotal-Voter Models in Small-Scale Elections: Evidence from Texas Liquor Referenda," *Journal of Public Economics*, 2008, *92* (3), 582–596.

Degan, Arianna and Antonio Merlo, "A Structural Model of Turnout and Voting in Multiple Elections," *Journal of the European Economic Association*, 2011, *9* (2), 209–245.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao, "Voting to Tell Others," Technical Report 2015.

DeNardo, James, "Turnout and the Vote: The Joke's on the Democrats," *American Political Science Review*, 1980, *74* (02), 406–420.

Dhillon, Amrita and Susana Peralta, "Economic Theories of Voter Turnout," *Economic Journal*, 2002, *112* (480), F332–F352.

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

Duffy, John and Margit Tavits, "Beliefs and Voting Decisions: A Test of the Pivotal Voter Model," *American Journal of Political Science*, 2008, *52* (3), 603–618.

Esponda, Ignacio and Demian Pouzo, "Conditional Retrospective Voting in Large Elections," *American Economic Journal: Microeconomics, Forthcoming*, 2016.

Feddersen, Timothy and Alvaro Sandroni, "A Theory of Participation in Elections," *American Economic Review*, 2006, pp. 1271–1282.

Feddersen, Timothy J, "Rational Choice Theory and the Paradox of Not Voting," *Journal of Economic Perspectives*, 2004, pp. 99–112.

Fujiwara, Thomas, "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil," *Econometrica*, 2015, *83* (2), 423–464.

Gordon, Brett R and Wesley R Hartmann, "Advertising Effects in Presidential Elections," *Marketing Science*, 2013, *32* (1), 19–35.

Haile, Philip A and Elie Tamer, "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, 2003, *111* (1), 1–51.

Hansford, Thomas G and Brad T Gomez, "Estimating the Electoral Effects of Voter Turnout," *American Political Science Review*, 2010, *104* (02), 268–288.

Harsanyi, John C, "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior," *Theory and Decision*, 1980, *12* (2), 115–133.

Hotz, V Joseph and Robert A Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 1993, *60* (3), 497–529.

Husted, Thomas A and Lawrence W Kenny, "The Effect of the Expansion of the Voting Franchise on the Size of Government," *Journal of Political Economy*, 1997, pp. 54–82.

Jäckel, Peter, "A Note on Multivariate Gauss-Hermite Quadrature," Technical Report 2005. Sourced from: www.pjaeckel.webspace. virginmedia.com/ANoteOnMultivariateGaussHermiteQuadrature.pdf.

Kanazawa, Satoshi, “A Possible Solution to the Paradox of Voter Turnout,” *Journal of Politics*, 1998, *60* (04), 974–995.

Karp, Jeffrey A and Susan A Banducci, “Political Efficacy and Participation in Twenty-Seven Democracies: How Electoral Systems Shape Political Behaviour,” *British Journal of Political Science*, 2008, *38* (02), 311–334.

Kroth, Verena, Valentino Larcinese, and Joachim Wehner, “A Better Life For All? Democratization and Electrification in Post-Apartheid South Africa,” Technical Report 2013.

Lalley, P Steven and E Glen Weyl, “Quadratic Voting,” Technical Report 2015.

Ledyard, John O, “The Pure Theory of Large Two-Candidate Elections,” *Public Choice*, 1984, *44* (1), 7–41.

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

Levine, David K and Thomas R Palfrey, “The Paradox of Voter Participation? A Laboratory Study,” *American Political Science Review*, 2007, *101* (01), 143–158.

Lijphart, Arend, “Unequal Participation: Democracy’s Unresolved Dilemma Presidential Address,” *American Political Science Review*, 1997, *91* (01), 1–14.

Martinez, Michael D and Jeff Gill, “The Effects of Turnout on Partisan Outcomes in US Presidential Elections 1960–2000,” *Journal of Politics*, 2005, *67* (4), 1248–1274.

McDonald, Michael P., “1980-2014 November General Election,” 2016. accessed May 26th, 2016.

McDonald, Michael P and Samuel L Popkin, “The Myth of the Vanishing Voter,” *American Political Science Review*, 2001, *95* (4), 963–974.

Merlo, Antonio, “Whither Political Economy? Theories, Facts and Issues,” in Whitney Newey Richard Blundell and Torsten Persson, eds., *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. I, Cambridge: Cambridge University Press, 2006, pp. 381–421.

Miller, Grant, "Women's Suffrage, Political Responsiveness, and Child Survival in American History," *The Quarterly Journal of Economics*, 2008, *123* (3), 1287.

Minozzi, William, "Endogenous Beliefs in Models of Politics," *American Journal of Political Science*, 2013, *57* (3), 566–581.

Myatt, David, "A Rational Choice Theory of Voter Turnout," Technical Report 2012.

Palfrey, Thomas R and Howard Rosenthal, "A Strategic Calculus of Voting," *Public Choice*, 1983, *41* (1), 7–53.

_ **and** _ , "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 1985, *79* (01), 62–78.

Riker, William H and Peter C Ordeshook, "A Theory of the Calculus of Voting," *American Political Science Review*, 1968, *62* (01), 25–42.

Shachar, Ron and Barry Nalebuff, "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 1999, pp. 525–547.

Smith, Mark A, "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout," *American Journal of Political Science*, 2001, pp. 700–706.

Strömberg, David, "How the Electoral College Influences Campaigns and Policy: the Probability of Being Florida," *American Economic Review*, 2008, *98* (3), 769–807.

Tucker, Harvey J and James DeNardo, "Does Heavy Turnout Help Democrats in Presidential Elections?," *American Political Science Review*, 1986, *80* (04), 1291–1304.

8 Appendix

8.1 Appendix A: Data Construction

In this Appendix, we explain how we construct the joint distribution of demographic characteristics and citizenship status at the county level. We first use the 5% Public Use

Microdata Sample of the 2000 U.S. Census (hereafter PUMS), which is an individual-level dataset, to estimate the covariance matrix between the demographic variables and citizenship information within each public use microdata area (PUMA). In particular, we estimate the joint distribution of the discrete demographic characteristics (Race, Hispanic, Citizenship) by counting the frequency of occurrence, and then estimate a covariance matrix for the continuous demographic variables (Age, Income, years of schooling) for each bin. Because the PUMA and counties do not necessarily coincide, we estimate covariance matrices for each PUMA and then use the correspondence chart provided in the PUMS website to obtain estimates at the county level.

In the second step, we construct the joint distribution of demographic characteristics by combining the covariance matrix estimated in the first step and the marginal distributions of each of the demographic variables at the county level obtained from Census Summary File 1 through File 3. We discretize continuous variables into coarse bins. We discretize age into 3 bins: (1)18-34 years old, (2) 35-59 years old, and (3) above 60 years old, income into 6 bins: (1) \$0-\$25,000, (2) \$25,000-\$50,000, (3) \$50,000-\$75,000, (4) \$75,000-\$100,000, (5) \$100,000-\$150,000, (6) above \$150,000, and years of schooling into 5 bins: (1) Less than 9th grade, (2) 9th-12th grade with no diploma, (3) high school graduate, (4) some college with no degree or associate degree, (5) bachelor degree or higher. Thus, there are 540 bins in total. The joint distribution of demographic characteristics that we create gives us a probability mass over each of the 540 bins for each county.

Finally, we augment the census data with religion data obtained from Religious Congregations and Membership Study 2000. This data has information on the share of the population with adherence to either “Evangelical Denominations” or “Church of Jesus Christ of Latter-day Saints” at the county level. Because the Census does not collect information on religion, we do not know the correlation between the religion variable and the demographic characteristics in the Census. Thus, we assume independence of the religion variable and other demographic variables. As a result, there are 1,080 bins in our demographics distribution.

8.2 Appendix B: Identification of $c(\cdot)$, $p(\cdot)$, and F_η in the general case

In this Appendix, we show that $c(\cdot)$, $p(\cdot)$, and F_η are identified even when the max operator in equation (11) binds with positive probability. Note that our argument in the main text considered only the case in which the max operator never binds. Recall that

$$\underbrace{\frac{\Phi^{-1}\left(1-v_{D,m}\right)-\Phi^{-1}\left(v_{R,m}\right)}{2}}_{\equiv Y_m}=\max\left\{0,\frac{c(\bar{\mathbf{x}}_m)}{p(\bar{\mathbf{x}}_m)}+\frac{\eta_m}{p(\bar{\mathbf{x}}_m)}\right\},\eta_m\perp\bar{\mathbf{x}}_m. \tag{16}$$

In this appendix, we work with the normalization that the value of $p(\cdot)$ at some $\bar{\mathbf{x}}_m = \mathbf{x}_0$ as $p(\mathbf{x}_0)=1$. This amounts to a particular normalization of variance of η. Note that the distribution of Y_m (the left hand side of equation (16)) conditional on $\bar{\mathbf{x}}_m=\mathbf{x}_0$ is a truncated distribution with mass at zero. Figure 10 illustrates this when the mass at zero is less than 50% and F_η is symmetric and single peaked at zero.

First, we present our identification discussion for the case that F_η is symmetric and single peaked at zero. As Figure 10 illustrates, the median of Y_m conditional on $\bar{\mathbf{x}}_m=\mathbf{x}_0$ directly identifies $c(\mathbf{x}_0)$ under these assumptions. Also, the density of η, f_η, is identified above the point of truncation. Formally, $f_\eta(F_\eta^{-1}(t))$ is identified for any $t>t(\mathbf{x}_0)$, where

$$t(\mathbf{x}_0)=\Pr\left(Y_m=0|\,\mathbf{x}_0\right).$$

Hence, $f_\eta(0)$ is identified from the height of the density of Y_m at the median.

Now consider $\mathbf{x}_1\neq\mathbf{x}_0$. Assume again that $t(\mathbf{x}_1)<0.5$. Then, $c(\mathbf{x}_1)/p(\mathbf{x}_1)$ is identified from the conditional median of Y_m and $p(\mathbf{x}_1)f_\eta(0)$ is identified by the height of the conditional density of Y_m at the median. Given that $f_\eta(0)$ is identified, $c(\mathbf{x}_1)$ and $p(\mathbf{x}_1)$ are both identified. Moreover, F_η is identified over its full support if there exists sufficient variation in $\mathbf{x}$, i.e., $\inf_{\mathbf{x}} t(\mathbf{x})=0$

We now consider the case in which F_η is not restricted to be symmetric and single peaked and $t(\mathbf{x}_0)$ may be less than 0.5. The distribution of Y_m is identified above $t(\mathbf{x}_0)$ as before. Now consider $\mathbf{x}_1\neq\mathbf{x}_0$. Similar as before, we identify $p(\mathbf{x}_1)f_\eta(F_\eta^{-1}(\tau))$ for τ above $t(\mathbf{x}_1)$.[31] If we let τ be any number larger than $\max\{t(\mathbf{x}_0),t(\mathbf{x}_1)\}$, both

[31] Note that we identify $f_{\eta/p(\mathbf{x}_1)}\left(F^{-1}_{\eta/p(\mathbf{x}_1)}(t)\right)$, where $f_{\eta/p(\mathbf{x}_1)}(\cdot)$ and $F^{-1}_{\eta/p(\mathbf{x}_1)}(\cdot)$ are the density

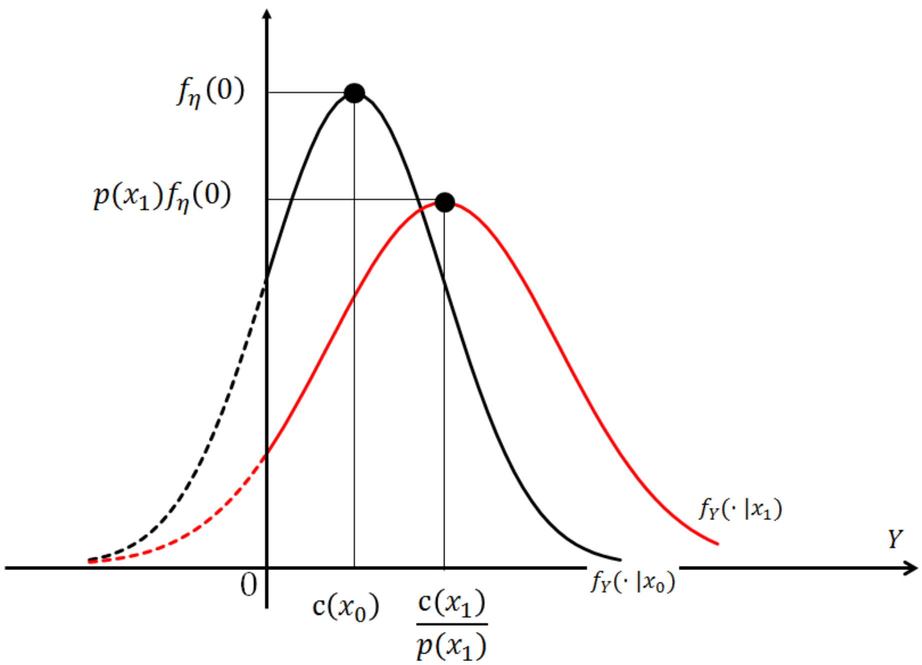

Figure 10: The distribution of Y_m conditional on $\mathbf{x} = \mathbf{x}_0$ and $\mathbf{x} = \mathbf{x}_1$ when the distribution of η is symmetric and single peaked, and $t(\mathbf{x}_0)$, $t(\mathbf{x}_1) < 0.5$, where $t(\mathbf{x})$ is the probability that Y_m is equal to zero conditioanl on $\mathbf{x}$. Note that the distribution of Y_m is truncated at zero. The conditional median of Y_m identifies $c(\mathbf{x}_0)$ and $c(\mathbf{x}_1)/p(\mathbf{x}_1)$, and the height of the density at the conditional median identifies $f_\eta(0)$ and $p(\mathbf{x}_1)f_\eta(0)$.

$f_\eta(F_\eta^{-1}(\tau))$ and $p(\mathbf{x}_1)f_\eta(F_\eta^{-1}(\tau))$ are identified. Hence $p(\mathbf{x}_1)$ is identified. Similarly, $p(\cdot)$ is identified for all $\mathbf{x}$.

We now consider identification of $c(\cdot)$. We present two alternative assumptions on F_η and show that $c(\cdot)$ can be identified under either assumption. First, assume that the median of η is zero, $Med(\eta) = 0$, and that there exists $\mathbf{x} = \mathbf{x}_2$ such that $t(\mathbf{x}_2) < 1/2$. The latter assumption means that more than half of the counties have turnout less than 100% when $\mathbf{x} = \mathbf{x}_2$. Then, the median of Y_m conditional on $\mathbf{x}_2$ identifies $c(\mathbf{x}_2)/p(\mathbf{x}_2)$. Now consider any $\mathbf{x}_1 \neq \mathbf{x}_2$ and let τ be any number larger than $\max\{t(\mathbf{x}_2), t(\mathbf{x}_1)\}$. Let z_1 and z_2 be the τ quantile of Y_m conditional on $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. z_1 and z_2 are clearly identified. Then,

$$\begin{aligned}
&p(\mathbf{x}_1)\left[F^{-1}_{\eta/p(\mathbf{x}_1)}(\tau) - F^{-1}_{\eta/p(\mathbf{x}_1)}(1/2)\right] = p(\mathbf{x}_2)\left[F^{-1}_{\eta/p(\mathbf{x}_2)}(\tau) - F^{-1}_{\eta/p(\mathbf{x}_2)}(1/2)\right] \\
&\quad\Leftrightarrow p(\mathbf{x}_1)\left[z_1 - \frac{c(\mathbf{x}_1)}{p(\mathbf{x}_1)}\right] = p(\mathbf{x}_2)\left[z_2 - \frac{c(\mathbf{x}_2)}{p(\mathbf{x}_2)}\right] \\
&\quad\Leftrightarrow \frac{c(\mathbf{x}_1)}{p(\mathbf{x}_1)} = z_1 - \frac{p(\mathbf{x}_1)}{p(\mathbf{x}_2)}\left(z_2 - \frac{c(\mathbf{x}_2)}{p(\mathbf{x}_2)}\right).
\end{aligned} \tag{17}$$

Given that all of the terms on the right hand side of (17) are identified, $c(\mathbf{x}_1)/p(\mathbf{x}_1)$ is identified.

Alternatively, assume that $\mathbf{E}(\eta) = 0$ and $\inf_{\mathbf{x}} t(\mathbf{x}) = 0$. We now show that $c(\cdot)$ is identified under these alternative assumptions. Intuitively, this latter assumption means that there exist values of $\mathbf{x}$ for which the max operator is never binding. In this case, we can fully recover the distribution of $F_\eta(\cdot)$. Then we can identify the distribution of $c(\mathbf{x})/p(\mathbf{x}) + \eta_m/p(\mathbf{x})$ for any $\mathbf{x}$. Hence we identify $c(\cdot)/p(\cdot)$.

of $\eta/p(\mathbf{x}_1)$ and the inverse distribution of $\eta/p(\mathbf{x}_1)$, respectively. Note that $f_{\eta/p(\mathbf{x}_1)}\left(F^{-1}_{\eta/p(\mathbf{x}_1)}(t)\right) = p(\mathbf{x}_1)f_\eta(F_\eta^{-1}(t))$.

8.3 Online Appendix: Additional Tables and Discussion on Fit

	Estimate	SE		Estimate	SE
Alabama	0 (Normalized)		Nevada	0.009	(0.059)
Arizona	0.132	(0.050)	New Jersey	0.247	(0.041)
Arkansas	0.363	(0.034)	New Mexico	0.055	(0.069)
California	0.236	(0.060)	New York	0.239	(0.036)
Colorado	0.057	(0.055)	North Carolina	0.088	(0.033)
Florida	0.101	(0.039)	North Dakota	0.006	(0.040)
Georgia	-0.107	(0.028)	Ohio	0.166	(0.035)
Idaho	-0.129	(0.044)	Oklahoma	0.057	(0.041)
Illinois	0.243	(0.030)	Oregon	0.130	(0.054)
Indiana	0.099	(0.031)	Pennsylvania	0.190	(0.037)
Iowa	0.300	(0.032)	South Carolina	-0.030	(0.031)
Kansas	-0.194	(0.038)	South Dakota	0.000	(0.039)
Kentucky	0.251	(0.032)	Tennessee	0.345	(0.034)
Louisiana	-0.163	(0.030)	Texas	-0.127	(0.034)
Maryland	0.107	(0.053)	Utah	-0.079	(0.062)
Michigan	0.246	(0.031)	Virginia	0.114	(0.035)
Minnesota	0.297	(0.032)	Washington	0.223	(0.055)
Mississippi	-0.128	(0.033)	West Virginia	0.216	(0.037)
Missouri	0.217	(0.029)	Wisconsin	0.335	(0.033)
Montana	-0.098	(0.044)	Wyoming	-0.230	(0.057)

Table 6: Estimates of State Preference Fixed Effects Relative to $\lambda_{Alabama}$. Standard errors are reported in parentheses. Higher values imply stronger preference for Democrats.

	Estimate	SE		Estimate	SE
Alabama	1 (Normalized)		Nevada	0.854	(0.075)
Arizona	0.660	(0.109)	New Jersey	0.725	(0.054)
Arkansas	0.739	(0.043)	New Mexico	0.976	(0.083)
California	0.751	(0.053)	New York	0.795	(0.052)
Colorado	1.165	(0.073)	North Carolina	0.787	(0.040)
Florida	0.975	(0.054)	North Dakota	1.062	(0.056)
Georgia	0.842	(0.039)	Ohio	1.323	(0.057)
Idaho	1.274	(0.081)	Oklahoma	0.919	(0.041)
Illinois	1.067	(0.041)	Oregon	1.441	(0.086)
Indiana	0.814	(0.042)	Pennsylvania	0.838	(0.046)
Iowa	1.319	(0.055)	South Carolina	0.751	(0.045)
Kansas	0.877	(0.043)	South Dakota	1.680	(0.105)
Kentucky	1.052	(0.037)	Tennessee	0.867	(0.041)
Louisiana	1.316	(0.079)	Texas	0.769	(0.041)
Maryland	0.767	(0.050)	Utah	1.229	(0.087)
Michigan	1.180	(0.048)	Virginia	0.754	(0.043)
Minnesota	1.765	(0.103)	Washington	1.080	(0.051)
Mississippi	1.193	(0.067)	West Virginia	0.814	(0.043)
Missouri	1.159	(0.040)	Wisconsin	1.762	(0.108)
Montana	1.066	(0.054)	Wyoming	1.151	(0.060)

Table 7: Estimates of State-level Fixed Effects of Voting Efficacy. Standard errors are reported in parentheses. Alabama is set to 1 for normalization.