Why People Vote: Comparing Models of Voter Turnout

Luke Miller and Maxime Cugnon de Sévricourt

September 17, 2024

Draft: Preliminary, please do not quote or cite

Abstract

This paper compares the performance of several leading theoretical models of voter turnout in fitting election results for U.S. House elections and U.S. state special elections. I show a model of turnout where individuals are divided into two sides, vote for the candidate that maximizes their utility while taking into account how their vote impacts the outcome of the election, and face heterogenous voting costs, outperforms other models. I also consider models whereby leaders make strategic decisions to mobilize or persuade voters to turnout, and where individuals vote according to the social norms of their group.

1 Introduction

Voting is a fundamental aspect of democratic societies and a key component to many political economy models. An estimated 4 billion people are eligible to vote in 2024, with 1 billion people expected to actually cast a ballot in an election.* Despite being such an important aspect of society, and the primary means by which individuals participate in the political process, social scientists still do not have a widely accepted model of voter turnout. Instead, there exists a number of models, each which purports to explain why individuals vote, and how they choose who to vote for.

This paper focuses on determing which of the existing models of turnout best fits the data, and whether these models perform well in absolute terms. To understand why there exists so many models of turnout, consider a basic model where peole vote if the following inequality holds:

$$piv\Delta u - c + d > 0 \tag{1}$$

Here piv is the probability an individual's vote is pivotal to the outcome of the election, Δu is the utility differential between their candidate winning and

*https://www.cnn.com/2024/07/08/world/global-elections-2024-maps-charts-dg/index.html losing, c is the cost of voting, and d is any intrinsic benefits the voter receives from participating in the election.

Early work focused on shutting down the effects from d and seeing how well the model could perform. However, in large elections, unless the expected vote shares between the two parties are incredibly close, the probability of one's vote being pivotal to the outcome of the election is astronomically small. For instance, since 2000, the average U.S. House of Representatives election has had 500,000 eligible voters and an average turnout rate of 50%. If the election has two candidates, D and R, and there is a 24.75% chance an individual votes for the D candidate, and a 25.25% chance they vote for the R candidate, leading to an expected victory margin of .5%, then the probability of being pivotal is approximately $1e{-}9$. If this spread increases to a 1% expected victory margin, the probability of being pivotal falls to approximately $1e{-}25$, and with an expected 3% victory margin, the probability of being pivotal is approximately $1e{-}199$. Given the average victory margin in U.S. House elections is 15%, one must assume large utility differentials between the two candidates to get large turnout rates.

To calculate just how large these utility differentials must be, suppose the cost of voting is equal to \$1. The minimum utility differential between the two candidates is:

$$\Delta u_{min} = \frac{c}{piv}$$

When the probability of being pivotal is $1e{-}9$, Δu_{min} is approximately \$350,000,000. When the probability of being pivotal is $1e{-}25$, this rises to approximately \$1e24, and by the time the probability of being pivotal is $1e{-}199$ (which is the case when the expected victory margin is only 3%), the minimum utility differential is approximately \$1e198.

Given this the implausibility of these numbers, authors have either modified the *piv* term to be more flexible, or focused on other aspects of equation 1 to explain turnout. This has manifested in to largely three branches of literature: (1.) Calculus of Voting models, (2.) Follow the Leader models, and (3.) Group Based models.

Calculus of Voting models update the *piv* term to be more flexible. For instance Castanheira (2003) allows for individuals to care not just about the probability of changing the outcome of the election, but the overall victory margin. With larger victory margins, politicians may find it easier to implement their preferred policies.

Follow the Leader models focus on how leaders can influence the d and c portions of equation 1. For instance, leaders can create a desire to vote through advertising or by holding campaign events. Alternatively, leaders can lower the cost of voting by holding voter registration drives or by providing transportation to the polls. In such models, individuals are no longer deciding to vote based on the decisions of other voters. Instead, leaders strategically decide how to mobilize or influence voters to turnout.

In a similar vein, Group Based models focus on how individuals are influenced by the social norms of their group. Again, individual voters are not deciding to vote based on the decisions of other voters. Now, they do so based on the social norms of their group. These social norms may, for instance, raise the cost of abstaining by creating a social stigma around not voting.

Despite the large number of models built to explain turnout, there has been little work directly comparing how well these different models fit the data. This paper aims to fill that gap. To do so, I estimate five leading theoretical models of turnout on two separate datasets. Two are based around models found in the Calculus of Voting literature, two are based around models found in the Follow the Leader literature, and one is based around a model found in the Group Voting literature.

I estimate these models on two separate datasets: U.S. House elections from 2000-2020 and a novel data set of U.S. state special elections from 2011-2023. The U.S. House elections provide a large number of comparable elections. However, they often coincide with other elections such as the U.S. Presidential Election or Senate races. Therefore, estimations may overestimate the impact of the U.S. House elections on turnout. To address this issue, I also estimate the models on U.S. state special elections. These occur when a state legislature seat becomes vacant before the end of the legislature session and a special election is called to fill the seat for the remainder of the term. These special elections are unique because they occur when no other elections are taking place, ensuring that individuals are choosing to turnout based solely on the special election.

To compare the performance of these models in fitting the data, I use the Vuong test, which is designed to compare the performance of two non-nested models. This enables me to compare the performance of each model to every other model. In addition to this, I test the absolute performance of each model by checking how well they fit the data.

This paper aims to make several contributions. First and foremost, I compare the performance of several models of turnout in fitting the exact same set of election results and check their overall performance in matching the data. This provides a clear picture of which models are performing well and which are not. In the process of doing so, I also estimate several models of turnout that have not been previously brought to the data. For instance, I estimate a version of the leader persuasion model developed by Strömberg (2008). However, I add the ability of voters to abstain. Additionally I estimate a new model of voter mobilization based around Bouton et al. (2023) where candidates aim to lower the voting costs of their supporters.

I find a version of the calculus of voting model where individuals vote taking into the account their *perceived* probability of being pivotal, a term taken from Kawai, Toyama, and Watanabe (2021), performs the best on the U.S. House election data. However, contrary to Kawai, Toyama, and Watanabe (2021), I maintain that the perceived probability of being pivotal remains an equilibrium condition in that one's decision to vote depends on other voters' decisions.

For the U.S. state special elections, the leader mobilization model following Bouton et al. (2023) performs the best. However, the Vuong test does not reject the null hypothesis that the calculus of voting model matches the data equally as well.

2 Literature Review

Early work modeling individual turnout decisions comes from Downs (1957), Riker and Ordeshook (1968), and Palfrey and Rosenthal (1983). In these models, two candidates are typically running for office and individuals are a supporter of one of the two candidates. Each individual must decide whether to vote for their preferred candidate or abstain. Their formulation matches equation 1, where individuals vote taking into account the probability their vote is pivotal, the utility differential between the two candidates, costs of voting, and intrinsic benefits from voting. They do not focus on any other agents in the system, such as leaders or groups. These papers are the first to show how small the probability of being pivotal is in large elections unless expected turnout rates are nearly identical.

A number of papers have since expanded upon these models, working to increase the impact of the *piv* term in equation 1 in elections with large numbers of eligible voters. As mentioned in the introduction, Castanheira (2003) allows for individuals to care not just about the probability of changing the outcome of the election, but the overall victory margin. Myatt (2015) develops a model where individuals are uncertain about the overall support for each candidate. In both cases the authors show turnout rates can be higher than predicted by the original models.

Coate, Conlin, and Moro (2008) estimate a version of the Palfrey and Rosenthal (1983) model on data from Texas liquor referenda, assuming any intrinsic benefits from voting are zero and that individuals are able to perfectly calculate the probability of being pivotal. They find it does a very poor job in matching the victory margins of the election, consistently predicting much closer elections than actually occurred. Kawai, Toyama, and Watanabe (2021) estimate a calculus of voting model on U.S. Presidential Election data. However, contrary to the standard models, they do not enforce that the probability of being pivotal is an equilibrium condition. Instead, they estimate it to be a function of a voter's characteristics and state level characteristics.

A second strand of literature focuses on how leaders can influence turnout. One of the earliest papers in this literature is Shachar and Nalebuff (1999). The authors assume voters are non-strategic, meaning, they do not take into account the probability of being pivotal when deciding to vote. Instead, leaders exert effort to lower the voting costs of their supporters. They do this by, for example, helping voters get to the polls or providing information on how to vote. While individuals are not strategic in deciding whether to vote, leaders are strategic in deciding how much effort to exert in each district. This leads to the standard comparative static that more competitive elections see higher levels of turnout. Bouton et al. (2023) develop a similar model to understand how parties optimally draw redistricting lines to maximize the number of seats they win.

Strömberg (2008) additionally develops a follow the leader model. His model takes inspiration from the Probabilistic Voting Model as described by Lindbeck and Weibull (1987). He assumes campaigns can influence the individual's utility from voting for a candidate. Individuals are not strategic in deciding whether to vote, but candidates stategically decide where to campaign, leading to higher turnout in more competitive districts. The author estimates this data on results from U.S. Presidential Elections. Incerti (2018) estimates a similar model on U.S. Congressional Elections. In both cases, the authors find that the model fits the data well. However, note that neither of these models allow for individuals to abstain. Instead they only model support within a district, assuming that turnout is determined exogenously and does not play a role in campaign strategies.

A final strand of literature focuses on how groups can influence either the cost of abstaining or the utility from participation. Feddersen and Sandroni (2006) develop a model where both sides create a social norm that encourages their supporters to vote. However, free riders in both groups choose to ignore the social norm. Similarly, Coate and Conlin (2004) present a model where everyone follows the social norm, but the strength of support for the two sides is unknown before establishing the voting rule. Both models allow for high levels of turnout even in large electorates. The latter model is estimated on a dataset of Texas liquor referenda, where the authors find evidence that their group utility model fits the data well.

3 Models

In this section, I outline each of the models estimated in this paper. I first estimate two versions of the calculus of voting model. Both versions are closely related to the classical model by Palfrey and Rosenthal (1983). However, they differ in what causes heterogeneity in turnout among individuals. In the first model, differences in turnout are due to differences in costs of voting. In the second model, differences in turnout are due to differences in the intensity of preferences for the two candidates. For the former, I first estimate the model using the analytical solution for the probability of being pivotal from Palfrey and Rosenthal (1983). However, given its poor performance I do not consider it further.

I then estimate both models using a contest function that maps expected vote shares for both candidates to a perceived probability of being pivotal, along with a version setting the probability of being pivotal to 1 across all elections. The contest function is estimated in a way to ensure that turnout decisions are consistent with the perceived probability of being pivotal. To the best of our knowledge, this is the first time this model has been estimated on election data this way. Setting the probability of being pivotal to 1 across all elections shuts down any effect from the *piv* term in equation 1, capturing a model more akin to an intensity model where individuals vote based solely on the utility differential between the two candidates.

I next estimate two follow the leader models: one where leaders target the d term in equation 1 and one where leaders target the c term. The first model is based around one developed by Strömberg (2008), which was found to be successful in explaining candidate behavior in both U.S. Presidential and later by Incerti (2018) for U.S. Congressional Elections. However, neither model included the ability for individuals to abstain. We therefore add this feature to the model. The second model comes from Bouton et al. (2023). This model focuses on how candidates can lower the cost of voting for their supporters, leading to higher turnout in more competitive districts. We are the first to estimate this model on election data.

Finally, I estimate a group utility model. This model closely follows Coate and Conlin (2004), where individuals vote based on the social norms of their group. These social norms are chosen to maximize the group's aggregate utility, which is a function of the expected size of the group, the utility members of the group get from winning the election, and the total cost they face from voting. The authors of this model structurally estimated it on a dataset of Texas liquor referenda and found it to be quite successful in matching the data. For this reason, I include their model in my analysis.

3.1 Basic Setup

In each model, there are two candidates running for office: D and R. They compete in a single election indexed by $j \in J$. Whoever receives the most votes wins the election, with ties decided by a fair coin flip.

Individuals are indexed by i. Each individual i in election j faces a cost of voting c_{ij}. Additionally, each individual i in election j receives a utility u_{ij}^s if candidate $s \in \{D, R\}$ wins the election. Let $\Delta u_{ij} \equiv u_{ij}^R - u_{ij}^D$. This is the utility differential the individual receives from candidate R winning the election over candidate D.

Individuals choose to either vote for D, vote for R, or abstain (A). Denote the probability an individual votes for D, votes R, or abstains (A) in election j as $Pr(v_{ij} = v)$, where $v \in \{D, R, A\}$. Relatedly, I will use σ_{vj} to denote the probability a random individual in the population votes for candidate D, candidate R, or abstains in election j. Note that σ_{vj} is the expected fraction of individuals who choose action $v \in \{D, R, A\}$ in election j.

3.2 Calculus of Voting

We estimate two versions of the calculus of voting model. One where preferences are identical within a side, though individuals face heterogenous costs of voting, and a second where the costs of voting are identical across a district but individuals have heterogenous preferences for the two candidates.

In both cases, individuals consider the probability of being pivotal within election j when deciding to vote, denoted by piv_j. In addition, there is assumed to be a finite and known number of eligible voters in each election n_j.

3.2.1 Probability of Being Pivotal

Denote $p_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j)$ as the analytical solution for the probability of casting the pivotal vote in favor of candidate $s \in \{D, R\}$ in election j. When $n_j - 1$ is even, $p_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j)$ is calculated as:

$$p_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j) = \frac{1}{2} \sum_{k=0}^{(n_j-1)/2} \binom{n_j}{k} \sigma_{Dj}^k \sigma_{Rj}^k (1 - \sigma_{Dj} - \sigma_{Rj})^{n_j - k}$$

This function sums across all situations in which k individuals vote for the D candidate and k individuals vote for the R candidate, resulting in a tie. The remaining individual can then cast a vote to break the tie. It then multiplies each situation by the probability of that situation occurring. The $\frac{1}{2}$ term is due to the fact that had the individual not voted, their preferred candidate still would have won the election with 50% probability due to the coin flip. This probability is defined similarly when $n_j - 1$ is odd. When estimating this model, we use an approximation from Myerson (2000):

$$\tilde{p}_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j) = exp(n_j) \left(log(\frac{p^*_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j)}{n_j}) + 1 \right) \tag{2}$$

Where

$$p^*_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j) = \frac{\exp\left(n_j \left(2\sqrt{\sigma_{Dj}\sigma_{Rj}} - \sigma_{Dj} - \sigma_{Rj}\right)\right)}{4\sqrt{\pi n_j \sqrt{\sigma_{Dj}\sigma_{Rj}}}} \left(\frac{\sqrt{\sigma_{Dj}} + \sqrt{\sigma_{Rj}}}{\sqrt{\sigma_{sj}}} \right)$$

So long as there are a large number of voters in the election and the turnout rates are not too low, $p_{s,ana} \approx \tilde{p}_{s,ana}$. As mentioned in the introduction, unless σ_{Dj} and σ_{Rj} are very close to each other, the probability of being pivotal is very small for large elections. Therefore, we also estimate a second specification where we use a more flexible version of the probability of being pivotal:

$$p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2) = \frac{1}{1 + e^{\alpha^1(\alpha^2 \sigma_j - 0.5)}} \tag{3}$$

Where both α^1 and α^2 are estimated from the data, and

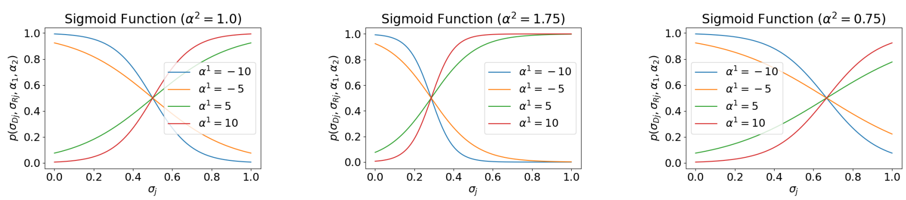
$$\sigma_j \equiv \begin{cases} \frac{\sigma_{Dj}}{\sigma_{Rj}} & \text{if } \sigma_{Dj} < \sigma_{Rj} \\ \frac{\sigma_{Rj}}{\sigma_{Dj}} & \text{if } \sigma_{Dj} \geq \sigma_{Rj} \end{cases}$$

The σ_j term captures the competitiveness of the election. Values of σ_j close to 1 indicate approiximately even number of expected votes for each candidate, while values of σ_j closer to 0 indicate one candidate is expected to win by a large margin.

Figure 1 plots $p(\sigma_{Dj}, \sigma_{Rj}, \alpha^1, \alpha^2)$ for varying values of α^1 and α^2. For positive values of α^1, one's perceived probability of being pivotal increases as the election becomes more competitive. However, the functional form also allows for the possibility that inidividuals are less likely to vote in competitive elections relative to uncompetitive ones. Further, the magnitude of α^1 determines the rate of transition from low to high probability of being pivotal. When α^1 is positive, larger values indicate a more sudden transition, which is more consistent with the analytical solution from Palfrey and Rosenthal (1983).

The second parameter, α_2, determines the point where the probability of being pivotal is equal to 0.5. Larger values of α_2 cause the probability of being pivotal to be equal to 0.5 closer to where $\sigma_j = 0$, and smaller values cause this point to be closer to where $\sigma_j = 1$. The latter case is more consistent with the analytical solution from Palfrey and Rosenthal (1983) where only in very close elections is the probability of being pivotal above zero for large elections. Figure 1 plots $p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2)$ for varying values of α^1 and α^2.

Figure 1: $p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2)$ for varying values of α^1 and α^2



3.2.2 Calculus of Voting with Hetergenous Costs

Cost of voting is heterogenous across individuals:

$$c_{ij} = c_j - \epsilon_{ijc}$$

where c_j is the mean cost of voting in election j and $\epsilon_{ijc} \sim H$ is a mean zero cost shock. Whether Δu_{ij} is positive or negative depends on a random variable $x_{ij} \sim F[-1, \mu_j]$, where F is the uniform distribution. If $x_{ij} > 0$, then $\Delta u_{ij} > 0$. We label this person a supporter of R. Else, $\Delta u_{ij} < 0$, and we label this person a supporter of D. We assume the magnitude of the utility differential is identical across both sides and the same for all individuals within a given side. Therefore, for simplicity we can define $u_j \equiv |\Delta u_{ij}|$.

Finally, there is an idiosyncratic shock to abstaining, $\epsilon_{ijA} \sim H$, which is independent of the cost shock ϵ_{ijc}. Individuals vote if the following inequality holds:

$$piv_j u_j - c_j + \epsilon_{ijc} > \epsilon_{ijA} \tag{4}$$

where we use either $\tilde{p}_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j)$ (equation 2), $p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2)$ (equation 3), or 1 for piv_j.

We assume the ϵ_{ijc} and ϵ_{ijA} shocks are independent and identically distributed across individuals and elections and are drawn from a type 1 extreme value distribution.[†]

The probability an individual votes for candidate D is the probability $x_{ij} < 0$ and the probability the expected utility from voting for D is greater than the expected utility from abstaining:

$$\begin{aligned}\sigma_{Dj} &= Pr(x_{ij} \leq 0) Pr(piv_j u_j - c_j + \epsilon_{ijc} > \epsilon_{ijA}) \\ &= F(0) \frac{e^{(piv_j u_j - c_j)}}{1 + e^{(piv_j u_j - c_j)}} \end{aligned} \tag{5}$$

Likewise, the probability an individual votes for candidate R (σ_{Rj}) is:

$$\begin{aligned}\sigma_{Rj} &= Pr(x_{ij} > 0) Pr(piv_j u_j - c_j + \epsilon_{ijc} > \epsilon_{ijA}) \\ &= (1 - F(0)) \frac{e^{(piv_j u_j - c_j)}}{1 + e^{(piv_j u_j - c_j)}} \end{aligned} \tag{6}$$

Notice that whether we use $\tilde{p}_{s,ana}(\sigma_{Dj}, \sigma_{Rj}, n_j)$ or $p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2)$, the expected turnout rates are themselves functions of the expected turnout rates. When estimating either of these versions, we estimate them ensuring this equilibrium condition is satisfied. We also estimate the model where the probability of being pivotal is set to 1 across all elections.

3.2.3 Calculus of Voting with Hetergenous Preferences

Cost of voting is now constant across individuals in a given district:

$$c_{ij} = c_j$$

and the cost of abstaining is set to 0. However, the utility differential between the two candidates now varies across individuals:

$$\Delta u_{ij} = u_j + \epsilon_{ij}$$

where u_j is the mean utility differential in election j and $\epsilon_{ij} \sim H$ is a mean zero utility shock. We again use a type 1 extreme value distribution for the ϵ_{ij} shocks to match the other models. As before, if Δu_{ij} is positive, then individual i is considered a supporter of R, meaning they receive more utility from candidate R winning the election compared to D winning the election. Else, they are considered a supporter of D.

An individual votes for R if the following inequality holds:

$$piv_j(u_j + \epsilon_{ij}) - c_j > 0 \tag{7}$$

[†]While this exact form of the cdf is not necessary for our estimation procedure, in other models we estimate, it is. Therefore, we maintain this assumption for consistency.

Voting for D depends on the utility differential from D winning vs. R winning:

$$\begin{aligned} \text{utility if D wins} - \text{utility if R wins} = \\ -(\text{utility if R wins} - \text{utility if D wins}) = \\ -\Delta u_{ij} = -u_j - \epsilon_{ij} \end{aligned}$$

Therefore, an individual votes for D if the following inequality holds:

$$piv_j(-u_j - \epsilon_{ij}) - c_j > 0 \tag{8}$$

Else, they abstain.

We can use equations 7 and 8 to calculate the probability an individual in district j votes for D and R, respectively. Notice an individual votes for D if (A.) they are a supporter of D and (B.) the expected utility from voting for D is greater than the utility from abstaining. We can write this as:

$$\begin{aligned} Pr(v_{ij} = D) &= Pr(u_j + \epsilon_{ij} < 0 \ \& \ piv_j(-u_j - \epsilon_{ij}) - c_j > 0) \\ &= Pr(\epsilon_{ij} < -u_j \ \& \ \epsilon_{ij} < -u_j - \frac{c_j}{piv_j}) \end{aligned}$$

Given that both c_j and piv_j are greater than zero, the probability an individual votes for D simplifies to:

$$Pr(v_{ij} = D) = Pr(\epsilon_{ij} < -u_j - \frac{c_j}{piv_j}) = H(-u_j - \frac{c_j}{piv_j}) \tag{9}$$

Likewise, we can calculate the probability an individual votes for R:

$$Pr(v_{ij} = R) = Pr(\epsilon_{ij} > -u_j + \frac{c_j}{piv_j}) = 1 - H(-u_j + \frac{c_j}{piv_j}) \tag{10}$$

We can then easily calculate the probability an individual abstains:

$$Pr(v_{ij} = A) = H(-\mu_j + \frac{c_i}{piv_j}) - H(-\mu_j - \frac{c_i}{piv_j}) \tag{11}$$

These probabilities are equivalent to σ_{Dj}, σ_{Rj}, and σ_{Aj}, respectively. We estimate this version both using $p(\sigma_{Dj}, \sigma_{Rj}; \alpha^1, \alpha^2)$ and seting piv_j equal to 1 across all elections. When using the former specification, we ensure that the turnout rates are consistent with the probability of being pivotal.

3.3 Leader Persuasion

In the leader persuasion model, individuals are not strategic in deciding whether to vote. Leaders exert effort to influence the utility individuals receive from voting for a candidate. This effort is denoted by e_j^s for $s \in \{D, R\}$. This enters the utility function of each individual through the function $a(e_j^s)$, where $s \in$

$\{D, R\}$. $a(e_j^s)$ can be thought of as the effect of advertising or other campaign efforts on an individual's utility from voting for candidate s. We will set $a(e_j^s) = \theta log(e_j^s)$, where θ is a parameter to be estimated.

3.3.1 Individual Voting Decisions

Individuals receive the following utility from voting for candidate D, R, or abstaining:

$$u_{ijD} = a(e_j^D) - u_j - c_j - \eta_j^D - \delta + \epsilon_{ijD}$$

$$u_{ijR} = a(e_j^R) + u_j - c_j - \eta_j^R + \delta + \epsilon_{ijR}$$

$$u_{ijA} = \epsilon_{ijA}$$

In addition to $a(e_j^s)$, there is a district mean utility differential which gives the values for if candidate R wins vs. D, $u_j \equiv u_j^R - u_j^D$. This value can be positive or negative, and is the same for all individuals within a district. Positive values indicate the district as a whole prefers candidate R to candidate D, while negative values indicate the district as a whole prefers candidate D to candidate R. The cost of voting, c_j, is constant across individuals in a district.

In addition, there are three types of shocks in this model. First, there are candidate specific shocks, η_j^s for $s \in \{D, R\}$. These have mean zero and are the same for all individuals within a district. Positive draws of η_j^s makes individuals less predisposed to vote for candidate s, while negative draws make individuals more predisposed to vote for candidate s. These only affect the utility from voting for the candidate they are associated with.

Second, there are national level shocks δ. Again these are mean zero and are the same for all individuals voting in a given year. If the realized δ is positive, this increases the utility from voting for candidate R and decreases the utility from voting for candidate D for all members in each district. The opposite is true if δ is negative. Both the η_j^s and δ shocks are drawn from a normal distribution with mean zero. However, we estimate the variance of these shocks (σ_δ, σ_η) from the data.

Finally, there are idiosyncratic shocks to voting for each candidate and abstaining, ϵ_{ijD}, ϵ_{ijR}, and ϵ_{ijA}. These are assumed to be drawn from a type 1 extreme value distribution with mean zero and variance normalized to 1.

Given this distribution, we can get the following turnout rates:

$$\sigma_j^D = \frac{e^{a(e_j^D) - u_j - c_j - \eta_j^D - \delta}}{1 + e^{a(e_j^D) - u_j - c_j - \eta_j^D - \delta} + e^{a(e_j^R) + u_j - c_j - \eta_j^R + \delta}}$$

$$\sigma_j^R = \frac{e^{a(e_j^R) + u_j - c_j - \eta_j^R + \delta}}{1 + e^{a(e_j^D) - u_j - c_j - \eta_j^D - \delta} + e^{a(e_j^R) + u_j - c_j - \eta_j^R + \delta}}$$

3.3.2 Candidate Effort

Parties choose the amount of effort to exert in each district for a given year. This effort is denoted by e_j^s for $s \in \{D, R\}$. We can think of this as the amount of resources the party spends on campaigning in each district.

Parties are only concerned with how δ and η_s for $s \in \{D, R\}$ influences the expected vote turnout rates for each candidate (σ_j^s). Although each district has a finite number of voters, and each voter has a certain probability of voting for either candidate, I assume parties do not account for the fact that the observed fraction of votes might not precisely match σ_j^s. However, in large elections, the observed fraction will closely approximate σ_j^s. The probability the expected vote shares for candidate D is greater than the expected vote shares for candidate R in district j is:

$$\begin{aligned} Pr(\sigma_j^D > \sigma_j^R) &= Pr(a(e_j^D) - c_j - \eta_j^D - u_j - \delta > a(e_j^R) - c_j - \eta_j^R + u_j + \delta) \\ &= Pr((a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta > \eta) \\ &= \Phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{\sqrt{2}\sigma_\eta}) \end{aligned}$$

Where $\eta \equiv \eta_j^D - \eta_j^R$ and Φ is the cdf of the standard normal distribution function. The expected number of seats party D wins (s^D) across the country is:

$$E(s^D) = \int \sum_j \Phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{\sqrt{2}\sigma_\eta})h(\delta)d\delta$$

Similarly, the expected number of seats party R (s^R) wins across the country is:

$$E(s^R) = \int \sum_j 1 - \Phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{\sqrt{2}\sigma_\eta})h(\delta)d\delta$$

This gives the following objective functions for the parties:

$$\max_{e_j^D} E(s^D) \text{ given } \sum_j e_j^D = E \tag{12}$$

and

$$\max_{e_j^R} E(s^R) \text{ given } \sum_j e_j^R = E \tag{13}$$

Note that we are assuming both parties have equal resources to spend on campaigning. We will explore this assumption further in the data section.

As shown in A.2, the equilibrium effort for each party is: $e_j^{D^*} = e_j^{R^*}$ for all j. Further, in the same section we show that, letting

$$Q_j \equiv \int \left(\phi(\frac{2\delta + 2u_j}{2\sigma_\eta}) \right) h(\delta) d\delta$$

and

$$Q = \sum_j Q_j$$

, we have:

$$e_j^{D^*} = e_j^{R^*} = Q_j \frac{E}{Q}$$

That is, parties exert equal levels of effort in a given district. Therefore, the expected turnout rates are:

$$\sigma_j^D = \frac{e^{a(e_j^{*D}) - c_j - \eta_j^D - \mu_j - \delta}}{1 + e^{a(e_j^{*D}) - c_j - \eta_j^D - u_j - \delta} + e^{(a(e_j^{*R}) - c_j - \eta_j^R + u_j + \delta)}} \tag{14}$$

$$\sigma_j^R = \frac{e^{a(e_j^{*R}) - c_j - \eta_j^R + \mu_j + \delta}}{1 + e^{a(e_j^{*D}) - c_j - \eta_j^D - u_j - \delta} + e^{(a(e_j^{*R}) - c_j - \eta_j^R + u_j + \delta)}} \tag{15}$$

$$\sigma_j^A = \frac{1}{1 + e^{a(e_j^{*D}) - c_j - \eta_j^D - \mu_j - \delta} + e^{(a(e_j^{*R}) - c_j - \eta_j^R + u_j + \delta)}} \tag{16}$$

with

$$e_j^{D^*} = e_j^{R^*} = Q_j \frac{E}{Q}$$

3.4 Leader Mobilization

The next model is based on a model of turnout developed in Bouton et al. (2023), which takes inspiration from Shachar and Nalebuff (1999). Candidates exert effort to lower the voting costs of their supporters. As races become more competitive, candidates exert more effort leading to higher turnout rates.

3.4.1 Individual Voting Decisions

Individuals receive the following utility from voting for candidate D or R:

$$u_{ij}^D = -u_j + \delta + \epsilon_{ij}$$

$$u_{ij}^R = u_j - \delta - \epsilon_{ij}$$

where $u_j \equiv u_j^R - u_j^D$ is the mean utility differential if candidate R wins vs. D wins in district j.

In addition, there exists a national shock δ that is the same for all individuals voting in a given year. As in the leader persuasion model, positive values of δ increases the utility from voting for candidate R and decreases the utility from voting for candidate D for all members in each district. The opposite is true if δ is negative.

Finally, there is an idiosyncratic shock $\epsilon_{ij} \sim F$ (with F a Type I Extreme Value distribution with mean 0 and variance normalized to 1) that affects the utility of voting for each candidate. In both cases, positive values of the random variables increases an individual's utility towards candidate R.

Given the symmetry between the utility functions of the two candidates, all individuals with utility $u_j - \delta - \epsilon_{ij} > 0$ prefer R, while all individuals with utility $u_j - \delta - \epsilon_{ij} < 0$ prefer D. Therefore, candidate R expects support from fraction $F(u_j - \delta)$ of the population, while candidate D expects support from fraction $1 - F(u_j - \delta)$ of the population.

Support does not necessarily translate into votes as individuals face a cost of voting:

$$c_{ij} \sim U[0, 1]$$

The idiosyncratic cost of voting is uncorrelated with the ϵ_{ij} shock on preferences. Let e_j^s be the effort candidate s exerts in district j. All individuals who support candidate s with a cost below e_j^s will vote for candidate s. Candidates face a cost of mobilizing voters. Given that candidates receive no benefit from mobilizing voters beyond the point where all supporters of the candidate vote for them, e_j^s will be at most 1 for all j.

Therefore, the probability an individual votes for candidate D is given by:

$$\begin{aligned} Pr(v_{ij} = D) &= Pr(c_{ij} < e_j^D \ \& \ u_j - \delta - \epsilon_{ij} < 0) \\ &= e_j^D F(u_j - \delta) \\ &= \sigma_j^D \end{aligned}$$

Likewise, the probability an individual votes for candidate R is:

$$\begin{aligned} Pr(v_{ij} = R) &= Pr(c_{ij} < e_j^R \ \& \ u_j - \delta - \epsilon_{ij} > 0) \\ &= e_j^R (1 - F(u_j - \delta)) \\ &= \sigma_j^R \end{aligned}$$

Where σ_j^D and σ_j^R are the expected fraction of eligible voters who turnout for each candidate.

3.4.2 Candidate Effort

Contrary to the Leader Persuasion Model, we assume candidates do not coordinate their efforts with a central party. Instead, they choose their effort levels to maximize the expected fraction of eligible voters who support them. Importantly, these effort levels are chosen before the realization of the national shock δ. Following Bouton et al. (2023), let $\hat{\delta}$ be the value of δ such that, given e_j^R and e_j^D, the candidates expect equal vote shares. In other words, for any $\delta < \hat{\delta}$, R expects to win the district. Note that, as in the Leader Persuasion Model, I assume candidates only think about the expected vote shares in the district. However, I still assume there are finite number of voters in the district, each with idiosyncratic shocks on utility and cost. This keeps with the assumption that the number of votes is drawn from a multinomial distribution that is found in the other models, making them comparable.

Candidate D's objective function is:

$$\max_{e_j^D} 1 - H(\hat{\delta}(e_j^R, e_j^D)) - c(e_j^D)$$

Similarly, candidate R's objective function is:

$$\max_{e_j^R} H(\hat{\delta}(e_j^R, e_j^D)) - c(e_j^R)$$

As shown in A.3, assuming that $c(e_j^s) = e_j^{s^{\alpha_j}}/\alpha_j$, the equilibrium effort for each party is: $e_j^{D^*} = e_j^{R^*} \equiv e_j^*$ for all j, where e_j^* is given by:

$$e_j^* = (\frac{h(\mu_j)}{4f(0)})^{1/\alpha}$$

Here, f is the pdf of the ϵ_{ij} random variable. Therefore, equilibrium turnout rates are:

$$\sigma_j^D = e_j^* F(\mu_j - \delta) \tag{17}$$

$$\sigma_j^R = e_j^*(1 - F(\mu_j + \delta)) \tag{18}$$

$$\sigma_j^A = 1 - \sigma_j^D - \sigma_j^R \tag{19}$$

where $e_j^* = (\frac{1}{4h(\mu_j)f(0)})^{1/\alpha}$ for all j.

3.5 Group Voting

This model is unique in that there is a continuum of potential voters instead of a finite number. For fraction μ_j of the population,

$$\Delta u_{ij} = -u_j^D < 0$$

and for fraction $1 - \mu_j$ of the population,

$$\Delta u_{ij} = u_j^R > 0$$

Following Coate and Conlin (2004), I assume $u_j^D = \bar{u}_j^D exp(\epsilon)$ and $u_j^R = \bar{u}_j^R exp(\epsilon)$, where $\epsilon \sim N(0,1)$. This random variable captures district level shocks that affect the utility of voting for each candidate. They impact total utility evenly across the two groups.

The cost of voting for each individual is drawn from a uniform distribution:

$$c_{ij} \sim U[0, \bar{c_j}]$$

Each side chooses a threshold λ_j^s such that all members of the group with cost below the threshold vote for the candidate they support. The treshold rule is chosen before the realization of μ_j, e.g. the fraction of people supporting D in district j. I will assume that μ_j is drawn from a beta distribution, with parameters α_1 and α_2. These parameters are known to both groups. As soon as μ is realized, each group knows if they are victorious or not. Side D wins if the fraction of people people who show up to vote for D is greater than the fraction of people who show up to vote for R. This is given by:

$$\mu_j \frac{\lambda_j^D}{c_j} > (1-\mu_j)\frac{\lambda_j^R}{c_j}$$

Rearranging, we see that side D wins if:

$$\mu_j > \frac{\lambda_j^R}{\lambda_j^R + \lambda_j^D}$$

Group D chooses a threshold rule (given by λ_j^D) to maximize total aggregate utility. This is composed of the expected utility from their candidate winning minus the expected cost of voting. Therefore, group D's objective function is:

$$\max_{\lambda_j^D} u_j^D \int_{\frac{\lambda_j^R}{\lambda_j^R+\lambda_j^D}}^{1} \mu f(\mu) d\mu - E[\mu_j] \int_0^{\lambda_j^D} cf(c)dc$$

Likewise, group R's objective function is:

$$\max_{\lambda_j^R} u_j^R \int_0^{\frac{\lambda_j^R}{\lambda_j^R+\lambda_j^D}} (1-\mu) f(\mu) d\mu - E[1-\mu_j] \int_0^{\lambda_j^R} cf(c)dc$$

As shown in appendix section A.4, this yields the following equilibrium strategies:

$$\lambda_j^{*D} = \left(c_j u_j^D h(\kappa_j)\kappa_j^2 \frac{(\alpha_{j1}+\alpha_{j2})}{\alpha_{j1}} \frac{\alpha_{j2}^{1/3} u_{jD}^{1/3}}{\alpha_{j1}^{1/3} u_{jR}^{1/3} + \alpha_{j2}^{1/3} u_{jD}^{1/3}} \right)^{1/2} \tag{20}$$

$$\lambda_j^{*R} = \left(c_j u_j^R h(\kappa_j)(1-\kappa_j)^2 \frac{(\alpha_{j1}+\alpha_{j2})}{\alpha_{j2}} \frac{\alpha_{j1}^{1/3} u_{jR}^{1/3}}{\alpha_{j1}^{1/3} u_{jR}^{1/3} + \alpha_{j2}^{1/3} u_{jD}^{1/3}} \right)^{1/2} \tag{21}$$

Where κ_j and $\bar{u_j}$ are functions of u_{jD}, u_{jR}, α_{j1}, and α_{j2}, and h is the pdf of the beta distribution.

4 Estimation

In order to make the comparison between the models fair, I parameterize the models very closely to each other. Let X_{aj} be a vector of exogenous variables for district j, where a indicates the parameter being estimated (e.g. u_j, c_j, μ_j, etc.). Let β_a be the vector of coefficients to be estimated for parameter a. The functional forms of the models are given in Table 1. Each parameter is a function of the dot product of the exogenous variables and the coefficients to be estimated. When a parameter is limited to positive values, I transform the dot product via the exponential function. Else, I use the identity function. All models are estimated via maximum likelihood.

Table 1: Functional Forms of Models

Model	Calc. of Voting (Costs)	Calc. of Voting (Preferences)	Leader Persuasion	Leader Mobilization	Agg. Group Utility
u_j	$exp(\beta_u \cdot X_{uj})$	$\beta_u \cdot X_{uj}$	$\beta_u \cdot X_{uj}$	$\beta_u \cdot X_{uj}$	$exp(\beta_u \cdot X_{uj})$
c_j	$exp(\beta_c \cdot X_{cj})$	$exp(\beta_c \cdot X_{cj})$	$exp(\beta_c \cdot X_{cj})$		$exp(\beta_c \cdot X_{cj})$
μ_j	$exp(\beta_\mu \cdot X_{\mu j})$				
α_j^1	β_{α^1}	β_{α^1}			$exp(\beta_{\alpha_1} \cdot X_{\alpha_1 j})$
α_j^2	β_{α^2}	β_{α^2}			$exp(\beta_{\alpha^2})$
θ_j				$exp(\beta_\theta \cdot X_{\theta j})$	
$\sigma_{\delta j}$				$exp(\beta_{\sigma_\delta} \cdot X_{\sigma_\delta j})$	

4.1 Likelihood for Calculus of Voting and Follow the Leader Models

In the Calculus of Voting models and the Follow the Leader models, the number of votes are a draw from a multinomial distribution:

$$v_{jD}, v_{jR}, v_{jA} \sim Multinomial(\sigma_j^D, \sigma_j^R, n_j)$$

For each gues of β, we solve for the parameters of the models, and calculate the equilibrium turnout rates. The likelihood of the data is given by:

$$L = \prod_j Pr(v_{jD}, v_{jR}, v_{jA} | \beta, n_j) \tag{22}$$

4.2 Likelihood for Aggregate Group Utility Model

For each district, there are two random variables: μ_j and ϵ_j. The former gives the fraction of people in district j who support D, while the latter is a common shock to utility for both sides. The fraction of individuals who support D is drawn from a beta distribution, with parameters α_{j1} and α_{j2}. The shock on utility is drawn from a standard normal distribution.

Given observed turnout and a vector of coefficients β, we can calculate the parameters of the model (u_{Dj}, u_{Rj}, c_j, α_{j1}, α_{j2}). These values map to unique values of μ_j and ϵ_j, as shown in appendix A.4. We want to calculate the probability of observing the vote totals in each district, which themselves are transformations of the random variables μ_j and ϵ_j. Therefore, the likelihood is given by:

$$L = \prod_j Pr(\mu_j|\alpha_{j1}, \alpha_{j2}, \beta) Pr(\epsilon_j|\beta) det(J_j|\beta)$$

where $det(J_j)$ is the determinant of the Jacobian of the transformation from μ_j and ϵ_j to the total votes.

4.3 Estimation Procedure

For the Calculus of Voting models using the Myerson (2000) formula, I use simulated annealing. Given the large number of individuals in each district, unless the turnout rates between the two sides are nearly identical, the probability of being pivotal is essentially zero. Because the effect of utility on turnout is modulated by this probability, the gradient of the likelihood function is very close to zero for most values of the coefficients associated with the utility parameters. This in turn makes estimating the model using gradient descent (which I use for the other models) ineffective. Simulated annealing is a global optimization technique that is able to look over the entire parameter space, and is therefore well-suited to this problem.

For the remaining models, I use gradient descent. For the Aggregate Group Utility Model, I follow Coate and Conlin (2004) and use maximum-likelihood estimation but add constraints on the turnout strategies to ensure that the turnout strategies $(\lambda_j^R, \lambda_j^D)$ remain below the cost of voting for each side. A value of λ_j^s above the cost of voting for side s would imply that more individuals on side s vote than are on side s.

For the Aggregate Group Utility Model, Leader Persuasion Model, and Leader Mobilization Model, I can use autodifferentiation to calculate the gradient directly using the equilibrium conditions. However, for the Calculus of Voting Models, because I use a root finding algorithm to solve for the equilibrium strategies, I use the implicit function theorem to calculate the gradient of the likelihood function. More information on the procedure can be found in Appendix section A.5.

Finally, for the Leader Persuasion Model, I use a two-step procedure where I first estimate u_j, δ, δ_σ, and δ_η using random effects regression. I then estimate the remaining parameters using maximum likelihood. More information can be found in the appendix section A.6.

All models save the Aggregate Group Utility Model have not been previously estimated. To test if the new models are identified, I simulate elections from the models using the U.S. House Elections data. That is, I randomly generate a vector of coefficients, map this to the parameters of the model, and simulate the number of votes in each district using the equilibrium strategies. I then estimate the model using the simulated data. I do this a number of times for each model. I then compare the recovered parameters to the true parameters. The results of the simulations are found in Appendix A.8. All models do a good job of recovering the true parameters.

5 Data

5.1 Election Results

I use two sets of elections to estimate the models. The first data set uses results from U.S. House Elections over 2000 - 2020. I drop elections where only one candidate is on the ballot. Across the 20 year period, I have 3,996 observations. The second data set uses results from U.S. state special elections. These occur when a seat in the state legislature is vacated before the end of the term, typically due to the incumbent resigning or dying. For this, I have data from 2010 - 2022. Given the special nature of these elections, I only have 336 elections. Table 2 gives summary statistics for the elections. As can be seen, the average turnout is much higher in the U.S. House Elections than in the state special elections. While on average about 50% of the population partcipates in the U.S. House Elections, only about 15% of the population participates in the state special elections.

Table 2: Summary Statistics for Election Results

(a) U.S. House Elections

	Mean	Std	Min	25%	75%	Max
Frac. Vote Dem.	0.246	0.098	0.032	0.172	0.308	0.589
Frac. Vote Rep.	0.244	0.114	0.001	0.155	0.329	0.632
Frac. Abstain	0.510	0.134	0.103	0.409	0.607	0.905
Margin of Victory	0.142	0.086	0.000	0.079	0.196	0.556
Total Elections	3996					

(b) State Special Elections

Index	Mean	Std	Min	25%	75%	Max
Frac. Vote Dem.	0.081	0.064	0.001	0.040	0.100	0.426
Frac. Vote Rep.	0.078	0.057	0.003	0.042	0.104	0.417
Frac. Abstain	0.842	0.111	0.179	0.801	0.911	0.986
Margin of Victory	0.034	0.035	0.000	0.013	0.043	0.404
Total Elections	336					

5.2 Demographic Data

I use demographic data from the U.S. Census. The Census does not provide data at the district level for all years of interest. Therefore I use a combination of block level data and block group level data and aggregate the data to the district level. When a block group is split across districts, I use population weights from the block level data to split the block group data. Table 3 shows the summary statistics for the demographic data used in the estimation. The average district in the U.S. House Election data has a total voting age population of around 500,000, which is about 6 times larger than the average district in the state special elections.

Table 3: Summary Statistics for Demographics

(a) U.S. House Elections

Index	Mean	Std	Min	25%	75%	Max
Total Voting Age	501888	58413	332036	461023	534745	850038
Fraction Black	0.111	0.135	0.001	0.027	0.133	0.785
Fraction Hispanic	0.151	0.170	0.003	0.036	0.194	0.871
Fraction White	0.646	0.222	0.023	0.517	0.825	0.971
No HS Fraction	0.168	0.112	0.000	0.094	0.209	0.742
Some College Fraction	0.280	0.066	0.061	0.245	0.318	0.820
College Fraction	0.171	0.077	0.002	0.122	0.210	0.673
College Plus Fraction	0.102	0.073	0.000	0.061	0.125	0.890
Fraction 18 - 29	0.156	0.026	0.081	0.137	0.171	0.278
Fraction 30 - 49	0.268	0.030	0.171	0.247	0.289	0.369
Fraction 65+	0.134	0.033	0.054	0.114	0.150	0.370

(b) State Special Elections

	Mean	Std	Min	25%	75%	Max
Total Voting Age	85003	121006	3336	29784	92115	790748
Fraction Black	0.145	0.209	0.000	0.020	0.163	1.016
Fraction Hispanic	0.135	0.183	0.005	0.033	0.149	1.036
Fraction White	0.929	0.295	0.084	0.785	1.159	1.391
No HS Fraction	0.115	0.059	0.015	0.069	0.157	0.383
Some College Fraction	0.301	0.056	0.121	0.273	0.335	0.545
College Fraction	0.181	0.074	0.041	0.128	0.218	0.419
College Plus Fraction	0.107	0.068	0.026	0.059	0.135	0.416
Fraction 18 - 29	0.205	0.061	0.092	0.167	0.234	0.628
Fraction 30 - 49	0.333	0.041	0.153	0.311	0.357	0.464
Fraction 65+	0.200	0.049	0.065	0.171	0.229	0.452

Table 4: Spending Data

(a) All FEC Disbursement Data

	Mean	Std	Min	25%	75%	Max
Total Disbursements Rep.	783	690	0.265	107	1236	2688
Total Disbursements Dem.	734	691	0.200	85	1193	2625

Note: Values are in thousands of dollars.

(b) FEC Disbursement Data: Advertising and Campaign Events Only

Index	Mean	Std	Min	25%	75%	Max
Total Spending Rep.	75	99	0.005	5	109	402
Total Spending Dem.	45	69	0.002	4	53	353

Note: Values are in thousands of dollars.

5.3 Spending Data

For the Leader Persuasion Model, I need a measure of effort by the candidates. For this, I use data on campaign spending. I use two sources of data: DIME data by Bonica (2018) and raw Federal Election Commission (FEC) data. The DIME data aggregates all FEC spending by candidates to get total spending. This data goes all the way back to 2000. However, as discussed in Schuster (2020), aggregate spending may not be a good measure of effort. This measure includes many different types of spending, i.e. administrative costs, transfers to other campaigns, advertising, refunds of contributions, etc. Schuster (2020) highlights two types of spending that likely impact individual voters: Advertising and Campaign Events. For this reason, I separately estimate the Leader Persuasion Model with FEC data filtered to only include these two types of spending. The results are not qualitatively different from the results using the DIME data.

Table 4 shows the summary statistics for the spending data. Promisingly, the mean spending in the aggregate FEC data between the two candidates is very similar. As predicted by Schuster (2020), the mean spending on advertising and campaign events is much lower than the total spending. However, over the 20 year period, the average spending by Republicans is higher than the average spending by Democrats, challenging the underlying assumption behind the Leader Persuasion Model that both parties have equal resources. Of course, this data set only includes spending by the candidates themselves. So, it is possible that PACs and other outside groups are spending more on the Democratic candidates, evening out the spending between the two parties.

Unfortunately I only have spending data for Federal elections, so I am unable to estimate the Leader Persuasion Model on the state special elections data.

5.4 Cost

For cost of voting, I use a recently compiled data set from Schraufnagel, Pomante, and Li (2022). This data set provides estimates on the cost of voting for each state due to various legal restrictions. There are a total of seven indexes that measure the cost of voting. These include how early one must register to vote prior to the election, the ease to which one can register to vote, restrictions on voter drives, how early 17 and 18 year olds can preregister to vote, the type of voter ID laws in place, and the number of early voting days.

5.5 Incumbency Status

Finally, for the U.S. House Elections, I use data on incumbency status. This of course is not available for the state special elections, as these are elections where the seat is vacated before the end of the term. For models where preferences exist on a continuum, such as the Calculus of Voting Model with Hetergenous Preferences and the Leader Mobilization Model, the incumbent is -1 if the incumbent is a Democrat and 1 if the incumbent is a Republican. This matches with the left - right continuum of preferences. In the other models, incumbency status is a dummy variable.

5.6 Data Selection

Given the large number of covariates, I first run a LASSO regression to select the most important covariates. I run two separate LASSO regressions: one with a dependent variable as the fraction of votes for the Democratic candidate out of the total population, and one with the dependent variable as the fraction of votes for the Republican candidate out of the total population. This is done on the U.S. House Elections data, as it is the larger dataset of the two. The results from the LASSO regressions are given in Table 5. I include only the covariates that are selected in the LASSO regression in the estimation of the models. The covariates represent district level demographics, election specific incumbency status, and state level voting laws. In addition to running the LASSO Regression, I also use a Random Forest model and a Gradient Boosting model to select the most important covariates. These results are found in Appendix A.7. The results are similar to the LASSO regression.

Table 6 show how the covariates are used in each model. Note that incumbent status is only used in the U.S. House Elections data.

Table 5: Results from LASSO Regression

Covariate	Democratic Turnout	Republican Turnout
Incumbent	-0.0402	0.0487
Registration Restrictions	-0.0180	-0.0012
Fraction Hispanic	-0.0091	
Fraction Black	0.0057	
Fraction with College Plus	0.0037	
Fraction with College Degree	0.0026	
Registration Deadlines	-0.0021	
Fraction White		0.0320
Fraction Aged 18-29		-0.0107
Fraction Aged 30-49		-0.0054
Fraction with Some College		0.0045
Preregistration		-0.0045
Fraction with No High School Degree		-0.0044
Median Income Ratio		-0.0025

Table 6: Covariates Used in Turnout Models

Model	Calc. of Voting Costs	Calc. of Voting Preferences	Leader Persuasion	Leader Mobilization	Agg. Group Utility
u_j	Year Dummy Variables	Incumbent Fraction Hispanic Fraction Black Fraction No HS Educ. Fraction Some Coll. Fraction Coll. Plus Fraction White Fraction 18-29 Fraction 30-49 Year Dummy Variables	Incumbent Fraction Hispanic Fraction Black Fraction No HS Educ. Fraction Some Coll. Fraction Coll. Plus Fraction White Fraction 18-29 Fraction 30-49 Intercept	Incumbent Fraction Hispanic Fraction Black Fraction No HS Educ Fraction Some Coll. Fraction Coll. Plus Fraction White Fraction 18-29 Fraction 30-49 Intercept	Year Dummy Variables
δ_j			Year Dummy Variables	Year Dummy Variables	
μ_j	Incumbent Fraction Hispanic Fraction Black Fraction No HS Educ. Fraction Some Coll. Fraction Coll. Plus Fraction White Fraction 30-49 Intercept				
c_j	Registration Restrictions Registration Deadlines Preregistration	Registration Restrictions Registration Deadlines Preregistration	Registration Restrictions Registration Deadlines Preregistration	Registration Restrictions Registration Deadlines Preregistration	Registration Restrictions Registration Deadlines Preregistration
α_j^1	Intercept	Intercept			Incumbent Fraction Hispanic Fraction Black Fraction No HS Educ. Fraction Some Coll. Fraction Coll. Plus Fraction White Fraction 30-49 Intercept
α_j^2	Intercept	Intercept			Intercept

6 Results and Discussion

I estimate the models on U.S. House Elections data and U.S. state special elections data. The only exception is the Leader Persuasion Model, which requires information on campaign spending. For this data, I use the FEC data, which is only available for federal elections.

6.1 Results from the Calculus of Voting Models with the Myerson Approximation

I first estimate the Calculus of Voting Model with Hetergenous Costs using the Myerson (2000) approximation. Given that I am using a global search algorithm, I limit the number of coefficients to estimate to be around ten. To increase the probability of success, I stack the deck in favor of the pivotal voter model by including presidential election results in the covariates. This should make it much easier for the model to match the data than only using demographic data, as I do in the other models.

Figure 2 shows the density plots from this estimation. Here I plot separately the density of predicted vote shares and the abstention rates for all 3,996 House elections. Perhaps surpringly, the model is able to match the Democratic and Republican vote shares quite well. However, the model predicts a near constant abstention rate of 50% across all districts. A closer examination of the density of the pivotal probability in Figure 3 shows that the model predicts the probability of being pivotal is near zero for all districts. Looking at the equilibrium turnout rates:

$$
\begin{aligned}
\sigma_j^D &= F(0)\frac{e^{(p(\sigma_{Dj},\sigma_{Rj},n_j)\mu_j - c_j)}}{1+e^{(p(\sigma_{Dj},\sigma_{Rj},n_j)\mu_j - c_j)}} \\
\sigma_j^R &= (1-F(0))\frac{e^{(p(\sigma_{Dj},\sigma_{Rj},n_j)\mu_j - c_j)}}{1+e^{(p(\sigma_{Dj},\sigma_{Rj},n_j)\mu_j - c_j)}}
\end{aligned}
$$

we can see that if we set both the pivotal probability and cost of voting equal to zero, we get the following turnout rates:

$$
\begin{aligned}
\sigma_j^D &= F(0)\frac{e^{(-0)}}{1+e^{(-0)}} = F(0)\frac{1}{2} \\
\sigma_j^R &= (1-F(0))\frac{e^{(-0)}}{1+e^{(-0)}} = (1-F(0))\frac{1}{2}
\end{aligned}
$$

Therefore, we still get an expected 50% turnout rate for a given side and a 50% abstention rate. The second graph in Figure 3 shows that the maximum likelihood solution for the cost parameter is indeed very close to zero. Further, the model predicts that the probability of being pivotal is near zero for all districts. The model is therefore able to achieve a good fit by matching the partisan vote share in each district using the presidential vote shares. However, the only people voting are doing so not because they believe they are pivotal, but because of civic duty or other similar reasons. Given that this implies the key mechanism of the Calculus of Voting is not at play, e.g. people voting because they believe they are pivotal, I do not consider this a successful model.

Figure 2: Density of Vote Shares for Calculus of Voting Model with Hetergenous Costs using the Myerson Approximation

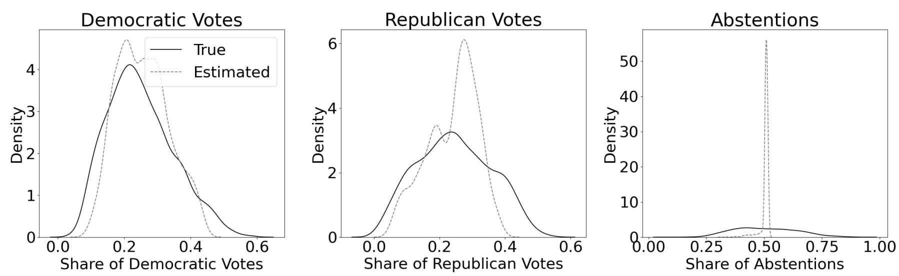
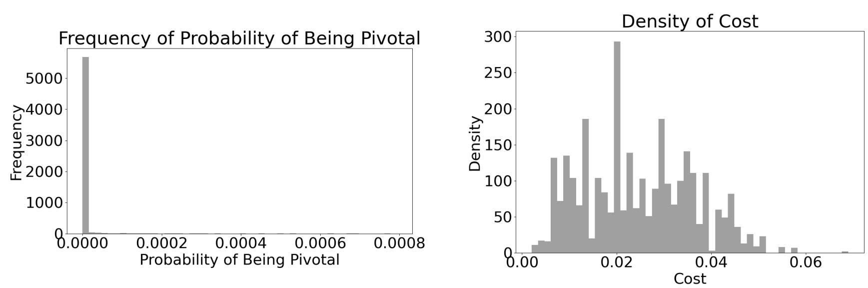


Figure 3: Density of Pivotal Probability and Cost for Calculus of Voting Model with Hetergenous Costs using the Myerson Approximation



6.2 Remaining Models

The coefficient estimates using the U.S. House Election Results for the remaining models are given in Table 11. I estimate both Calculus of Voting models with and without the probability of being pivotal influencing the turnout rates. To estimate the latter, I set $p(\sigma_j^D, \sigma_j^R, n_j) = 1$ for all districts. This shuts down any strategic voting in the model.

To more clearly understand how well the models fit the data, I produce two types of plots. First, similar to the plots for the Calculus of Voting Models with the Myerson approximation, I plot the density of both predicted and actual vote shares, along with abstention rates. These are given in Figure 4 for the U.S. House Elections data, and in Figure 7 for the U.S. State Special Elections data. However, the density plots may be able to match the overall distribution of the data without correctly predicting the turnout rates of any given election. Therefore, I also plot the predicted votes and abstentions against the actual votes and abstentions for each election in the data. To enhance the clarity, for each plot I order the districts in terms of the actual number of votes or abstentions. Districts furthest to the left have the fewest votes or abstentions, while districts furthest to the right have the most votes or abstentions. I additionally include a best fit line for each plot using a polynomial of degree 3. These are given in Figure 5 (U.S. House Elections) and Figure 8 (U.S. State Special Elections).

The differences among the models is most apparent from the U.S. House Elections data. From Figure 5 one can see that most models struggle to match the data on at least one dimension. For instance, while the Calculus of Voting Model with Hetergenous Preferences, the Leader Persuasion Model, and the Leader Mobilization Model all show upward sloping lines for total Democratic and Republican votes, they show flat lines for abstentions. This means these models predict near constant rates of abstention across all districts. Both the Aggregate Group Utility Model and the Calculus of Voting Model with Hetergenous Costs do a very good job matching the overall density of the vote shares and abstention rates.

The success of the Calculus of Voting Model with Hetergenous Costs is further confirmed by the log-likelihood results from the U.S House Election data shown in in Table 11. Among all models with votes drawn from a multinomial distribution, the Calculus of Voting Model with Hetergenous Costs has the highest log-likelihood. However, to test if this is significant, I use the Vuong test to compare the models. Note I leave out the Aggregate Group Utility Model model from the Vuong test, as it contains different assumptions from the other models, namely a continuum of voters vs. a finite number of voters. The Vuong test creates a test statistic that is asymptotically normal. Positive values indicate Model 1 is more likely, while negative values indicate Model 2 is more likely. Table **??** shows the results of the Vuong test for the U.S House Elections data. A pairwise comparison of the models indicates that the Calculus of Voting Model with Hetergenous Costs outperforms all other models statistically. Moreover, the model's performance declines when the probability of being pivotal is excluded, highlighting the significance of this factor in predicting voter turnout. I additionally run the Vuong test on the U.S. State Special Elections data. While the Leader Mobilization Model model beats each of the other models in a pairwise comparison, the difference between the Calculus of Voting Model with Hetergenous Costs and the Leader Mobilization Model is not statistically significant at the standard 5% level. As a robustness check, I also run the Clarke test on the results. The Clarke test differs from the Vuong test in that it is a non-parametric test. Instead of comparing likelihood ratios, the Clarke test examines how often one model outperforms the other across individual observations. It then conducts a binomial test to determine if this difference is statistically significant. The results from the Clarke test are consistent with the Vuong test (see tables 9 and 10).

6.3 Interpretation of the Results

The results from the Vuong test suggest that the Calculus of Voting Model with Hetergenous Costs is the best model for predicting turnout in U.S. House Elections, and a leading contender to modeling turnout in U.S. State Special Elections. While this model is closely related to the classic Palfrey and Rosenthal (1983) model, it differs by allowing for a more general form of the probability of being pivotal. Figure 6 shows exactly how the probability of being pivotal changes as the ratio of votes between the losing and winning candidate changes. I also include the estimated probability of being pivotal from the Calculus of Voting Model with Hetergenous Preferences.

In both cases, the probability of being pivotal increases as the ratio of vote shares approaches one, matching the intuition that the probability of being pivotal is highest when the election is closest. For the Calculus of Voting Model with Hetergenous Preferences, this change happens much more quickly than in the Hetergenous Cost model, and at a much lower level. When the ratio of vote shares is near 20%, the probability of being pivotal is already near one for this model. In large elections such as those for the U.S. House of Representatives, the analytical solution for the probability of being pivotal should remain largely flat until the ratio of vote shares is near one, at which point it should rapidly increase. The steady increase in the probability of being pivotal estimated from the Hetergenous Cost model suggests this is not how voters perceive the election.

In addition to matching the basic intuition that voters are more likely to vote when the election is close, the Calculus of Voting Model with Hetergenous Costs also matches the underdog effect. This is a common comparative static found in many models of voter turnout whereby the ratio of voters for the losing candidate relative to the losing candidate's overall support in the population is higher than the ratio of voters for the winning candidate relative to the winning candidate's overall support in the election. In equilibrium the turnout rates of each side are:

$$\tau_D = \frac{e^{p(\sigma_{Dj},\sigma_{Rj},n_j)}\mu_j - c_j}{1 + e^{p(\sigma_{Dj},\sigma_{Rj},n_j)}\mu_j - c_j} = \tau_R$$

Where τ_s gives the turnout rate of side $s \in \{D, R\}$. Without loss of generality, suppose $F(0) < 1 - F(0)$. In this case, R is winning candidate, as $\tau_R(1 - F(0)) > \tau_D F(0)$. The ratio of turnout rates relative to support for each side is:

$$\frac{\tau_R}{1 - F(0)} < \frac{\tau_D}{F(0)}$$

Therefore, the ratio of the turnout rate of the winning candidate to the winning candidate's overall support is less than the ratio of the turnout rate of the losing candidate to the losing candidate's overall support. Thus, despite relaxing the assumption voters are calculating the exact probability of being pivotal, the Calculus of Voting Model with Hetergenous Costs with the sigmoid function maintains key comparative statics found in other models of voter turnout.

Table 7: Pairwise Comparison using Vuong Test (U.S. House Elections)

	Model 1	**Calc. of Voting Costs**	Calc. of Voting Costs (No Piv)	Calc. of Voting Preferences	Calc. of Voting Preferences (No Piv)	Leader Persuasion	Leader Mobilization
Model 2	Calc. of Voting Costs		-11.13 (0.00)	-6.20 (0.00)	-11.43 (0.00)	-15.12 (0.00)	-16.23 (0.00)
	Calc. of Voting Costs (No Piv)	**11.13 (0.00)**		-2.20 (0.03)	-8.31 (0.00)	-14.11 (0.00)	-10.21 (0.00)
	Calc. of Voting Preferences	**6.20 (0.00)**	2.20 (0.03)		-11.79 (0.00)	-13.41 (0.00)	-10.26 (0.00)
	Calc. of Voting Preferences (No Piv)	**11.43 (0.00)**	8.31 (0.00)	11.79 (0.00)		-11.88 (0.00)	-1.92 (0.05)
	Leader Persuasion	**15.12 (0.00)**	14.11 (0.00)	13.41 (0.00)	11.88 (0.00)		8.04 (0.00)
	Leader Mobilization	**16.23 (0.00)**	10.21 (0.00)	10.26 (0.00)	1.92 (0.05)	-8.04 (0.00)	
	Log-Likelihood	**-65,922,099**	-70,302,313	-72,974,983	-80,458,665	-104,224,688	-82,935,868

Vuong test with p-values in parentheses. Null Hypothesis both models fit data equally well. Positive (negative) values indicate model 1 (model 2) is better fit.

Table 8: Pairwise Comparison using Vuong Test (U.S. House Elections)

	Model 1	**Calc. of Voting Costs**	Calc. of Voting Costs (No Piv)	Calc. of Voting Preferences	Calc. of Voting Preferences (No Piv)	Leader Persuasion	Leader Mobilization
Model 2	Calc. of Voting Costs		-11.13 (0.00)	-6.20 (0.00)	-11.43 (0.00)	-15.12 (0.00)	-16.23 (0.00)
	Calc. of Voting Costs (No Piv)	**11.13 (0.00)**		-2.20 (0.03)	-8.31 (0.00)	-14.11 (0.00)	-10.21 (0.00)
	Calc. of Voting Preferences	6.20 (0.00)	2.20 (0.03)		-11.79 (0.00)	-13.41 (0.00)	-10.26 (0.00)
	Calc. of Voting Preferences (No Piv)	11.43 (0.00)	8.31 (0.00)	11.79 (0.00)		-11.88 (0.00)	-1.92 (0.05)
	Leader Persuasion	15.12 (0.00)	14.11 (0.00)	13.41 (0.00)	11.88 (0.00)		8.04 (0.00)
	Leader Mobilization	16.23 (0.00)	10.21 (0.00)	10.26 (0.00)	1.92 (0.05)	-8.04 (0.00)	
	Log-Likelihood	-65,922,099	-70,302,313	-72,974,983	-80,458,665	-104,224,688	-82,935,868

Vuong test with p-values in parentheses. Null Hypothesis both models fit data equally well. Positive (negative) values indicate model 1 (model 2) is better fit.

Table 9: Clarke Test Statistics (U.S. House Elections)

	Model 1	Calc. of Voting Costs	Calc. of Voting Costs (No Piv)	Calc. of Voting Preferences	Calc. of Voting Preferences (No Piv)	Leader Persuasion	Leader Mobilization
Model 2	Calc. of Voting Costs		1878, 2117 (0.00)	1942, 2053 (0.08)	1823, 2172 (0.00)	1496, 2499 (0.00)	1389, 2606 (0.00)
	Calc. of Voting Costs (No Piv)	2117, 1878 (0.00)	aN	1985, 2010 (0.70)	1841, 2154 (0.00)	1498, 2497 (0.00)	1456, 2539 (0.00)
	Calc. of Voting Preferences	2053, 1942 (0.08)	2010, 1985 (0.70)		1800, 2195 (0.00)	1237, 2758 (0.00)	1653, 2342 (0.00)
	Calc. of Voting Preferences (No Piv)	2172, 1823 (0.00)	2154, 1841 (0.00)	2195, 1800 (0.00)		1106, 2889 (0.00)	1708, 2287 (0.00)
	Leader Persuasion	2499, 1496 (0.00)	2497, 1498 (0.00)	2758, 1237 (0.00)	2889, 1106 (0.00)		2182, 1813 (0.00)
	Leader Mobilization	2606, 1389 (0.00)	2539, 1456 (0.00)	2342, 1653 (0.00)	2287, 1708 (0.00)	1813, 2182 (0.00)	

The first number in each cell is the number of times Model 1 outperforms Model 2, while the second number is the number of times Model 2 outperforms Model 1. The number in parentheses is the p-value from a binomial test, with the null hypothesis that the models are equally likely to outperform each other.

Table 10: Clarke Test Statistics (U.S. State Special Elections)

	Model 1	Calc. of Voting Costs	Calc. of Voting Costs (No Piv)	Calc. of Voting Preferences	Calc. of Voting Preferences (No Piv)	Leader Mobilization
Model 2	Calc. of Voting Costs		166, 169 (0.91)	181, 154 (0.16)	162, 173 (0.58)	186, 149 (0.05)
	Calc. of Voting Costs (No Piv)	169, 166 (0.91)		178, 157 (0.27)	166, 169 (0.91)	182, 153 (0.13)
	Calc. of Voting Preferences	154, 181 (0.16)	157, 178 (0.27)		153, 182 (0.13)	166, 169 (0.91)
	Calc. of Voting Preferences (No Piv)	173, 162 (0.58)	169, 166 (0.91)	182, 153 (0.13)		180, 155 (0.19)
	Leader Mobilization	149, 186 (0.05)	153, 182 (0.13)	169, 166 (0.91)	155, 180 (0.19)	

The first number in each cell is the number of times Model 1 outperforms Model 2, while the second number is the number of times Model 2 outperforms Model 1. The number in parentheses is the p-value from a binomial test, with the null hypothesis that the models are equally likely to outperform each other.

Table 11: Estimated Coefficients for U.S. House Election Data

Coefficient	Calc. Of Voting (Cost)	Calc. Of Voting (Cost, No Piv)	Calc. Of Voting (Utility)	Calc. Of Voting (Utility, No Piv)
Constant (μ)	-1.68 (1.272e-05)	-1.517 (1.339e-05)	NaN	NaN
Incumbent	-0.009 (1.370e-06)	-0.065 (1.478e-06)	0.177 (3.569e-07)	0.191 (3.919e-07)
Frac. Hispanic	1.501 (1.072e-05)	2.076 (1.189e-05)	-0.023 (3.635e-06)	0.339 (4.898e-06)
Frac. Black	0.841 (1.014e-05)	0.719 (1.080e-05)	-0.106 (3.232e-06)	-0.061 (4.056e-06)
Frac. No HS	0.048 (1.167e-05)	-0.025 (1.237e-05)	0.018 (5.226e-06)	-0.075 (6.744e-06)
Frac. Some Coll	1.872 (1.288e-05)	1.742 (1.376e-05)	0.362 (5.447e-06)	0.119 (6.718e-06)
Frac. Coll+	-0.46 (9.865e-06)	-0.881 (9.235e-06)	-0.057 (5.601e-06)	-0.344 (5.528e-06)
Frac. Coll	0.787 (1.114e-05)	0.722 (1.154e-05)	0.109 (6.067e-06)	0.059 (7.037e-06)
Frac. White	2.629 (9.361e-06)	2.702 (9.922e-06)	0.322 (2.731e-06)	0.38 (3.372e-06)
Frac. 18-29	-4.156 (2.362e-05)	-4.527 (2.496e-05)	-0.655 (1.102e-05)	-0.645 (1.370e-05)
Frac. 30-49	-1.456 (1.908e-05)	-1.596 (2.017e-05)	-0.809 (1.007e-05)	-0.643 (1.323e-05)
Median Income Ratio	-0.04 (1.633e-07)	-0.028 (1.930e-07)	-0.007 (4.341e-08)	-0.0 (7.099e-08)
Constant (Cost)	-0.66 (5.756e-06)	-0.894 (4.669e-06)	-0.883 (9.795e-07)	-0.777 (5.956e-07)
Registration Restrictions	0.149 (7.906e-07)	0.182 (7.407e-07)	0.093 (2.081e-07)	0.102 (2.099e-07)
Preregistration	0.027 (4.488e-07)	0.052 (4.718e-07)	0.022 (1.118e-07)	0.022 (1.131e-07)
Registration Deadlines	0.034 (6.044e-07)	0.037 (5.906e-07)	0.017 (2.065e-07)	0.013 (2.086e-07)
Year 2002	-0.335 (3.736e-06)	-0.491 (3.385e-06)	-0.012 (1.183e-06)	0.025 (1.514e-06)
Year 2004	0.486 (1.741e-06)	0.347 (1.304e-06)	-0.028 (1.052e-06)	-0.008 (1.226e-06)
Year 2006	-0.166 (2.760e-06)	-0.316 (2.596e-06)	-0.093 (1.271e-06)	-0.066 (1.405e-06)
Year 2008	0.468 (1.689e-06)	0.317 (1.353e-06)	-0.062 (1.098e-06)	-0.076 (1.303e-06)
Year 2010	-0.523 (4.479e-06)	-0.719 (4.174e-06)	0.044 (1.161e-06)	0.062 (1.504e-06)
Year 2012	0.321 (1.831e-06)	0.121 (1.794e-06)	-0.071 (1.129e-06)	-0.054 (1.383e-06)
Year 2014	-0.709 (5.782e-06)	-1.037 (5.932e-06)	-0.025 (1.210e-06)	0.003 (1.603e-06)
Year 2016	0.343 (1.888e-06)	0.137 (1.831e-06)	-0.059 (1.216e-06)	-0.02 (1.438e-06)
Year 2018	0.001 (2.591e-06)	-0.231 (2.711e-06)	-0.087 (1.252e-06)	-0.065 (1.477e-06)
Year 2020	0.281 (2.031e-06)	0.07 (2.047e-06)	-0.002 (1.136e-06)	0.015 (1.369e-06)
Constant (α_1)	-35.424 (5.562e-03)	0.0 (0.000e+00)	-134.625 (2.046e-02)	0.0 (0.000e+00)
Constant (α_2)	-0.084 (1.313e-05)	0.0 (0.000e+00)	-0.026 (3.909e-06)	0.0 (0.000e+00)
Log Likelihood	-65,922,099	-70,302,313	-72,974,983	-80,458,665

7 Conclusion

In this paper, I compare several models of voter turnout. While a number of authors have proposed theoretical models, and others have empirically tested isolated aspects of these models, this paper is unique in bringing together several models and testing them all on the same data set. I test five models of turnout.

I begin with versions similar to the earliest models of turnout as introduced by Downs (1957) and Palfrey and Rosenthal (1983). I test two different versions. In one, all candidates on a side receive the same utility if their side wins. However, voters face idiosyncratic costs in voting. In the other version, voters have identical costs of voting, but different utilities from each candidate winning.

I first test how well both of these models match the data using a close approximation to the analytical form of the probability of being pivotal as calculated by Myerson (2000). Estimates using this formula suggest that differences in the probability of being pivotal do not explain differences in turnout rates across districts. Instead, the model predicts a near constant abstention rate of 50% across all districts.

Given that the key mechanism of the Calculus of Voting is not at play, I adjust the models by introducing a more general form of the probability of being pivotal. While the form I use can allow for higher perceived probability of being pivotal at lower turnout rates relative to the Myerson approximation, it still enforces either monotonically increasing or monotonically decreasing probabilities of being pivotal with respect to the ratio of votes. In addition to this adjustment, I also test the models without the probability of being pivotal influencing the turnout rates. This allows me to test the extent to which the models can match the data without strategic voting.

I next build a model of turnout based around Strömberg (2008). In this model, candidates exert effort to persuade voters to support them and turnout on election day. Contrary to previous versions of this model that have been tested, I add the ability for individual voters to abstain. Additionally, I test a leader-based model built around Bouton et al. (2023). In this model, leaders choose an effort level. Individuals who support a given leader vote for this leader if their cost of voting is below the effort level chosen by the leader. Finally, I test a group based model which is closely related to the Coate and Conlin (2004). Here, individuals follow a voting rule decided by their group to maximize overall group utility.

Many models struggled to fit all dimensions of the data. Of the models which assume a finite number of voters with votes drawn from a multinomial distribution, the Calculus of Voting Model with Hetergenous Costs is the most successful. In a pairwise comparison with the other models using the Vuong test, the Calculus of Voting Model with Hetergenous Costs using the adjusted probability of being pivotal outperformed all other models on the U.S. House

Table 12: Estimated Coefficients for U.S. House Election Data

Coefficient	Leader Persuasion	Leader Persuasion, Fec	Leader Mobilization
constant	-0.345 (1.047e-01)	-0.345 (1.047e-01)	-1.579 (7.092e-06)
Incumbent	0.25 (4.008e-03)	0.25 (4.008e-03)	NaN
Frac. Hispanic	0.333 (6.730e-02)	0.333 (6.730e-02)	0.669 (4.198e-06)
Frac. Black	-0.162 (6.440e-02)	-0.162 (6.440e-02)	0.749 (4.207e-06)
Frac. No HS	0.028 (8.201e-02)	0.028 (8.201e-02)	0.268 (5.770e-06)
Frac. Some Coll	0.449 (8.679e-02)	0.449 (8.679e-02)	1.246 (6.654e-06)
Frac. Coll+	-0.239 (7.716e-02)	-0.239 (7.716e-02)	0.004 (5.497e-06)
Frac. Coll	0.174 (8.590e-02)	0.174 (8.590e-02)	0.503 (5.909e-06)
Frac. White	0.646 (6.362e-02)	0.646 (6.362e-02)	1.613 (4.271e-06)
Frac. 18-29	-0.817 (1.589e-01)	-0.817 (1.589e-01)	-2.87 (1.236e-05)
Frac. 30-49	-0.533 (1.864e-01)	-0.533 (1.864e-01)	0.272 (1.197e-05)
Median Income Ratio	-0.013 (7.429e-04)	-0.013 (7.429e-04)	-0.006 (4.177e-08)
Constant (Cost)	-1.008 (1.850e-06)	-1.219 (1.639e-06)	NaN
Registration Restrictions	0.176 (4.447e-07)	0.198 (4.783e-07)	-0.094 (3.034e-07)
Preregistration	0.071 (2.485e-07)	0.081 (2.671e-07)	-0.039 (1.780e-07)
Registration Deadlines	0.031 (4.319e-07)	0.04 (4.713e-07)	-0.042 (3.015e-07)
theta	0.016 (7.831e-08)	0.002 (2.970e-08)	NaN
Year 2002	0.045	0.045	-0.197 (1.449e-06)
Year 2004	-0.009	-0.009	-0.175 (1.124e-06)
Year 2006	-0.084	-0.084	-0.095 (1.358e-06)
Year 2014	0.051	0.051	-0.275 (1.613e-06)
Year 2018	-0.05	-0.05	-0.201 (1.380e-06)
Year 2020	0.055	0.055	-0.276 (1.188e-06)
Year 2000	-0.006	-0.006	NaN
Year 2008	-0.071	-0.071	-0.109 (1.122e-06)
Year 2010	0.119	0.119	-0.282 (1.466e-06)
Year 2012	-0.041	-0.041	-0.181 (1.245e-06)
Year 2016	-0.01	-0.01	-0.264 (1.325e-06)
σ_δ	0.063 (1.057e-01)	0.063 (1.057e-01)	NaN
Constant (α)	NaN	NaN	0.979 (1.440e-06)
Constant (δ)	NaN	NaN	-0.383 (1.096e-06)
Log Likelihood	-104,224,688	-102,769,337	-82,935,868

Table 13: Estimated Coefficients for U.S. House Election Data

Coefficient	Group Utility
Constant (α_1)	1.994 (4.074e-01)
Frac. Hispanic	-0.107 (2.302e-01)
Frac. Black	0.323 (2.250e-01)
Frac. No HS	0.806 (2.939e-01)
Frac. Some Coll	-1.079 (3.761e-01)
Frac. Coll+	0.922 (2.938e-01)
Frac. Coll	0.657 (3.345e-01)
Frac. White	-1.665 (2.313e-01)
Frac. 18-29	2.251 (6.549e-01)
Frac. 30-49	1.53 (6.822e-01)
Constant (α_2)	1.511 (2.575e-02)
Year 2002 (D)	-0.819 (1.846e-01)
Year 2004 (D)	0.131 (2.000e-01)
Year 2006 (D)	-0.36 (1.882e-01)
Year 2008 (D)	0.291 (2.111e-01)
Year 2010 (D)	-1.019 (1.809e-01)
Year 2012 (D)	-0.047 (2.178e-01)
Year 2014 (D)	-1.113 (2.019e-01)
Year 2016 (D)	-0.18 (2.445e-01)
Year 2018 (D)	-0.359 (2.486e-01)
Year 2020 (D)	-0.311 (2.727e-01)
Year 2002 (R)	-0.496 (1.849e-01)
Year 2004 (R)	0.302 (1.917e-01)
Year 2006 (R)	-0.562 (1.795e-01)
Year 2008 (R)	0.157 (1.864e-01)
Year 2010 (R)	-0.527 (1.912e-01)
Year 2012 (R)	0.07 (2.056e-01)
Year 2014 (R)	-0.721 (1.912e-01)
Year 2016 (R)	0.13 (2.257e-01)
Year 2018 (R)	-0.303 (2.329e-01)
Year 2020 (R)	0.044 (2.796e-01)
Constant (Cost)	-0.142 (2.766e-01)
Registration Restrictions	0.148 (5.924e-02)
Preregistration	0.037 (4.434e-02)
Registration Deadlines	0.035 (5.915e-02)
Log Likelihood	97,848

Elections data. This includes a similar version, but where I set the probability of being pivotal equal to one for all districts. Interestingly, the Leader Mobilization Model is quite successful in matching the data on U.S. State Special Elections data despite performing poorly on the U.S. House Elections data. However, it is not statistically different from the Calculus of Voting Model with Hetergenous Costs in the pairwise comparison from the Vuong test.

This paper sheds light on which models of turnout are most succesful in matching actual election results. As we begin understanding the mechanisms behind turnout, we can begin to understand how the electorate responds to changes in the political environment. This will help policymakers and advocates improve the voting environment and enhance a variety of political economy models with more accurate assumptions about voter behavior.

Figure 4: Density of Vote Shares (House Elections)

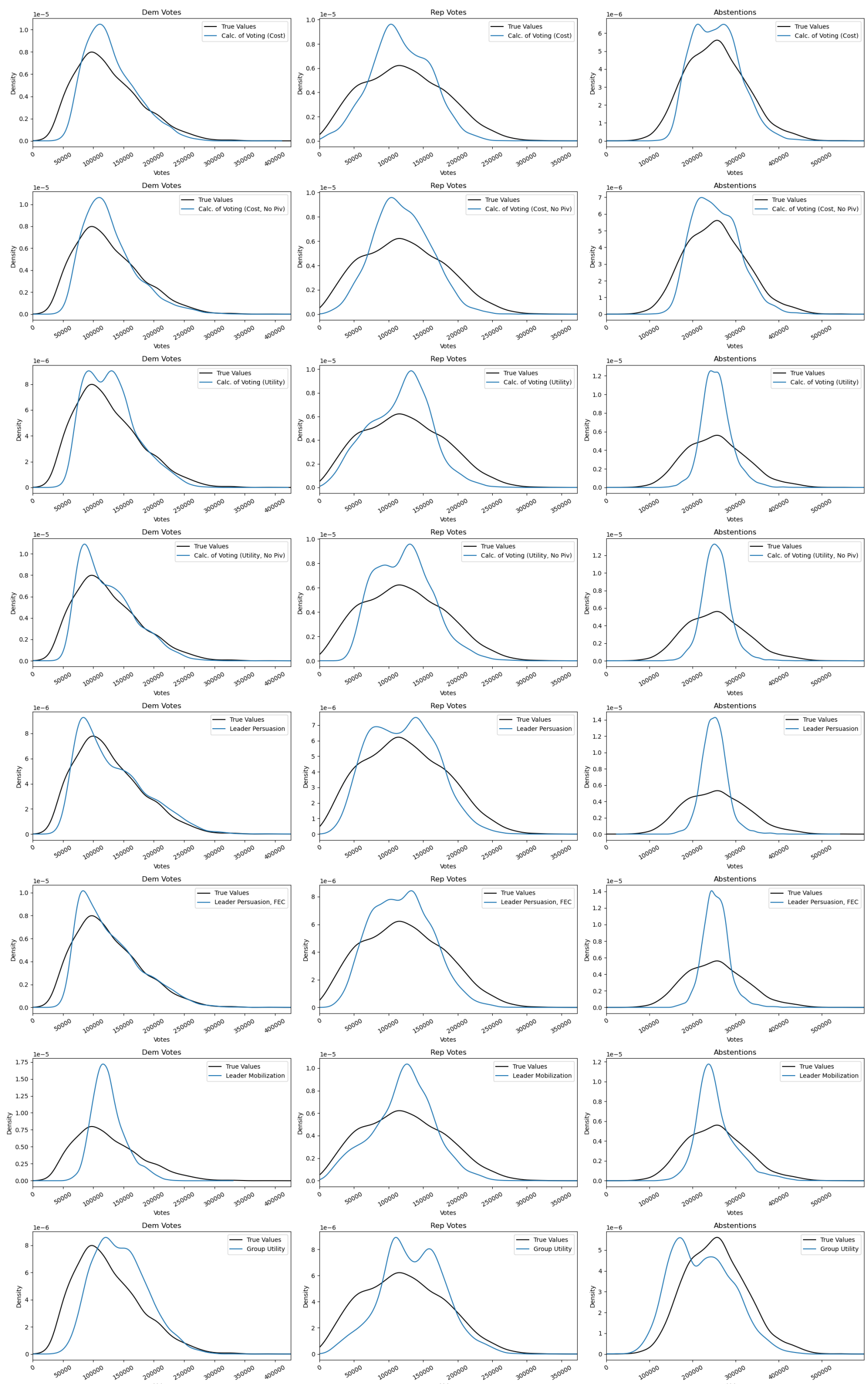


Figure 5: Ordered Vote Shares (House Elections)

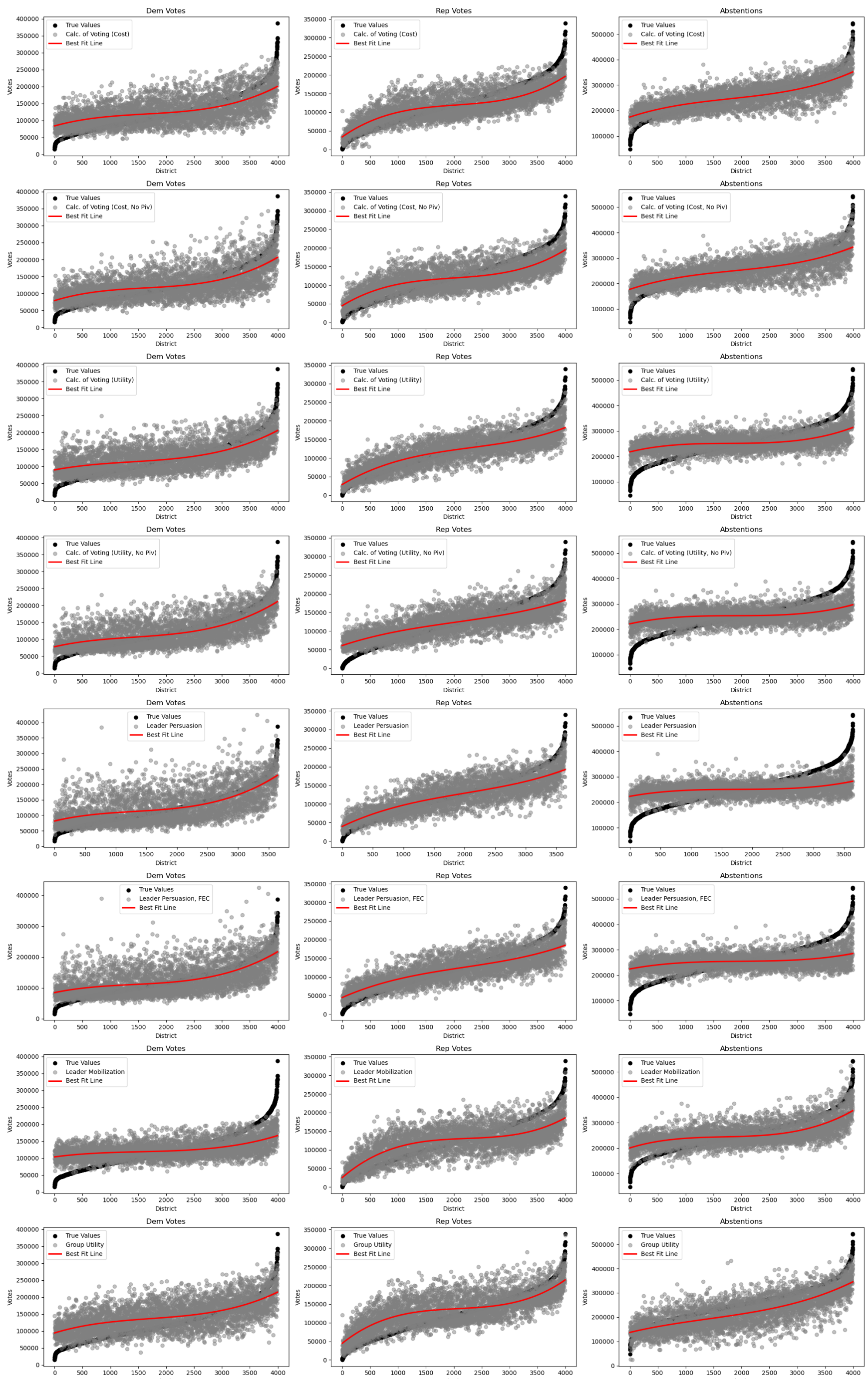
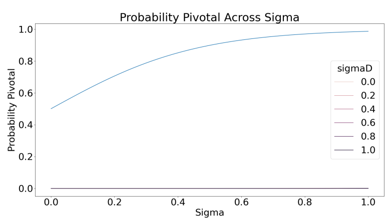
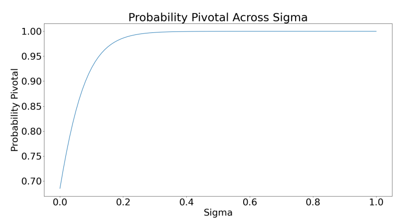


Figure 6: Anticipated Probability of Being Pivotal



Calculus of Voting Model with Hetergenous Costs



Calculus of Voting Model with Hetergenous Preferences

Figure 7: Density of Vote Shares (State Special Elections)

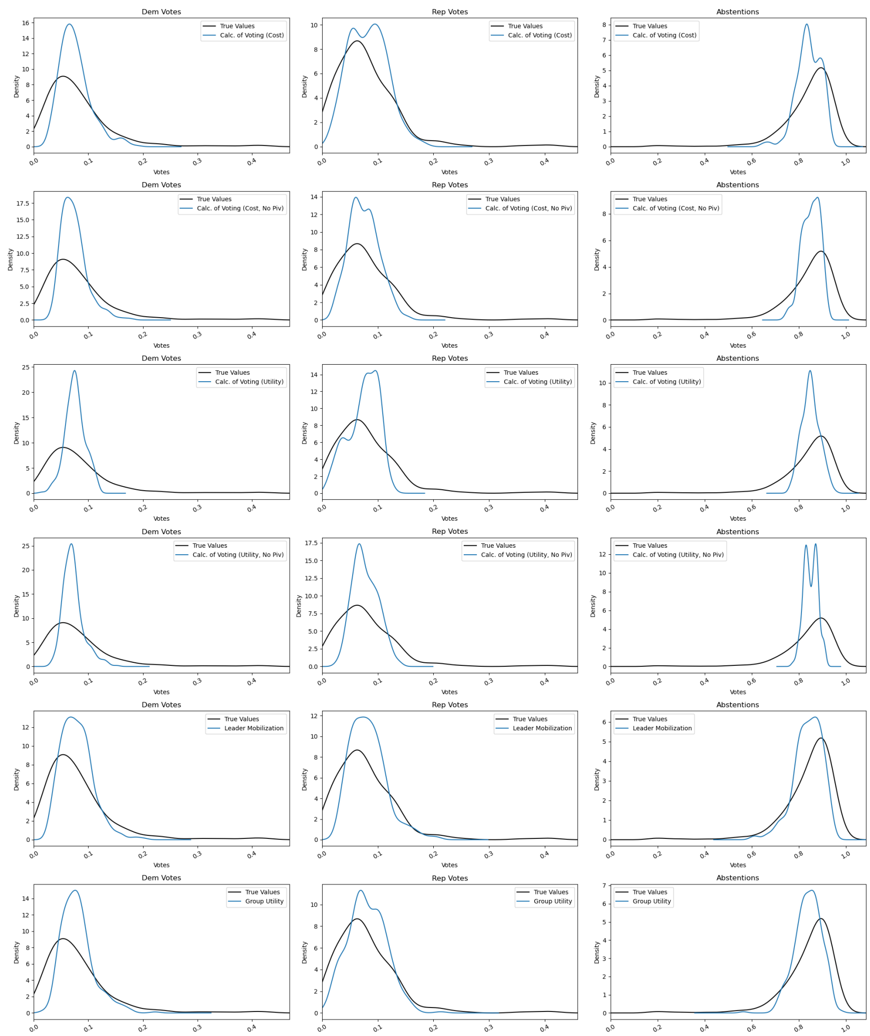
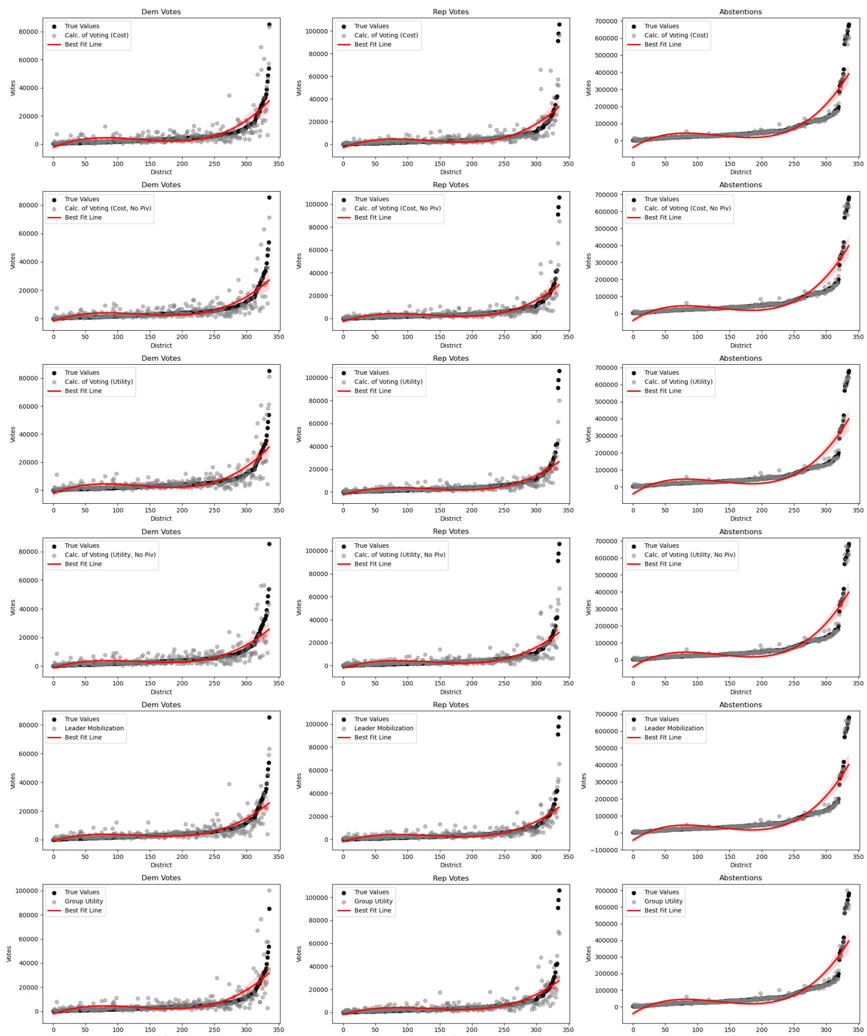


Figure 8: Ordered Vote Shares State (State Special Elections)



8 References

References

Bonica, Adam (2018). "Inferring roll-call scores from campaign contributions using supervised machine learning". In: *American Journal of Political Science* 62.4, pp. 830–848.

Bouton, Laurent et al. (2023). *Pack-Crack-Pack: Gerrymandering with Differential Turnout*. Tech. rep. National Bureau of Economic Research.

Castanheira, Micael (2003). "Victory margins and the paradox of voting". In: *European Journal of Political Economy* 19.4, pp. 817–841.

Coate, Stephen and Michael Conlin (2004). "A group rule—utilitarian approach to voter turnout: Theory and evidence". In: *American Economic Review* 94.5, pp. 1476–1504.

Coate, Stephen, Michael Conlin, and Andrea Moro (2008). "The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda". In: *Journal of Public Economics* 92.3-4, pp. 582–596.

Downs, Anthony (1957). *An Economic Theory of Democracy.*

Feddersen, Timothy and Alvaro Sandroni (2006). "A theory of participation in elections". In: *American Economic Review* 96.4, pp. 1271–1282.

Incerti, Devin (2018). "The optimal allocation of campaign funds in US House elections". In: *Electoral Studies* 56, pp. 102–113.

Kawai, Kei, Yuta Toyama, and Yasutora Watanabe (2021). "Voter turnout and preference aggregation". In: *American Economic Journal: Microeconomics* 13.4, pp. 548–586.

Lindbeck, Assar and Jörgen W Weibull (1987). "Balanced-budget redistribution as the outcome of political competition". In: *Public choice* 52, pp. 273–297.

Myatt, David P (2015). "A theory of voter turnout". In: *Manuscript, London Bus. School*.

Myerson, Roger B (2000). "Large poisson games". In: *Journal of Economic Theory* 94.1, pp. 7–45.

Palfrey, Thomas R and Howard Rosenthal (1983). "A strategic calculus of voting". In: *Public choice* 41.1, pp. 7–53.

Riker, William H and Peter C Ordeshook (1968). "A Theory of the Calculus of Voting". In: *American political science review* 62.1, pp. 25–42.

Schraufnagel, Scot, Michael J Pomante, and Quan Li (2022). "Cost of voting in the American States: 2022". In: *Election Law Journal: Rules, Politics, and Policy* 21.3, pp. 220–228.

Schuster, Steven Sprick (2020). "Does campaign spending affect election outcomes? New evidence from transaction-level disbursement data". In: *The Journal of Politics* 82.4, pp. 1502–1515.

Shachar, Ron and Barry Nalebuff (1999). "Follow the leader: Theory and evidence on political participation". In: *American Economic Review* 89.3, pp. 525–547.

Strömberg, David (2008). "How the Electoral College influences campaigns and policy: the probability of being Florida". In: *American Economic Review* 98.3, pp. 769–807.

A Appendix

A.1 Solving for Equilibriums in Calculus of Voting Models

For the Calculus of Voting Model with Hetergenous Costs, I use the Neweton-Raphson method to solve for the equilibrium turnout rates. I rewrite the turnout rates given in equations **??** and **??** as:

$$F(\sigma_j^D, \sigma_j^R) = \begin{bmatrix} \sigma_j^D - F(0)\frac{e^{(p(\sigma_{Dj}, \sigma_{Rj}, n_j)\mu_j - c_j)}}{1+e^{(p(\sigma_{Dj}, \sigma_{Rj}, n_j)\mu_j - c_j)}} \\ \sigma_j^R - (1-F(0))\frac{e^{(p(\sigma_{Dj}, \sigma_{Rj}, n_j)\mu_j - c_j)}}{1+e^{(p(\sigma_{Dj}, \sigma_{Rj}, n_j)\mu_j - c_j)}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Let σ_{jt}^D and σ_{jt}^R be the guess for the turnout rates at time t. The next guess is given by:

$$F(\sigma_{jt+1}^D, \sigma_{jt+1}^R) = F(\sigma_{jt}^D, \sigma_{jt}^R) - J_F(\sigma_{jt}^D, \sigma_{jt}^R)^{-1} F(\sigma_{jt}^D, \sigma_{jt}^R)$$

Where $J_F(\sigma_{jt}^D, \sigma_{jt}^R)$ is the Jacobian of F evaluated at σ_{jt}^D and σ_{jt}^R. I repeat this procedure until the difference between the turnout rates at time t and $t+1$ is less than some small ϵ (in practice, I use $1e-6$). This gives us equilibrium turnout rates where the turnout rates of both sides are best responses to the turnout rate of the other side.

I use a similar approach to solve for the equilibrium turnout rates in the Calculus of Voting Model with Hetergenous Preferences.

A.2 Equilibrium effort in Leader Persuasion Model

The expected number of seats for party D is given by:

$$E(s^D) = \int \sum_j \Phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{\sqrt{2}\sigma_\eta})h(\delta)d\delta$$

and the expected number of seats for part R is:

$$E(s^R) = \int \sum_j \left(1 - \Phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{\sqrt{2}\sigma_\eta})\right) h(\delta)d\delta$$

This gives us the following first order conditions:

$$\frac{\partial E(s^D)}{\partial e_j^D} = a'(e_j^D) \int \left(\phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{2\sigma_\eta})\right) h(\delta)d\delta = \lambda$$

$$\frac{\partial E(s^R)}{\partial e_j^R} = a'(e_j^R) \int \left(\phi(\frac{(a(e_j^D) - a(e_j^R)) - 2u_j - 2\delta}{2\sigma_\eta})\right) h(\delta)d\delta = \lambda$$

Let $Q_j \equiv \int \left(\phi(\frac{(a(e_j^D)-a(e_j^R))-2u_j-2\delta}{2\sigma_\eta}) \right) h(\delta)d\delta$ and $Q \equiv \sum Q_j$.
Then,

$$a'(e_j^D)Q_j = \frac{\theta}{e_j^D}Q_j = \lambda$$

Where we use that $a(e_j^D) = \theta \log(e_j^D)$. Therefore,

$$\frac{\theta}{e_j^D}Q_j = \frac{\theta}{e_{j'}^D}Q_{j'} \quad \forall j, j'$$

$$e_j^D = e_{j'}^D \frac{Q_j}{Q_{j'}} \quad \forall j, j'$$

Summing over all districts:

$$\sum_j e_j^D = \sum_j e_{j'}^D \frac{Q_j}{Q_{j'}}$$

$$\sum_j e_j^D = \frac{e_{j'}^D}{Q_{j'}} \sum_j Q_j$$

$$E^D = \frac{e_j^D}{Q_j}Q$$

$$e_j^D = Q_j \frac{E^D}{Q}$$

Likewise,

$$e_j^R = Q_j \frac{E^R}{Q}$$

Thus, using the fact that $E^D = E^R$, $e_j^D = e_j^R$ for all j. Given that $e_j^D = e_j^R$, $\Delta u_j = 0$ for all j. Therefore,

$$Q_j = \int \left(\phi(\frac{2\delta + 2\mu_j}{2\sigma_\eta}) \right) h(\delta)d\delta \tag{23}$$

This gives us us a unique solution for e_j^D and e_j^R:

$$e_j^D = e_j^R = Q_j \frac{E}{Q}$$

A.3 Equilibrium effort in Leader Mobilization Model

The following closely follows the derivation in Bouton et al. (2023)

Let $\hat{\delta}$ be the value of δ such that, given e_j^R and e_j^D, the candidates expect equal vote shares. That is, $\hat{\delta}$ solves:

$$e_j^R F(\mu_j - \hat{\delta}) = e_j^D (1 - F(\mu_j - \hat{\delta})) \tag{24}$$

In other words, for any $\delta < \hat{\delta}$, R expects to win the district. Let H be the cdf of this distribution.

The objective function for the R candidate is:

$$\max_{e_j^R} H(\hat{\delta}(e_j^R, e_j^D)) - c(e_j^R)$$

Where $c(e_j^R)$ is the cost of mobilizing voters. This is assumed to be increasing in e_j^R. Similarly, the objective function for the D candidate is:

$$\max_{e_j^D} 1 - H(\hat{\delta}(e_j^R, e_j^D)) - c(e_j^D)$$

R chooses e_j^R to maximize the probability of winning:

$$\max_{e_j^R} H(\hat{\delta}) - c(e_j^R)$$

FOC is given by:

$$h(\hat{\delta}) \frac{d\hat{\delta}}{de_j^R} - c'(e_j^R) = 0$$

Likewise, D chooses e_j^D to maximize the probability of winning:

$$\max_{e_j^D} 1 - H(\hat{\delta}) - c(e_j^D)$$

FOC is given by:

$$-h(\hat{\delta}) \frac{d\hat{\delta}}{de_j^D} - c'(e_j^D) = 0$$

Make the following assumptions:

$$c(e_j^s) = e_j^{s\alpha}/(\alpha)$$

Then, the FOCs become:

For R:

$$h(\hat{\delta}) \frac{F(\mu_j - \hat{\delta})}{(e_j^D + e_j^R) f(\mu_j - \hat{\delta})} - e_j^{R^{\alpha-1}} = 0$$

For D:

$$h(\hat{\delta})\frac{1 - F(\mu_j - \hat{\delta})}{(e_j^D + e_j^R)f(\mu_j - \hat{\delta})} - e_j^{D^{\alpha-1}} = 0$$

The ratio of the two FOCs is given by:

$$\frac{e_j^R}{e_j^D} = (\frac{F(\mu_j - \hat{\delta})}{1 - F(\mu_j - \hat{\delta})})^{\frac{1}{\alpha-1}}$$

We can use this with 24 to solve for $\hat{\delta}$. Doing so shows that:

$$\hat{\delta} = \mu_j - F^{-1}(\frac{1}{2}) = \mu_j$$

Plugging this into the FOC gives:

$$e_j^R = \left(h(\mu_j)\frac{F(0)}{(e_j^D + e_j^R)f(0)}\right)^{\frac{1}{\alpha-1}}$$

$$e_j^D = \left(h(\mu_j)\frac{1 - F(0)}{(e_j^D + e_j^R)f(0)}\right)^{\frac{1}{\alpha-1}}$$

This gives us that:

$$e_j^R = e_j^D \equiv e_j$$

A.4 Equilibrium strategies in Aggregate Group Utility Model

Given the independence across districts, I will drop the subscript j for the remainder of this section. Side D wins if $\mu\frac{\lambda_D}{cost} > (1-\mu)\frac{\lambda_R}{cost}$. Simplifying, we get that side D wins if

$$\mu > \frac{\lambda_R}{\lambda_D + \lambda_R}$$

Alternatively, side R wins if:

$$\mu < \frac{\lambda_R}{\lambda_D + \lambda_R}$$

Side D's objective function is to maximize overall expected utility minus cost. The cost must be paid whether or not side D wins the election. Therefore, side D's cost is simply:

$$cost_D = E[\mu]\int_0^{\lambda_D} cf(c) = \frac{\alpha_1}{\alpha_1 + \alpha_2}\frac{\lambda_D^2}{2c}$$

However, they only receive the utility from winning if their side wins, therefore their expected utility is given by:

$$E[u_D] = u_D \int_{\frac{\lambda_R}{\lambda_D+\lambda_R}}^{1} \mu h(\mu) d\mu \tag{25}$$

Together, this gives us the following objective function:

$$\max_{\lambda_D} u_D \int_{\frac{\lambda_R}{\lambda_D+\lambda_R}}^{1} \mu h(\mu) d\mu - \frac{\alpha_1}{\alpha_1+\alpha_2}\frac{\lambda_D^2}{2c}$$

Likewise, side R's objective function is:

$$\max_{\lambda_R} u_R \int_{0}^{\frac{\lambda_R}{\lambda_D+\lambda_R}} (1-\mu) h(\mu) d\mu - \frac{\alpha_2}{\alpha_1+\alpha_2}\frac{\lambda_R^2}{2c}$$

The first order conditions are:

$$u_D h(\frac{\lambda_R}{\lambda_R+\lambda_D})\frac{\lambda_R^2}{(\lambda_R+\lambda_D)^3} - \frac{\alpha_1}{\alpha_1+\alpha_2}\frac{\lambda_D}{c} = 0 \tag{26}$$

$$u_R h(\frac{\lambda_R}{\lambda_R+\lambda_D})\frac{\lambda_D^2}{(\lambda_R+\lambda_D)^3} - \frac{\alpha_2}{\alpha_1+\alpha_2}\frac{\lambda_R}{c} = 0 \tag{27}$$

We can rearrange 26 and 27 to get common terms between the two equations. Setting them equal, we then get:

$$h(\frac{\lambda_R}{\lambda_R+\lambda_D})\frac{c(\alpha_1+\alpha_2)}{(\lambda_R+\lambda_D)^3} = \frac{\alpha_1\lambda_D}{u_D\lambda_R^2} = \frac{\alpha_2\lambda_R}{u_R\lambda_D^2} \tag{28}$$

This gives us a relationship between λ_D and λ_R:

$$\frac{\alpha_1}{u_D}\lambda_D^3 = \frac{\alpha_2}{u_R}\lambda_R^3$$

Which can be used to calculate $\frac{\lambda_R}{\lambda_R+\lambda_D}$:

$$\frac{\lambda_R}{\lambda_R+\lambda_D} = \frac{\alpha_1^{1/3}u_R^{1/3}}{\alpha_1^{1/3}u_R^{1/3}+\alpha_2^{1/3}u_D^{1/3}} \equiv \kappa$$

Further, let

Using 26, 27, and 28, along with the above definition, we get:

$$\lambda_D^2 = cu_D h(\kappa)\kappa^2 \frac{(\alpha_1+\alpha_2)}{\alpha_1}\frac{\alpha_2^{1/3}u_D^{1/3}}{\alpha_1^{1/3}u_R^{1/3}+\alpha_2^{1/3}u_D^{1/3}} \tag{29}$$

$$\lambda_R^2 = cu_R h(\kappa)(1-\kappa)^2 \frac{(\alpha_1+\alpha_2)}{\alpha_2}\frac{\alpha_1^{1/3}u_R^{1/3}}{\alpha_1^{1/3}u_R^{1/3}+\alpha_2^{1/3}u_D^{1/3}} \tag{30}$$

This simplifies to:

$$\lambda_D^2 = cu_D h(\kappa)\kappa^2(1-\kappa)\frac{(\alpha_1+\alpha_2)}{\alpha_1} \tag{31}$$

$$\lambda_R^2 = cu_R h(\kappa)\kappa(1-\kappa)^2 \frac{(\alpha_1+\alpha_2)}{\alpha_2} \tag{32}$$

By assumption, $u_D = \bar{u_D} exp(\epsilon)$ and $u_R = \bar{u_R} exp(\epsilon)$.
We can rewrite 31 and 32 with ϵ:

$$\lambda_D^2 = c\bar{u_D} exp(\epsilon) h(\kappa)\kappa^2(1-\kappa) \frac{(\alpha_1+\alpha_2)}{\alpha_1} \tag{33}$$

$$\lambda_R^2 = c\bar{u_R} exp(\epsilon) h(\kappa)\kappa(1-\kappa)^2 \frac{(\alpha_1+\alpha_2)}{\alpha_2} \tag{34}$$

Additionally, the Democratic vote share is given by $v_D = \frac{\lambda_D \mu}{c}$ and the Republican vote share is given by $v_R = \frac{\lambda_R(1-\mu)}{c}$.

We can use these definitions, along with 29 and 30 to create a system of two equations and two unknowns, allowing us to solve for μ and ϵ given the parameters:

$$\frac{c^2 v_D^2}{\mu^2} = c\bar{u_D} exp(\epsilon) h(\kappa)\kappa^2(1-\kappa) \frac{(\alpha_1+\alpha_2)}{\alpha_1}$$

$$\frac{c^2 v_R^2}{(1-\mu)^2} = c\bar{u_R} exp(\epsilon) h(\kappa)\kappa(1-\kappa)^2 \frac{(\alpha_1+\alpha_2)}{\alpha_2}$$

Solving for common terms on the right hand side between the two equations, we get:

$$\frac{v_D^2 \alpha_1}{\mu^2 \bar{u_D}\kappa} = \frac{v_R^2 \alpha_2}{(1-\mu)^2 \bar{u_R}(1-\kappa)}$$

$$\frac{(1-\mu)^2}{\mu^2} = \frac{\alpha_2 v_R^2 \bar{u_D}\kappa}{\alpha_1 v_D^2 \bar{u_R}(1-\kappa)}$$

Therefore,

$$\mu = \left(\frac{(\alpha_1 v_D^2 \bar{u_R}(1-\kappa))^{1/2}}{(\alpha_1 v_D^2 \bar{u_R}(1-\kappa))^{1/2} + (\alpha_2 v_R^2 \bar{u_D}\kappa)^{1/2}} \right)$$

We can use this to solve for ϵ.

$$exp(\epsilon) = \frac{c v_D^2}{\bar{u_D} h(\kappa)\kappa^2(1-\kappa)} \frac{\alpha_1}{\alpha_1+\alpha_2} \left(\frac{(\alpha_1 v_D^2 \bar{u_R}(1-\kappa))^{1/2} + (\alpha_2 v_R^2 \bar{u_D}\kappa)^{1/2}}{(\alpha_1 v_D^2 \bar{u_R}(1-\kappa))^{1/2}} \right)^2$$

The fraction of votes observed for D and R is therefore a transformation of the random variables ϵ and μ. Note that the other models in this paper are in terms of total votes for each party, rather than the fraction of votes. I therefore replace v_D and v_R with $\frac{\text{Total Votes D}}{Num.EligibleVotes}$ and $\frac{\text{Total Votes R}}{Num.EligibleVotes}$ respectively.

This gives us the log-likelihood function for each district:

$$LL(\beta)_j = Pr(\mu_j; \alpha_1, \beta) Pr(\epsilon_j) |J|$$

Where μ_j is drawn from a Beta distribution with parameters α_1 and α_2, and ϵ_j is drawn from a Normal distribution with mean 0 and variance 1, and $|J|$ is the determinant of the Jacobian of the transformation from ϵ and μ to v_D and v_R, i.e.

$$J = \begin{bmatrix} \frac{\partial \mu}{\partial \tilde{v_D}} & \frac{\partial \epsilon}{\partial \tilde{v_D}} \\ \frac{\partial \mu}{\partial \tilde{v_R}} & \frac{\partial \epsilon}{\partial \tilde{v_R}} \end{bmatrix}$$

Where $\tilde{v_D}$ and $\tilde{v_R}$ are number of votes for D and R respectively divided by the number of eligible voters. Together, this gives us the log-likelihood function for the entire dataset:

$$LL(\beta) = \sum_j LL(\beta)_j$$

A.5 Solving for the gradient in the Calculus of Voting Models

The likelihood function for the Calculus of Voting Model with Hetergenous Costs is given by:

$$L = \prod_j Pr(v_{jD}, v_{jR}, v_{jA} | \beta, n_j) \tag{35}$$

Though note the chain of causality runs as follows:

$$\beta \to \text{utility, cost}, \mu, \alpha^1, \alpha^2, \to \text{turnout rates} \to \text{vote shares}$$

I use autodifferentiation from the JAX library to calculate the jacobian of the parameters with respect to β, and the jacobian of the vote shares with respect to the turnout rates. However, because I solve for the equilibrium turnout rates using a root finding algorithm, I cannot use autodifferentiation to calculate how the parameters affect vote shares.

Instead, I rewrite the equilibrium turnout rates as:

$$F(params, \sigma_j) = \begin{bmatrix} \sigma_j^D - F(0) \frac{e^{(p(\sigma_j, \alpha^1, \alpha^2)\mu_j - c_j)}}{1 + e^{(p(\sigma_j, \alpha^1, \alpha^2)\mu_j - c_j)}} \\ \sigma_j^R - (1 - F(0)) \frac{e^{(p(\sigma_j, \alpha^1, \alpha^2)\mu_j - c_j)}}{1 + e^{(p(\sigma_j, \alpha^1, \alpha^2)\mu_j - c_j)}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Where σ_j is the vector of turnout rates and *params* is the vector of parameters (utility, cost, μ, α^1, α^2).

By the implicit function theorem, the jacobian of the vote shares with respect to the parameters is given by:

$$\frac{\partial \sigma_j^*}{\partial \text{params}} = -[J_{F,\sigma_j}(params, \sigma_j^*)]^{-1} J_{F,\text{params}}(params, \sigma_j^*)$$

Where σ_j^* is the equilibrium turnout rates given the parameters, J_{F,σ_j} is the jacobian of F with respect to σ_j, and $J_{F,\text{params}}$ is the jacobian of F with respect to the parameters.

A.6 Estimaing Leader Persuasion Model

To esimate the Leader Persuasion Model, first take the log difference of equations 14 and 15:

$$s_j \equiv log(\sigma_j^D) - log(\sigma_j^R) = u(e_j^{*D}) - u(e_j^{*R}) - 2\mu_j - 2\delta - \eta_j$$

In equilibrium $u(e_j^{*D}) - u(e_j^{*R}) = 0$. Therefore,

$$s_j = -2\mu_j - 2\delta - \eta_j$$

Note that $\delta \sim N(0, \sigma_{\delta j})$, and $\eta_j \sim N(0, \sigma_{\eta j})$. We can estimate this equation using random effects techniques, recovering the coefficients on μ_j, in addition to δ, η_j, $\sigma_{\delta j}$, and $\sigma_{\eta j}$.

However, we still need to estimate c_j and θ. Given η_j^R and η_j^D, along with the estimates for δ and μ_j, for a given β_c and θ, we can calculate the turnout rates given by equations 14, 15, and 16. The number of votes in each district is then a draw from a multinomial distribution:

$$v_{jD}, v_{jR}, v_{jA} \sim Multinomial(\sigma_j^D, \sigma_j^R, n_j)$$

However, given that we only observe $\eta_j = \eta_j^D - \eta_j^R$, we must integrate out one of the η_j^s terms. Rewriting the turnout rates in terms of η_j^R:

$$\sigma_j^D = \frac{e^{u(e_j^{*D}) - c_j - (\eta_j + \eta_j^R) - \mu_j - \delta}}{1 + e^{u(e_j^{*D}) - c_j - (\eta_j + \eta_j^R) - \mu_j - \delta} + e^{(u(e_j^{*R}) - c_j - \eta_j^R + \mu_j + \delta)}}$$

$$\sigma_j^R = \frac{e^{(u(e_j^{*R}) - c_j - \eta_j^R + \mu_j + \delta}}{1 + e^{u(e_j^{*D}) - c_j - (\eta_j + \eta_j^R) - \mu_j - \delta} + e^{(u(e_j^{*R}) - c_j - \eta_j^R + \mu_j + \delta)}}$$

the likelihood becomes:

$$L = \prod_j \int_{\eta_{Rj}} Pr(v_{jD}, v_{jR} | \beta_c, \theta, \eta_{Rj}) f(\eta_{Rj}) d\eta_{Rj}$$

A.7 Data Selection with Random Forest and Gradient Boosting

In addition to a LASSO regression, I use a random forest and gradient boosting model to select the most important variables in predicting Democratic and Republican turnout rates across districts. As in the main text, I do this for Republican turnout and Democratic turnout. Random Forest and Gradient Boosting both use decision trees to make predictions, Gradient boosting builds

Table 14: Importance of Features for Democratic Turnout

Feature	Gradient Boosting	Random Forest
Frac. Coll+	0.039	0.041
Preregistration	0.025	
Registration Deadlines	0.045	0.035
Registration Restrictions	0.116	0.087
Incumbent	0.437	0.283
Frac. No HS	0.021	0.036
Frac. 18-29		0.036
Frac. 65+	0.016	0.039
Frac. Asian	0.019	0.034
Frac. Black	0.061	0.060
Frac. Hispanic	0.117	0.106

trees sequentially, with each tree predicting the residuals of the previous tree. Random Forest builds trees in parallel, with each tree being independent of the others, and the final prediction being the average of the predictions of all trees. In the context of this paper, the outcome variable is the turnout rate of a party in a district. Gradient Boosting differs from Random Forest in that it builds trees sequentially, while Random Forest builds trees in parallel. Tables 14 and 15 show the top ten features from both models in predicting Democratic and Republican turnout rates, respectively. The predictions from these models largely align with the results from the LASSO regression.

A.8 Results from Simulations

Tables show mean difference between estimated and true values, along with standard deviation of the difference.

Table 15: Importance of Features for Republican Turnout

Feature	Gradient Boosting	Random Forest
Poll Hours	0.014	
Preregistration	0.033	0.027
Registration Deadlines	0.012	
Registration Restrictions	0.016	
Incumbent	0.565	0.462
Median Income Ratio	0.031	0.043
Frac. No HS	0.020	0.023
Frac. Some Coll	0.023	0.025
Frac. 18-29	0.043	0.038
Frac. 30-49		0.025
Frac. 65+		0.024
Frac. Black		0.022
Frac. White	0.174	0.168

Table 16: Calculus of Voting Model with Hetergenous Costs

Variable	Mean	Std
variable 0	0.000	0.001
variable 1	-0.000	0.001
variable 2	0.000	0.004
variable 3	-0.000	0.002
variable 4	0.000	0.003
variable 5	0.000	0.002
variable 6	-0.000	0.001
variable 7	0.002	0.013
variable 8	0.001	0.005
variable 9	-0.003	0.026
variable 10	0.001	0.006
variable 11	0.001	0.005
variable 12	0.002	0.010
variable 13	0.081	0.603
variable 14	0.001	0.012
variable 15	-0.003	0.018
variable 16	0.000	0.003
variable 17	-0.021	0.295
variable 18	0.017	0.133

Note: This table shows the mean difference between estimated and true values, along with the standard deviation of the difference. A total of 40 simulations were run.

Table 17: Calculus of Voting Model with Hetergenous Preferences

Variable	Mean	Std
variable 0	0.008	0.055
variable 1	0.007	0.113
variable 2	0.008	0.120
variable 3	-0.026	0.283
variable 4	-0.017	0.182
variable 5	-0.003	0.161
variable 6	-0.005	0.129
variable 7	0.004	0.092
variable 8	0.005	0.145
variable 9	-0.000	0.105
variable 10	0.012	0.115
variable 11	0.001	0.097
variable 12	0.002	0.054
variable 13	-0.005	0.117
variable 14	0.006	0.102
variable 15	0.006	0.070
variable 16	-0.001	0.093
variable 17	-0.001	0.120

Note: This table shows the mean difference between estimated and true values, along with the standard deviation of the difference. A total of 400 simulations were run.

Table 18: Leader Persuasion Model

Variable	Mean	Std
variable 0	0.053	0.264
variable 1	-0.007	0.028
variable 2	-0.014	0.029
variable 3	-0.005	0.026
variable 4	0.029	0.055
variable 5	0.006	0.067
variable 6	0.009	0.043
variable 7	0.017	0.032
variable 8	0.001	0.038
variable 9	0.011	0.031
variable 10	0.000	0.000
variable 11	-0.007	0.051
variable 12	0.000	0.001
variable 13	-0.002	0.005
variable 14	-0.000	0.001
variable 15	-0.002	0.002
variable 16	0.060	0.261
variable 17	0.000	0.002

Note: This table shows the mean difference between estimated and true values, along with the standard deviation of the difference. A total of 1400 simulations were run.

Table 19: Leader Mobilization Model

Variable	Mean	Std
variable 0	0.018	0.196
variable 1	0.019	0.180
variable 2	-0.001	0.056
variable 3	0.024	0.194
variable 4	0.008	0.068
variable 5	-0.044	0.503
variable 6	0.013	0.134
variable 7	0.001	0.006
variable 8	-0.001	0.018
variable 9	-0.001	0.018
variable 10	0.002	0.015
variable 11	0.000	0.008
variable 12	-0.000	0.013
variable 13	0.001	0.010
variable 14	0.002	0.016
variable 15	0.001	0.009
variable 16	0.001	0.008
variable 17	0.129	1.199

Note: This table shows the mean difference between estimated and true values, along with the standard deviation of the difference. A total of 100 simulations were run. Note high standard deviation for variable 17 was driven by 2 trials that failed to converge.