In Pursuit of Direct Democracy

Analyzing Predictors of U.S. State Ballot Measure Support and Roll-off

Andrew Kim

I. Introduction

Although often perceived as less salient than major candidate races at the top of electoral ballots each year, ballot measures—policy issues appearing on an electoral ballot for voters to decide on—have long been presented to U.S. voters as a mechanism of direct democratic policymaking. As of the end of 2024, the National Conference of State Legislatures (2024) reports 8,612 ballot measures that have been proposed or will soon be proposed to U.S. residents since 1902, and in 2024 alone, 159 statewide ballot measures across 41 states have been certified for the ballot (Ballotpedia 2024a). Ballot measures can be either referred to public vote by the legislature or petitioned for placement on the ballot by citizens, known as an initiative. In either case, measures can range in subject from criminal justice law to tax collection to abortion rights, implicating a vast array of constituent-relevant policies. Yet, despite ballot measures' far-reaching role in the law, they tend to receive relatively little attention from the American public and, in turn, lower levels of informed public participation.

In this paper, I explore the role of factors from ballot text to campaign finance to voter access as they impact ballot measures' electoral support and roll-off (the share of people that vote in the top races on a ballot but not in lower ones). First, I conduct a literature review on ballot measures and the relationship between characteristics of voting experience and ballot measure outcomes. I conclude this section with a series of hypotheses to guide my analysis on ballot measure support and roll-off. Next, I define my methods for collecting and analyzing electoral and demographic data related to all state ballot measures from 2004 to 2024. Using a gradient-boosted tree model, I identify high-importance features in the prediction of support and roll-off, then incorporate these features into a series of linear regression models to understand the significance and magnitude of influence that various variables have on the outcome variables.

After presenting the results of these analyses, I discuss the implications of my findings and offer several policy recommendations to enhance ballot text accessibility and regulate issue campaign finance in response to threats to ballot measures' efficacy as a direct democratic institution.

II. Background Research

A. Ballot Measures in U.S. Elections

For the directness of their impact on policy outcomes, ballot measures do not enjoy the same salience as federal and statewide candidate races. One Arkansas study held that only about one in four voters displayed any familiarity with state ballot measures before entering the ballot box, even about sensitive policy issues such as marijuana, alcohol, and wages (Barth, Burnett, and Parry 2019). Further, the majority of voters did not expect to see any ballot measures on their ballot in the first place. These findings illustrate that the ballot itself is the first place the vast majority of voters will reckon with consequential policy changes put before them, in turn emphasizing the importance of ballot measures' textual presentation on voters' ultimate decision to support the measure, oppose it, or abstain from voting altogether.

When making the decision to vote in an election, individuals are commonly believed to conduct a personal calculus of voting—weighing the likely benefits of casting their ballot against the costs required to do so (Riker and Ordeshook 1968). Benefits to voters may include electoral competitiveness (the extent to which they believe their vote will change the outcome), relevance of consequent policy change, and a sense of civic fulfillment. On the other hand, a time-consuming registration process, polling places with limited locations and opening hours, and a complex ballot may dissuade voters from making the effort to vote. As it relates to ballot measures, high costs in difficult-to-understand text with weak benefits in perceived policy relevance can lead voters to default to voting against or abstaining from measures they are uncertain about, especially when they lack indicators such as party affiliation that voters often use as a heuristic to satisfice in candidate races. It is a product of this calculus of voting that it is well established that items at the end of a ballot—as ballot measures nearly universally are—are

more likely to be skipped or voted against (a risk-mitigating choice in favor of the status quo) out of voter fatigue (Augenblick and Nicholson 2016).

B. Ballot Text

On average, the 41 statewide ballot measures presented to U.S. voters in 2023 required 19 years of education (a third-year graduate school reading level) to comprehend according to the Flesch reading ease and Flesch-Kincaid grade level formulas for textual complexity (Ballotpedia 2023a). That same year, the United States Census Bureau (2024) reported that only 14.3 percent of U.S. residents ages 25 and older had completed advanced education (and only 36.2 percent had obtained a Bachelor's degree or higher). The literature broadly concludes that long, complex measure text is an obstacle to voting. Redish et al. (2010) found that voters completing a ballot with instructions in plain language (short sentences composed of simple, everyday words) are not only capable of voting according to their intentions at higher rates, but notice and prefer such a ballot. Ulbig and Reilly (2021) refine this discovery by adding that English speakers of all proficiency levels presented with a textually complex ballot measure struggle to vote in a manner consistent with their preferences compared to a plain language measure. Likewise, Poole (2019) concludes that voters change their vote choice depending on whether their ballot contains traditional or simplified language, and largely prefer the ballot with simplified wording. The bottom line is clear: simple, concise ballot language makes tangible improvements to the affective voting experience and voting according to one's true preferences.

But the Poole study underscores a different concern, that ballot text can alter voting behavior, even independent from misinterpretation. Other studies concur: poorer readability increases abstentions and votes against ballot measures (Reilly and Richey 2009; Shulman et al.

2022), similar to the aforementioned voter fatigue effect observed in down-ballot races. Dyck and Pearson-Merkowitz (2021) conclude that voters are more easily induced to vote against ballot measures that are technical and require high levels of political sophistication, education, and/or effort to understand (e.g., financial regulation). Even seemingly inconsequential alterations can have outsize effects on voter support; Rossier (2021) determined a change in ballot question framing from an absolute to percentage-based tax increase resulted in a drastic drop in support. Similarly, Poole (2019) presented voters with identical ballot initiatives with either the option to vote "for" vs. "against" or "yes" and "no" and found that the latter options increased support by 20 percent—the difference between the ballot failing and passing.

Regretfully, the latent effects of ballot text have been reported as a vote-swaying tool by partisan legislators across the ideological spectrum. The right-leaning Cato Institute calls attention to deceptive ballot language in California in 2020, which omitted mention of a $10 billion annual commercial property tax increase from the title of a measure seeking to raise education funding (Joffe 2024). The progressive Brennan Center for Justice raises a similar example of a 2024 Ohioan anti-gerrymandering initiative to devolve redistricting powers from politicians to an independent citizen commission (Rudensky 2024). In September 2024, the coalition behind the initiative sued the Ohio Ballot Board for approving language describing the new commission as itself "required to gerrymander" and insulated from public accountability (Smyth 2024). Given the latter initiative had been signed by over 535,000 voters to circumvent legislative procedure and reach the ballot, deceptive ballot measure text not only confuses voters at the ballot box but can stifle mechanisms of citizen-initiated policy change.

The academic consensus is that complex ballot measure text effectively disenfranchises a majority of American voters from casting a confident and informed vote. Minor changes to ballot

language can also make consequential impacts in electoral outcomes. Ultimately, when ballot measures contain language that is inaccessible—or worse, deliberately misleading by the politicians and lobbyists who write them—they are rendered a disingenuous means of conveying the electorate's will in a direct democracy (Hvasta 2020).

C. Campaign Finance

One other ballot text survey experiment by Burnett and Kogan (2015) corroborates that adjustments to ballot measure wording (i.e., "limiting same-sex marriage" vs. "eliminating the right to same-sex marriage") significantly changes voters' support for the question, although exposure to political campaign information can mitigate this effect. Indeed, issue campaigns can act as a boon for voter education and ballot measure salience.

Youth civic engagement research organization CIRCLE at Tufts University (2014) observed that states with one or more controversial ballot measures alongside a competitive statewide election (for Governor or Senate) in 2014 experienced a significant increase by 5 points in youth turnout since 2010 compared to states with a competitive statewide election only (decrease of 1 point). Researchers pointed to targeted outreach and increased media coverage on ballot measure issues as major contributors to turnout. Relatedly, a more recent study posits that while ballot readability does influence whether a voter abstains from a measure, this relationship is moderated by the amount of issue information and perceived importance a voter possesses going into the ballot box (Milita 2017).

Theoretical literature commends civic education campaigns as tools that help citizens to understand how to connect their everyday interests with specific legislation and actionable voting decisions, thus promoting civic participation. Beyond turnout, civic education can moderate less

confident voters' instinct to abstain or vote for the status quo under uncertainty by arming them with ideological and issue-specific knowledge (Galston 2004). The empirical research appears to agree with political campaigns' efficacy for candidate races, with caveats: in general, campaign spending significantly increases vote share and win probability, especially in competitive races. Yet, this relationship is weak to negligible for incumbent candidates compared to their challenging contenders because of diminishing returns on top of the incumbent's natural electoral advantage (Gerber 2004; Le et al. 2024). When spending is effective in candidate races, Schuster (2020) finds, it is most so on turning out additional voters rather than switching the vote of already likely voters, and low-information voters are particularly likely to respond strongly to campaign spending. In conjunction, these findings lend support to the theory that political campaigns meaningfully contribute to voter education when they lack issue knowledge or indicators for vote satisficing.

Research concentrated on ballot measure issue campaigns (as distinct from candidate-centered) cites a less clear effect. Rogers and Middleton (2014) synthesize conflicting studies that ballot initiative campaign spending does not affect initiative outcomes, that it may increase the number of votes rejecting but not approving initiatives, or that it can influence votes in either direction. Their own field experiment supports the latter hypothesis that ballot measure campaigns can influence both approval and rejection of initiatives by changing vote choice, as opposed to changing turnout or ballot completion. Conversely, Childers and Binder (2015) find that strong issue campaigns and high spending on controversial measures mobilizes additional turnout, while Primo (2013) contends with another survey experiment that voters on the margin receive few informational benefits to knowing which interest groups campaign for or against certain measures. All in all, while most of the literature points to a significant impact of

campaign spending on ballot measure outcomes, it is uncertain whether it is a product of increased vote share in one direction or changing turnout for that measure.

In either case, spending on these issue campaigns is behemoth in the U.S.; OpenSecrets reports that committees supporting or opposing ballot measures in 2024 raised a collective $417 million as of October 28 that year (Serna Jr. 2024). While the majority of these races saw no spending at all, a select number of controversial issues (e.g., abortion, marijuana) drew massive funding—Florida's Measure 3 to legalize the use of marijuana for adults 21 and over garnered over $125 million alone (Serna Jr. 2024). With such a staggering sum of money being apportioned to ballot measure campaigns, the question of its positive effect on direct democracy follows. Interest group spending on ballot measure elections can increase welfare by educating the public on ballot measures, which represent more secure policy changes than voting for a legislative representative who is not guaranteed to advance a voter's policy interests (Stratmann 2010). Yet, other types of spending do not increase welfare—legal scholar Richard Briffault (1996) warns that, in the absence of other cues typical of candidate races such as a transparent nominating process or party affiliation indicators, unrestricted campaign finance poses a greater threat of manipulating ballot measure electoral outcomes and undermining the very process designed to subvert the effects of special interests over the government's policymaking.

D. Partisanship

Individual partisanship has long been known as the most important driver of vote choice in presidential elections throughout the 20th century and into the 21st (Stapleton and Langehennig 2024). Analogously, scholars have suggested that the partisan composition of a state's electorate is associated with support for certain ballot measures (Berman and Yawn 2001).

Explanations range from substantive ideological differences on particular issues, to willingness of different partisans to defer to legislative judgments, to propensity to support measures referred by an ideologically congruent legislature. One study testing all three hypotheses (but notably focusing on Arizona voting patterns from 1912 to 1996) found support for the deference theory—Republicans were more likely to support measures submitted by state legislatures—as well as for partisan cues—Republicans especially supported measures from legislatures controlled by a Republican majority (Berman and Yawn 2001).

While the direction of partisan effects on general propensity to vote on and support ballot measures may have changed in the years since the Arizona study due to partisan realignment, modern legislative behavior offers some insights. In the last few years, Republicans in state legislatures across the country have led efforts to raise the requirements for citizen-initiated measures to reach the ballot and increase the approval percentage for passage (Brower 2023; Carter, Clapman, and Comella 2024). While this might not speak to inherent partisan likelihood of supporting, opposing, or abstaining from a ballot measure, the trend is indicative of the general momentum of modern ballot initiatives towards progressive causes such as enshrining abortion access, legalizing marijuana, and expanding Medicaid (Brower 2023).

In spite of expected partisan divide over the content of ballot measures, earlier examples of manipulative ballot text from both sides of the aisle indicates bipartisan support for one ballot reform: ensuring that voters are presented with measure text that preserve their ability to cast a fully educated vote.

E. Hypotheses

Although this research project is exploratory in nature, the above literature offers preliminary theoretical support for several hypotheses as they relate to voter support for ballot measures and ballot roll-off. These general hypotheses will loosely guide my analysis of U.S. state ballot measure outcomes from 2004 to 2024:

1. As ballot measure text gets more complex, support for the measure decreases, and roll-off on the measure increases.
2. As the length of a ballot increases, support for the measure decreases, and roll-off on the measure increases.
3. As the dollars raised supporting a measure increases, support for the measure increases, and roll-off on the measure decreases.
4. As the dollars raised opposing a measure increases, support for the measure decreases, and roll-off on the measure decreases.
5. As turnout in an election decreases, support for ballot measures increases, and roll-off on the measure increases.

III. Methods

I seek to examine the effects of various state-level and ballot measure-related factors on two outcome variables: ballot measure support and roll-off. I quantify support for a ballot measure as the proportion (from zero to one) of votes in favor of the measure's approval out of the total number of votes cast on the measure. Since I lack precise total voter turnout data for every election year for which I have ballot measure data, I operationalize roll-off as the percentage (from zero to one) decrease in total votes on a ballot measure compared to the ballot

measure with the greatest number of votes in that state's election. Under this operationalization, the most voted-on ballot measures in a state's given election have a roll-off of 0.

I programmed a web scraper to gather all data on ballot measures from Ballotpedia's web pages on ballot measures (annual tables and individual web pages) from 2004 to 2024, including each ballot measure's election year, type (e.g., legislatively referred, citizen-initiated, automatically referred), title, number of votes for, number of votes against, approval status, and ballot text. From the vote share metrics, I was then able to pull the support and roll-off variables from each measure. Additionally, I processed the scraped ballot text—which included any combination of the ballot title, ballot summary, and a fiscal impact statement—to determine several additional variables related to accessibility of the ballot text, including the number of words contained, whether the measure is framed as a question, whether the measure is broken into multiple sentences, whether the measure included a fiscal impact statement, and the reading grade level of the full text based on the Flesch-Kincaid grade level readability formula.

The remainder of my data collection was to obtain additional features for my predictive analysis of support and roll-off. Using a similar web scraping method, I retrieved state-level turnout and partisanship data from Ballotpedia for the full time range of the ballot measure data. For turnout data, I employed linear interpolation to approximate turnout in the very small portion of years a state did not report turnout percentages. Party data included each U.S. state's governor's political party, majority party of U.S. Senators, majority party of U.S. representatives, and most recent majority vote for U.S. President. I operationalized each of these four offices' partisan makeup as one quarter of that state's final partisanship score, ranging from -1.0 (most Democratic) to 1.0 (most Republican). Independent officeholders and split offices (e.g., one Democratic and one Republican Senator) count as a point in neither direction. For instance, a

Democratic governor (-.25), split Senators (.00), majority Republican Representatives (.25), and most recent Republican presidential vote majority (.25) denotes a partisanship score of .25.

The last dataset I acquired through web scraping was from OpenSecrets' online ballot measure campaign finance database. For each measure for which finance data was available, I recorded the total dollars raised in support of a measure and the total dollars raised in opposition. Finally, I added these numbers to determine the total number of campaign dollars raised for a given ballot measure.

I found the remaining datasets for my analysis through publicly available online sources. Given the composition of the electorate is a hypothesized predictor of a ballot measure's support and roll-off, I also sought to collect data on the difficulty of registering and turning out to vote as a loose operationalization of the determination and voting familiarity of an electorate. For this purpose, I used the Election Law Journal's Cost of Voting Index (COVI), which assigns COVI scores and rankings to each U.S. state's relative voting accessibility based on their election laws from 1996 to 2024 (Pomante II 2024). Like with turnout, I imputed annual COVI scores between ranking years using linear interpolation. Finally, I download annual demographic data, including state population, gender, median age, education level, median income, and other control measures from the U.S. Census Bureau's American Community Survey. For several appropriate features (but most notably my campaign finance dollar quantities), I used state population data to convert them to per capita (e.g., dollars raised in support of a ballot measure per capita).

After merging all of these datasets into a single dataset, I had 2,184 rows for analysis, each representing one ballot measure. I first obtained descriptive statistics such as ballot measure passage rates and state-by-state differences for different variables. Then, for predicting both the ballot measure support and roll-off outcome variables, I built a gradient-boosted tree to derive

feature importances from my full dataset. To tune each model, I utilized 5-fold cross-validation across two settings for each of five model parameters to derive the parameters with the lowest root mean squared error. Using these parameters, I ran a final gradient-boosted tree and recorded the weights (feature importances) assigned to each feature.

Upon discerning the most important predictive features for support and roll-off, I proceeded to run a multivariate linear regression model for each with the full set of predictors. During this process, I determined that the high number of missing values in my Flesch-Kincaid reading grade level variable (which required ballot text of at least 100 words to calculate) was diminishing the number of observations included in my regressions. In response, I reran a reduced model without the grade level variable, followed by a final, third model using only the subset of high-importance variables as defined by the gradient-boosted tree stage. Ultimately, I evaluated my hypotheses and the effects of my array of predictive independent variables on ballot measure support and roll-off by interpreting the magnitude, direction, and statistical significance of coefficients assigned to each variable at a significance level of $\alpha = .05$.

IV. Results

A. Descriptive Statistics

Of the over 2,000 state ballot measures that had concluded by the time of my exploratory data analysis, approximately 67.22 percent passed. Even-numbered general election years present substantially more ballot measures (100 to 200 in a year) to voters than in odd-numbered off-year elections (under 50 per year) (see Figure A). Further, ballot measures appear to pass at higher rates in the off years, potentially as a product of a more informed and politically invested electorate during lower-salience elections.

The average vote share in favor of a state ballot measure was 57.41 percent—a histogram of ballot measure approval vote percentages displays an approximately normal and spread out distribution, with non-insignificant shares of measures receiving under 40 percent and over 80 percent support (see Figure B).

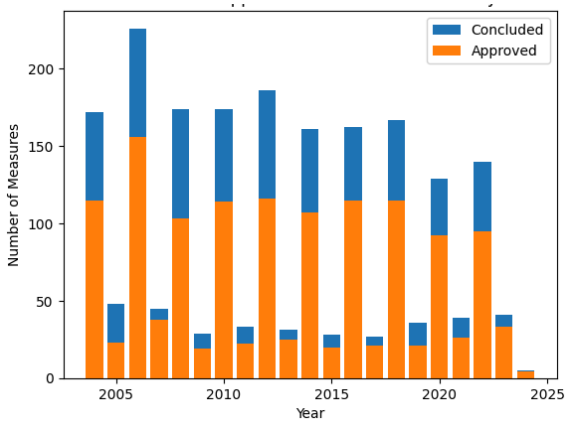


Figure A: Ballot measures approved over total measures, by year

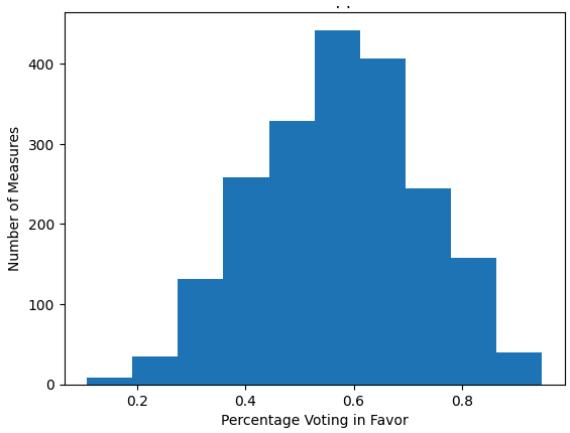


Figure B: Histogram of ballot measure percentage vote in favor

Discounting ballot measures with zero roll-off (recall that roll-off is operationalized as the percentage difference in total votes compared to the ballot measure with the greatest number of votes in the same election), most ballot measures see relatively minimal roll-off of under 5 percent from the top measure (see Figure C). The average roll-off in the ballot measure dataset was 2.77 percent.

Almost all elections proposed multiple measures on the same ballot; the average number of ballot measures per state in a year was 7.30. In line with the voter fatigue literature, this number signifies voters are tasked with deciding on a potentially overwhelming number of unfamiliar policy issues each election, and some states reported presenting as many as 15 to 20 ballot measures to voters in a single year (see Figure D).

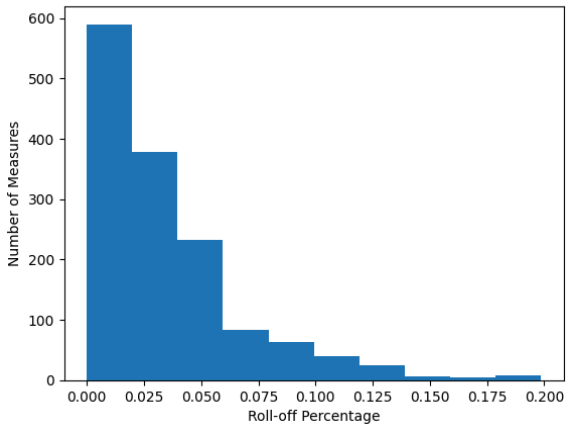


Figure C: Histogram of ballot measure percentage roll-off from top measure on ballot

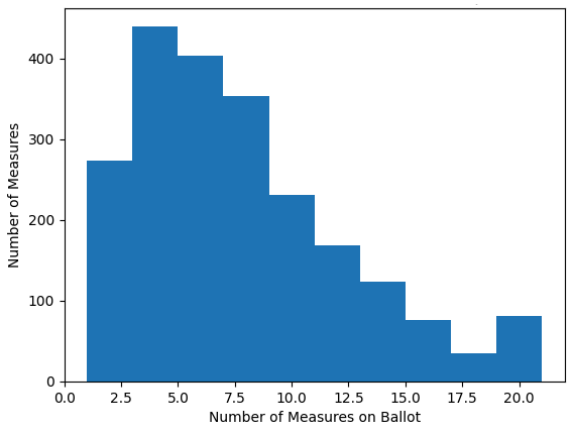


Figure D: Histogram of number of ballot measures in state per year

The most common subjects for ballot measures related to taxes, followed by state governance (e.g., government accountability, legislative mechanisms) (see Figure E). Although some subjects like taxes, crime, and civil rights are generally familiar and accessible to everyday voters for low-cost opinion formulation, many focus on technical policies in governmental function and state finances that can be more difficult for voters to comprehend and develop a

confident opinion on. Even in high-salience subjects, ballot text presentation can influence the extent to which a policy change is digestible to a presumptively unfamiliar voter.

A clear majority of ballot measures—1133 in this dataset—were legislatively referred constitutional amendments. While citizen-initiated state statutes and constitutional amendments were still prevalent (248 and 214 occurrences in the dataset, respectively), ballot measures primarily originate from state legislatures rather than the citizen initiative process (see Figure F).

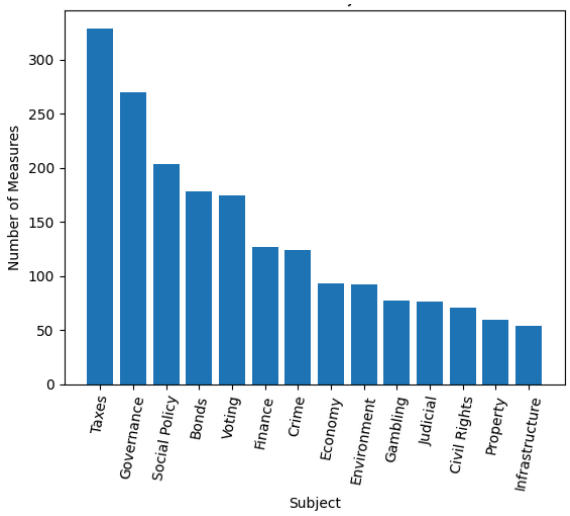


Figure E: Counts of ballot measure subjects

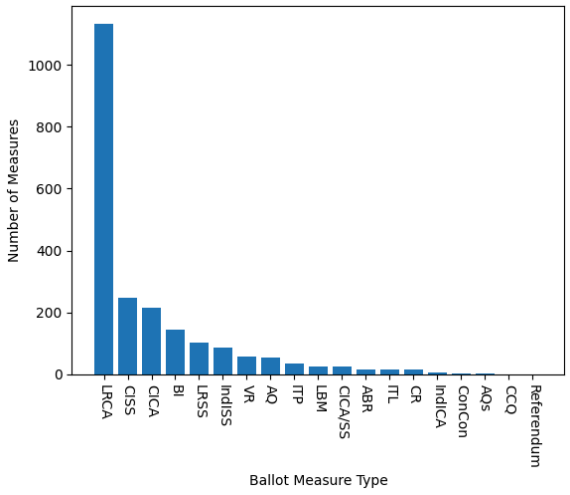


Figure F: Counts of ballot measure types

Not all states employ the ballot measure process equally. Several states such as California, Louisiana, and Colorado have posed over 100 ballot measures in the past 20 years, while others have witnessed under 10 (see Figure G). Disparities in the number of ballot measures in each state reflects not only varying propensities to put policy decisions up to the community at large, but varying citizen familiarity with ballot measures. State residents who almost never encounter non-candidate races on their ballot may be less prepared and practiced in comprehending measure text and casting an informed vote.

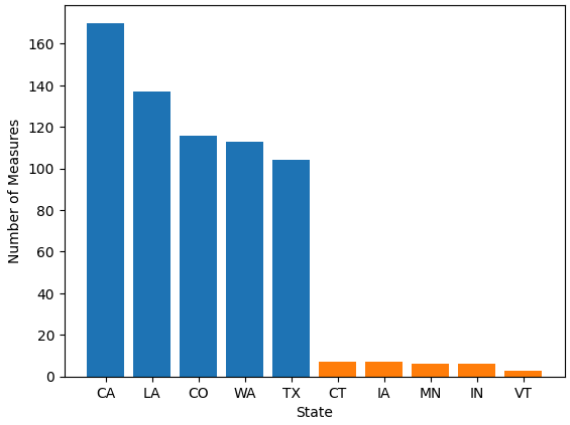


Figure G: States with most and least ballot measures (2004-2024)

Narrowing in on ballot measure text, word lengths also varied greatly—the average word length of a ballot measure was 126.21 words. Between the measure title and summary, a large share of ballot measures displayed concise text under 100 words, but others entered the high hundreds (see Figure H). Although voters may benefit from receiving the requisite amount of information to fully comprehend the measure's meaning, as with the number of races on a ballot, voter fatigue at high word counts can dissuade voters from reading the full text, instead voting against by default or abstaining.

Empirical textual readability raised more cause for concern; the average Flesch-Kincaid reading grade level (for measures over 100 words) was 25.07 grades of education. This metric may be skewed by a significant number of measures with inordinately high Flesch-Kincaid grade levels above 50, but most measures still fell between grade levels of 10 and 20 (see Figure I). A sobering 91.60 percent of measures with analyzable text required at least 12 years (a high school degree) of education to comprehend, and 62.76 percent required at least 16 years (a four-year Bachelor's degree).

Considering ballot text characteristics that theoretically aid voter comprehension, holding word length constant, 36.39 percent of measures were phrased as a question rather than a statement, 62.54 percent of measures were broken down into multiple sentences rather than a single run-on sentence (as required in some jurisdictions), and 7.00 percent of measures included a fiscal statement describing the impact of the measure on state finances (see Figure J).

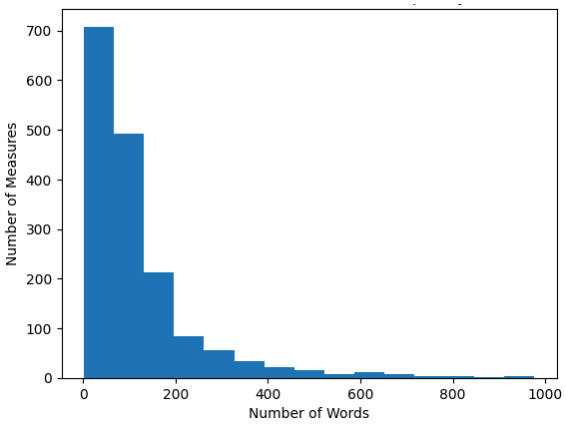


Figure H: Histogram of ballot measure word counts

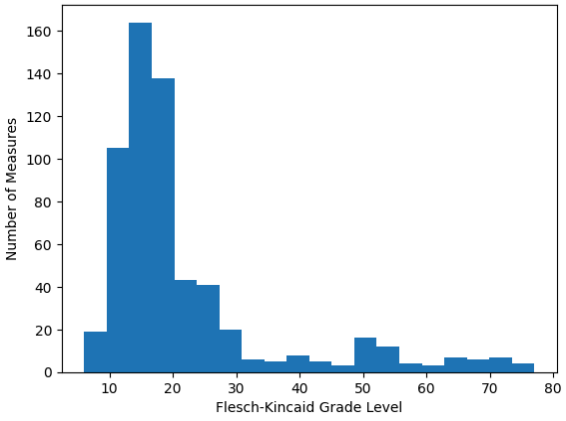


Figure I: Histogram of ballot measure Flesch-Kincaid reading grade levels

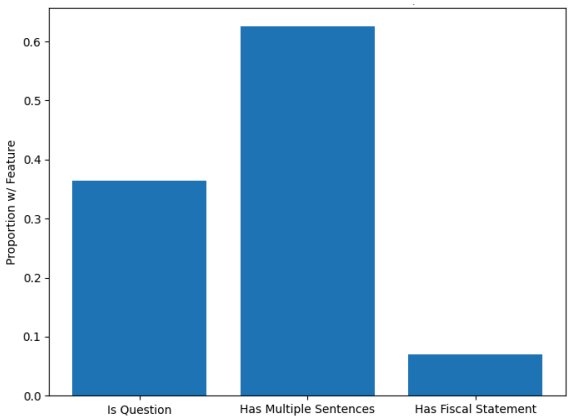


Figure J: Frequencies of ballot measure text characteristics

Finally, interest groups in support of ballot measures collectively outspent groups against them. On average, a ballot measure received $4,831.14 in fundraising in support of it per 10,000 residents in the state and $3,302.31 in opposition to it. But, most funding raised in a given cycle for ballot measure issue campaigns are for a select few number of controversial, high-impact measures such as reproductive rights and marijuana legalization. As seen in histograms of fundraising histograms, a very small minority of ballot measures raise over 25 cents per capita, while some raised over $2.50 per person living in the state to sway public opinion (see Figure K; Figure L).

Out of ballot measures where there was any fundraising at all, support fundraising outpaced opposition fundraising by an average of $2,230.99 per 10,000 residents. Although many races reported relatively low differences in support and opposition campaign funds, dozens over the last 20 years have seen differences upwards of $10,000 per 10,000 residents (see Figure M). Furthermore, 22.16 percent of races had uncontested fundraising by only one side of the ballot measure.

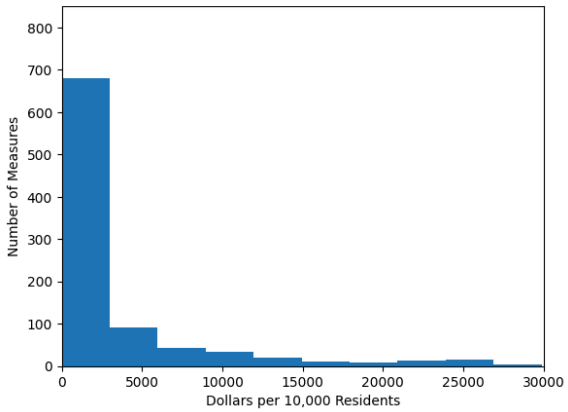


Figure K: Histogram of dollars fundraised in support of ballot measures

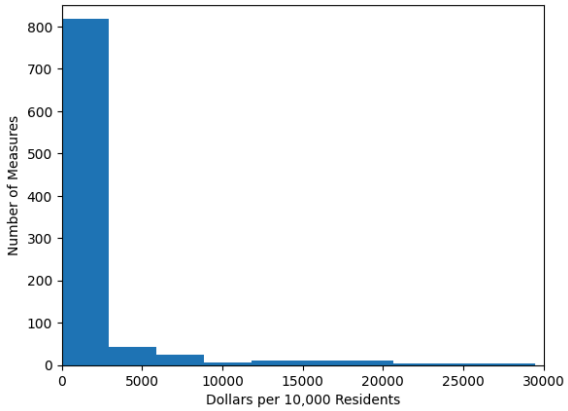


Figure L: Histogram of dollars fundraised in opposition to ballot measures

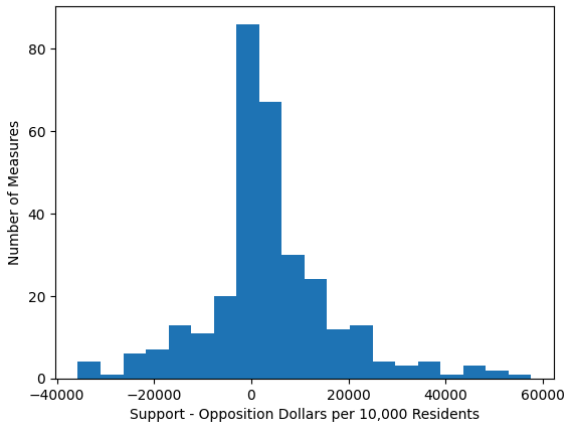


Figure M: Histogram of ballot measure fundraising differences per 10,000 residents

B. Gradient-Boosted Tree Models

After tuning a gradient-boosted tree according to the optimal learning parameters derived from 5-fold cross-validation for the ballot measure support and roll-off outcome variables, I obtained feature importances for each variable.

The optimal model for ballot measure support reported a root mean squared error (RMSE) of .1291 and R^2 of .3017. The most important predictors of support according to this model were U.S. state, ballot measure type, word length, roll-off, dollars raised in opposition to the measure, and dollars raised in support of the measure (see Figure N). I ran a second iteration of the model with only the variables with F-score weights of 15.0 or higher and observed comparable performance at an RMSE of .1297 and R^2 of .2948. Although predictive power based on R^2 was moderate to weak, RMSE significantly improved from my baseline model that always predicted the mean vote share, which yielded an RMSE of .1547.

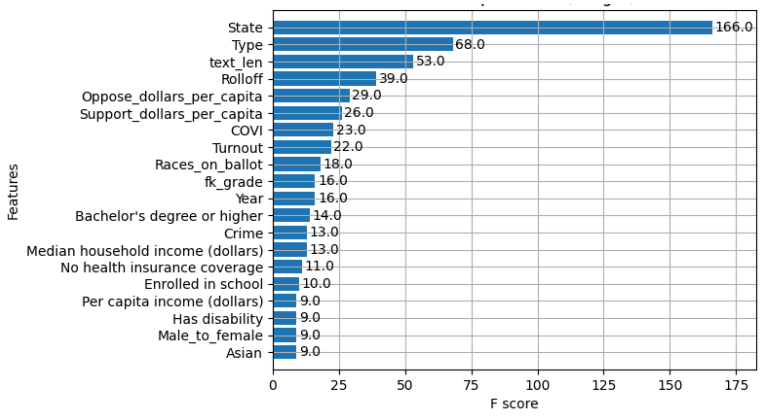


Figure N: Ballot measure support gradient-boosted tree feature importances

Utilizing the same process for ballot measure roll-off, my optimally tuned gradient-boosted tree model yielded an RMSE of .0284 and R^2 of .4236, an even more substantial improvement from my baseline model RMSE of .0375. High importance features when predicting roll-off included U.S. state, word count, ballot measure support, and number of races on the ballot, among others (see Figure O). The simplified model with only features with F-scores equal to or greater than 50 also had similar performance at an RMSE of .0287 and R^2 of .4136.

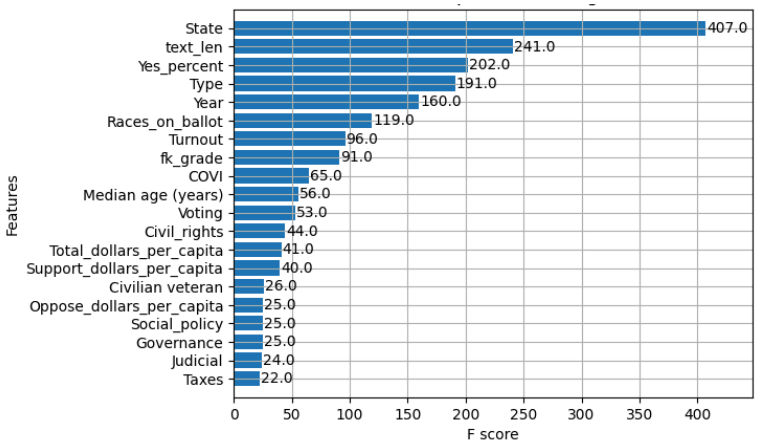


Figure O: Ballot measure roll-off gradient-boosted tree feature importances

C. Linear Regression Models

The final step of my analysis was to run linear regression models to discern effects of ballot text, campaign finance, voting difficulty, and other variables on ballot measure support and roll–off. For each outcome variable, I tested the assumptions for a linear regression model—outputs of these tests can be found in Appendix B. I ran three regressions for each outcome—one with the full set of predictors (“full” model), one omitting Flesch-Kincaid grade

level to increase observations ("reduced" model), and one with high-importance predictors defined in the gradient-descent tree models ("high-importance" model).

1. Ballot Measure Support

Of the ballot complexity variables—Flesch-Kincaid reading grade level, formatting as a question, multiple sentences, and presence of a fiscal statement—only multiple sentences was a statistically significant predictor of ballot measure support at $\alpha = .05$, and only in the reduced model that did not control for reading grade level. In this model, a ballot measure with multiple sentences is expected to have a 3.3 point decrease in support, but this effect is rendered statistically insignificant in the full model.

Ballot length (number of other ballot measures in the election) and ballot measure word length also had non-significant coefficients. Alongside the ballot complexity coefficients, neither ballot text complexity nor ballot length appear to significantly affect the share of individuals that vote in favor or against a given ballot measure.

On the other hand, the campaign finance variables proved significant predictors of ballot measure support. Both dollars raised per capita in support of a measure and dollars raised per capita in opposition were highly statistically significant, with coefficient values of .023 and -.025, respectively. This indicates that for every additional dollar raised per capita in support of a measure, the measure would experience an increase in support of 2.3 points, while every additional dollar raised per capita against the measure decreases support by 2.5 points. Total dollars raised per capita did not have a significant effect on support, nor did it yield a substantively large coefficient (-.002), suggesting both support and opposition issue campaigning are effective in swaying final vote shares, and to the degree that one is greater than the other.

Notably, opposition fundraising, which descriptive analysis showed as less frequent and smaller in magnitude than support fundraising, had the slightly larger absolute effect on measure support of the two.

Partisanship score did not appear to significantly change measure support in any of the models.

A state's relative COVI score did not have a significant effect on measure support, while statewide electoral turnout had a highly significant positive effect in the reduced model (with the greatest number of observations). Turnout's coefficient of .233 indicates that for every ten point increase in general turnout (not to be confused with measure-specific turnout), support for a measure increases by 2.33 points. Contrary to the negative effect hypothesized, this relationship may imply that an increase in voters (voters on the margin of turning out tend to have less political knowledge and voting experience) who are unconfident in their ballot measure preferences will abstain rather than necessarily vote against a measure by default.

Several variables outside of my initial hypotheses also reported high significance. In the full model, every 10 percentage point increase in the state population's Bachelor's degree attainment rate was significantly associated with a 1.09 point increase in ballot measure support (and marginally significantly associated with a .25 point increase in support in the reduced model). Black or African American population share was highly significant in the reduced model, where a 10 point increase in Black or African American share was expected to raise measure support by 1.05 points. Measures related to civil rights and judicial reform were both significant in the reduced model with negative coefficients of -.082 and -.066, respectively, meaning such measures received 8.2 and 6.6 points less in support than for miscellaneous subjects. Finally, bond issues and legislatively referred ballot measures (as opposed to initiatives)

had significantly positive effects on support in the reduced and high-importance models, ranging from large increases in support from 19.4 points to 32.7 points, revealing that bond issues and legislative referrals tend to be less controversial or unpopular than other types of measures.

A regression summary table of all variables of interest and significant predictors can be found in Appendix C.1.

2. Ballot Measure Roll-off

Like with ballot measure support, multi-sentence ballot measures were the only ballot text accessibility feature significantly associated with roll-off, where a coefficient of -.03 in the full model can be interpreted as a 3 point decrease in roll-off when a measure is divided from one sentence into more digestible segments. While 3 points may not at first seem significant, this is a large improvement in turnout for a given measure when roll-off generally sits below 5 percent.

While measure-specific word count was once again unrelated to roll-off, the number of other measures on the ballot had a significant and marginally significant positive effect on roll-off in the reduced and high-importance models, respectively. Both models yielded coefficients values around .003, which translates to a .3 point increase in ballot measure abstentions for each additional measure also found on the ballot.

Neither dollars raised per capita in support of a ballot measure nor dollars raised per capita in opposition were significant predictors of roll-off. Total dollars raised per capita had a significant effect in the reduced model, but its coefficient of -.001 indicates a near-negligible effect in practice. Campaign finance did not appear to change voters' decision to vote or abstain on a measure relative to others on the ballot. This does not speak to how campaign finance

affects turnout to the ballot box in general—large issue campaigns may still drive greater turnout, but for all measures on the ballot that the target measure is on.

This time, COVI had a significant positive effect on roll-off in the reduced, but negative, even if insignificant, coefficients in the other regressions point to an inconclusive effect. More concretely, overall electoral turnout had a positive effect in all models, with high significance in the reduced model. Its coefficient of .068 signals that for every 10 point increase in turnout, roll-off rises by .68 points. This supports the theory that low-information voters that participate in high-turnout elections tend to abstain at higher differentials from measures they are uncertain about compared to those they are more familiar with.

As with measure support, partisanship had no significant effect on roll-off.

Moving to variables outside my hypotheses, measures about state finances yielded a significant coefficient of .012 in the reduced model (a decrease in measure turnout of 1.2 points), which aligns with the expectation that voters will abstain at greater rates from measures whose subjects are complicated and unfamiliar. Also in the reduced model, states with greater shares of residents who speak English less than "very well" saw a significant decrease in roll-off. While this finding contrasts with the general idea that voters that have more difficulty comprehending a ballot measure will tend to abstain, it supports prior research that, when faced with complex ballots, voters with lower English language ability may be more willing to accept low certainty and cast ballots that conflict with their preferences compared to their English-proficient counterparts (Ulbig and Reilly 2021). Finally, while total population had a significant coefficient in the reduce model, it had a negligible coefficient of .000.

A regression summary table of all variables of interest and significant predictors can be found in Appendix C.2. The Python notebook containing all data collection, cleaning, and analysis for this project can be found in Appendix D.

V. Discussion

My analysis of predictors of ballot measure support and roll-off illuminate several key insights into how ballot presentation and issue campaigns influence direct democratic policy outcomes. For one, the complexity of ballot text and word count largely did not have significant effects on either support or roll-off compared to other electoral details. The exception was when a ballot measure was broken down into multiple sentences, which significantly improved measure response rates and calls for further research into latent ballot design features that improve voter comprehension.

Following the voter fatigue literature, more races on a ballot significantly increased the rate of abstentions from the highest-response measure on the ballot. In particular, elections with high turnout experienced greater rates of roll-off, likely from less frequent and lower-information voters. These individuals may lack prior knowledge and practice voting on complex ballot measures, in turn deciding not to vote at all. Though, there is evidence of an exception: voters with lower English proficiency. These voters may cast a ballot that they are not fully confident reflects their true policy beliefs, risking not only a vote away from their preference but in favor of the opposite. In either case, a rich direct democratic system strives to include and accurately represent these voters with ample civic education before reaching the ballot box and through accessible information on the ballot.

Strongly funded issue campaigns proved a significant player in influencing voters' preferences both in support of and against ballot measures, but not necessarily in reducing abstentions from those without or unable to develop a confident opinion. Thus, while ballot measure campaigns can provide compelling information to voters would might otherwise vote differently and/or with less knowledge, they appear to be ineffective in mobilizing voters that

decide to skip the question entirely. Civic education efforts may be better served by initiatives that instill general habits of civic participation and confidence translating preferences to a vote. At the same time, issue campaigns' proven ability to change the outcome of elections adds credence to the warning that unregulated money in elections can depart from welfare-enhancing public education into one-sided races engineered to favor the biggest check. My descriptive analysis indicated numerous ballot measure races have large disparities in and/or uncontested spending between supporters and opponents.

In the final section of this paper, I make several policy recommendations to address how lawmakers can enhance voter confidence, civic inclusion, and the integrity of direct democratic processes.

VI. Policy Recommendations

1. Ballot Text Accessibility

My analysis cast doubt on the extent to which general ballot complexity and word length alter support and roll-off (though it importantly does not disprove literature on diminishing voter satisfaction and ability to vote accurately according to one's preferences). However, the significant drop-off in abstentions for measures broken into multiple sentences and on shorter ballots with few other measures suggests that ballot writers should further research and consider how unexplored features of ballot presentation can improve voting confidence.

New York advanced a bill in 2023 requiring a descriptive ballot title of up to fifteen words, summary of up to thirty words, and statement of what a "yes" or "no" vote means, entirely in plain language (Ballotpedia 2023b), which passed unanimously and with bipartisan support (New York State Assembly 2023). But liberal states aren't the only ones to codify accessible ballot measures; Alabama adopted similar requirements for ballot measures to include a summary, effect of the measure if passed and if defeated, and its cost and source of funding, all in "plain, nontechnical language […] using words with common and every day meaning that are understandable to the average reader" (Ballotpedia 2023b). These examples, as well as known recommendations from organizations like the Center for Civic Design (Johnson 2023; Johnson 2024) provide a clear framework on how to standardize ballot design to be as accessible as possible, from explanations of technical changes to intuitively defined answer choices.

2. Issue Campaign Finance Regulation

Issue campaigns had a real impact on electoral outcomes, but proven one-sided spending can hinder the free flow of civic information on consequential policy decisions. The Centre for

Economic Policy Research (CEPR) proposes that sound campaign finance regulations level the playing field and enhances healthy electoral competition while preserving the representativeness of electoral outcomes (Tricaud et al. 2022). Countries like Brazil and the U.K. set mandatory spending limits, and others, including Canada, France, Italy, and South Korea, also incorporate public reimbursement of campaign expenditures to support candidates with fewer resources. The U.S., while more relaxed, still requires presidential candidates to heed a spending limit in order to receive public reimbursements. CEPR's paper on French elections (which follow the U.S.' plurality electoral system) concludes both spending caps and campaign expenditures reimbursement significantly equalize the competitiveness of non-incumbent candidates while maintaining winners that are equally representative of constituents' political orientations (Tricaud et al. 2022).

In the U.S., ballot measures have always been a state and local affair. U.S. states generally impose campaign finance reporting requirements on issue campaigns as they do candidate campaigns, but the relative absence of spending limits and public reimbursements enable outsized spending on behalf of well-resourced special interests. States should strongly consider adopting more robust campaign finance regulations for ballot measures, in particular public reimbursement. By financially incentivizing even poorer-resourced groups to educate the public about low-salience causes on the ballot, voters receive a stronger and more well-balanced stream of information about the issues they will ultimately decide on.

3. Caveat to Voter Access

Regression analysis demonstrated that the turnout of voters with less information and lower levels of formal education resulted in hesitancy to support or vote at all on complex ballot

measures. These individuals are particularly disadvantaged in the calculus of voting—marginalized communities, including young people, racial minorities, people with lower education levels, and those with lower incomes face greater voter education costs (Vij 2020) due to separation from voter education resources, weak targeting by issue campaigns, and lacking excess time outside of work to dedicate to research. One research survey found that while ballot measures lose support as they increase in complexity, the effect is smaller among highly educated voters (Hessami and Resnjanskij 2019).

U.S. states have ample work remaining to liberalize access to voting, whether facilitating easy-to-navigate voter registration processes or expanding opportunities to visit polling places to vote. Increasing the share of people who are habitual voters may improve how representative voters are of the population, but findings that marginalized and infrequent voters still struggle to digest ballot measures, even after taking pains to reach the ballot box, undermine the gains of greater voter access. Ballot text accessibility and issue campaign finance reform stand out as crucial goals to arm voters with the information to holistically understand their ballot.

In sum, the tenability of ballot measures as a good-faith practice of direct democracy rests on policy that promotes inclusive civic engagement, as electoral results are only as representative as the representativeness of confident votes behind them.

VII. References

A. Research References

Ballotpedia. 2023a. “Ballot measure readability scores, 2023.” *Ballotpedia*. https://ballotpedia.org/Ballot_measure_readability_scores,_2023 (December 30, 2024).

Ballotpedia. 2023b. “Readability laws for ballot measure language.” *Ballotpedia*. https://ballotpedia.org/Readability_laws_for_ballot_measure_language (December 30, 2024).

Ballotpedia. 2024a. “2024 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2024_ballot_measures (December 30, 2024).

Barth, Jay, Craig M Burnett, and Janine Parry. 2019. “Direct Democracy, Educative Effects, and the (Mis)Measurement of Ballot Measure Awareness.” *Springer Link*. https://doi.org/10.1007/s11109-019-09529-w (December 30, 2024).

Berman, David R., and Mike Yawn. 2024. “Partisanship and Support for Propositions on the Ballot: A State-Level Longitudinal Study.” *JSTOR*. https://www.jstor.org/stable/42955729 (December 30, 2024).

Briffault, Richard. 1996. “Ballot Propositions and Campaign Finance Reform.” *Columbia Law School Scholarship Archive*. https://scholarship.law.columbia.edu/cgi/viewcontent.cgi?article=1906&context=faculty_scholarship (December 30, 2024).

Brower, Mac. 2023. “Republicans Hint at Why They Are Restricting Ballot Measures in These States.” *Democracy Docket*. https://www.democracydocket.com/analysis/republicans-hint-at-why-they-are-restricting-ballot-measures-in-these-states/ (December 30, 2024).

Carter, Sara, Alice Clapman, and Alexi Comella. 2024. “Politicians Take Aim at Ballot Initiatives.” *Brennan Center for Justice*. https://www.brennancenter.org/our-work/research-reports/politicians-take-aim-ballot-initiatives (December 30, 2024).

Childers, Matt, and Mike Binder. 2015. “The Differential Effects of Initiatives and Referenda on Voter Turnout in the United States, 1890-2008.” *Dale E. Fowler School of Law*. https://www.chapman.edu/law/_files/publications/Volume%2019/clr-19-childers-binder.pdf (December 30, 2024).

CIRCLE at Tufts University. 2014. “The Effects of Competitiveness and Ballot Measures on Youth Turnout.” *CIRCLE*. https://circle.tufts.edu/latest-research/effects-competitiveness-and-ballot-measures-youth-turnout (December 30, 2024).

Galston, William A. 2004. “Civic Education and Political Participation.” *JSTOR*. https://www.jstor.org/stable/4488817?seq=2 (December 30, 2024).

Gerber, Alan. 2004. “Does Campaign Spending Work in the United States?” *The Abdul Latif Jameel Poverty Action Lab*.

https://www.povertyactionlab.org/evaluation/does-campaign-spending-work-united-states (December 30, 2024).

Hessami, Zohal, and Scen Resnjanskij. 2024. “Complex ballot propositions, individual voting behavior, and status quo bias.” *ScienceDirect*. https://www.sciencedirect.com/science/article/abs/pii/S0176268018302799 (December 30, 2024).

Joffe, Marc. 2024. “Deceptive Advertising on the Ballot.” *Cato at Liberty Blog*. https://www.cato.org/blog/deceptive-advertising-ballot (December 30, 2024).

Johnson, Isamu S. 2023. “Explaining ballot questions in plain language.” *Center for Civic Design*. https://civicdesign.org/wp-content/uploads/2023/06/2023-Clarity-85-Explaining-ballot-questions-in-plain-langauge.pdf (December 30, 2024).

Johnson, Isamu S. 2024. “Ballot questions are hard to understand. Here are 6 ways to fix them.” *Center for Civic Design*. https://civicdesign.org/ballot-questions-are-hard-to-understand-here-are-6-ways-to-fix-them (December 30, 2024).

Milita, Kerri. 2017. “Beyond Roll-off: Individual-Level Abstention on Ballot Measure Voting.” *Taylor & Francis Online*. https://www.tandfonline.com/doi/full/10.1080/17457289.2017.1300157 (December 30, 2024).

National Conference of State Legislatures. 2024. “NCSL Ballot Measures Database.” *National Conference of State Legislatures*. https://app.powerbi.com/view?r=eyJrIjoiYjEwNDI2NTctZDFkMy00ZGM4LWFkMTItNTcwYTdkZmMxMGIxIiwidCI6IjM4MmZiOGIwLTRkYzMtNDEwNy04MGJkLTM1OTViMjQzMmZhZSIsImMiOjZ9 (December 30, 2024).

New York State Assembly. 2023. “6-8-23 Session.” *New York State Assembly*. https://nystateassembly.granicus.com/MediaPlayer.php?view_id=7&clip_id=7672&meta_id=211248 (December 30, 2024).

Pomante II, Michael J. 2023. “Data Details.” *Cost of Voting Index*. https://costofvotingindex.com/data (December 30, 2024).

Primo, David M. 2013. “Information at the Margin: Campaign Finance Disclosure Laws, Ballot Issues, and Voter Knowledge.” *University of Rochester Department of Political Science*. https://www.sas.rochester.edu/psc/primo/PrimoELJ2013.pdf (December 30, 2024).

Riker, William H., and Peter C. Ordeshook. 1968. “A Theory of the Calculus of Voting.” *JSTOR*. https://www.jstor.org/stable/pdf/1953324.pdf (December 30, 2024).

Rogers, Todd, and Joel Middleton 2014. “Are Ballot Initiative Outcomes Influenced by the Campaigns of Independent Groups? A Precinct-Randomized Field Experiment Showing That They Are.” *Scholars at Harvard*. https://scholar.harvard.edu/files/todd_rogers/files/rogers_and_middleton.2014.pdf (December 30, 2024).

Rudensky, Yurij. 2024. “Ohio Politicians Try to Deceive Voters About Anti-Gerrymandering Ballot Measure.” *Brennan Center for Justice*. https://www.brennancenter.org/our-work/analysis-opinion/ohio-politicians-try-deceive-voters-about-anti-gerrymandering-ballot (December 30, 2024).

Schuster, Steven S. 2020. “Does Campaign Spending Affect Election Outcomes? New Evidence from Transaction-Level Disbursement Data.” *The University of Chicago Press*. https://www.journals.uchicago.edu/doi/10.1086/708646 (December 30, 2024).

Serna Jr., Albert.. 2024. “State ballot measures attract more than $417 million ahead of Election Day.” *OpenSecrets*. https://www.opensecrets.org/news/2024/10/state-ballot-measures-attract-more-than-417-million-ahead-of-election-day/ (December 30, 2024).

Smyth, Julie C. 2024. “Ohio lawsuit seeks rewrite of redistricting ballot language dubbed ‘biased, inaccurate, deceptive.’” *AP News*. https://apnews.com/article/election-2024-redistricting-ohio-ballot-language-lawsuit-d51a71bb3992d15da2d95bdceb4e4f01 (December 30, 2024).

Stapleton, Carey E., and Stefani R. Langehennig. 2024. “Partisanship and voting behavior reconsidered in the age of polarization.” *ScienceDirect*. https://www.sciencedirect.com/science/article/abs/pii/S0261379424000106 (December 30, 2024).

Stratmann, Thomas. 2010. “Campaign Spending and Ballot Measures.” *SpringerLink*. https://link.springer.com/chapter/10.1057/9780230248656_2 (December 30, 2024).

Ulbig, Stacy G., and Shauna Reilly. 2021. “Compounded Confusion? Ballot Language Complexity, English Proficiency, and Minority Language Voter Behavior.” *Cost of Voting Index*. https://digitalcommons.coastal.edu/jops/vol49/iss1/2/ (December 30, 2024).

United States Census Bureau. 2024. “Educational Attainment.” *Census Bureau Data*. https://data.census.gov/table?q=Educational%2BAttainment (December 30, 2024).

Vij, Sarina. 2020. “Why Minority Voters Have a Lower Voter Turnout: An Analysis of Current Restrictions.” *American Bar Association*. https://www.americanbar.org/groups/crsj/publications/human_rights_magazine_home/voting-in-2020/why-minority-voters-have-a-lower-voter-turnout/ (December 30, 2024).

B. Data Sources

Ballotpedia. n.d. “2004 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2004_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2005 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2005_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2006 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2006_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2007 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2007_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2008 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2008_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2009 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2009_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2010 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2010_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2011 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2011_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2012 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2012_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2013 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2013_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2014 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2014_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2015 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2015_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2016 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2016_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2017 ballot measures.” *Ballotpedia*.
https://ballotpedia.org/2017_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2018 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2018_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2019 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2019_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2020 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2020_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2021 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2021_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2022 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2022_ballot_measures (December 30, 2024).

Ballotpedia. n.d. “2023 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2023_ballot_measures (December 30, 2024).

Ballotpedia. 2024b. “Historical and potential changes in trifectas.” *Ballotpedia*. https://ballotpedia.org/Historical_and_potential_changes_in_trifectas (December 30, 2024).

Ballotpedia. 2024c. “Presidential voting history by state.” *Ballotpedia*. https://ballotpedia.org/Presidential_voting_history_by_state (December 30, 2024).

Ballotpedia. 2024d. “Voter turnout in United States elections.” *Ballotpedia*. https://ballotpedia.org/Voter_turnout_in_United_States_elections (December 30, 2024).

Ballotpedia. 2024e. “2024 ballot measures.” *Ballotpedia*. https://ballotpedia.org/2024_ballot_measures (December 30, 2024).

OpenSecrets. 2024. “Ballot Measures.” *OpenSecrets*. https://www.opensecrets.org/ballot-measures/ (December 30, 2024).

Pomante II, Michael J. 2023. “Data Details.” *Cost of Voting Index*. https://costofvotingindex.com/data (December 30, 2024).

United States Census Bureau. 2024. “DP05: ACS DEMOGRAPHIC AND HOUSING ESTIMATES.” *Census Bureau Data*. https://data.census.gov/table/ACSDP5Y2010.DP05 (December 30, 2024).

United States Census Bureau. 2024. “S0201: SELECTED POPULATION PROFILE IN THE UNITED STATES.” *Census Bureau Data*. https://data.census.gov/table/ACSSPP1Y2010.S0201 (December 30, 2024).

VIII. Appendix

A. Gradient-Boosted Tree Tuning Parameters

The optimal parameter outputs of my ballot measure support and roll-off predictive models with 5-fold cross-validation were as follows:

A.1. Ballot Measure Support

Optimal parameters:

- Subsample ratio of columns for each tree = .8
- Learning rate = .1
- Max tree depth = 3
- Number of boosting rounds = 100
- Subsample ratio of training instance = .8

Best RMSE: .1283

A.2. Ballot Measure Roll-off

Optimal parameters:

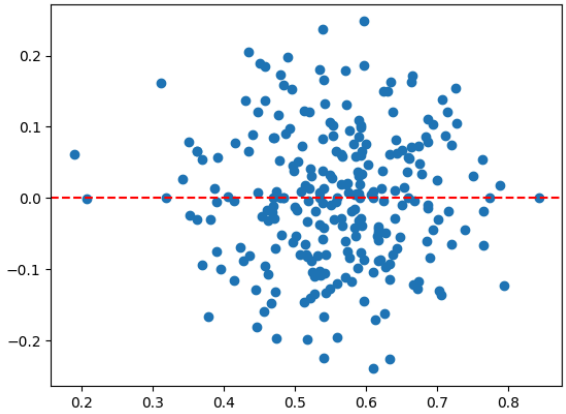
- Subsample ratio of columns for each tree = .8
- Learning rate = ..05
- Max tree depth = 5
- Number of boosting rounds = 100
- Subsample ratio of training instance = .8

Best RMSE: .0267

B. Linear Regression Assumption Tests

B.1. Ballot Measure Support

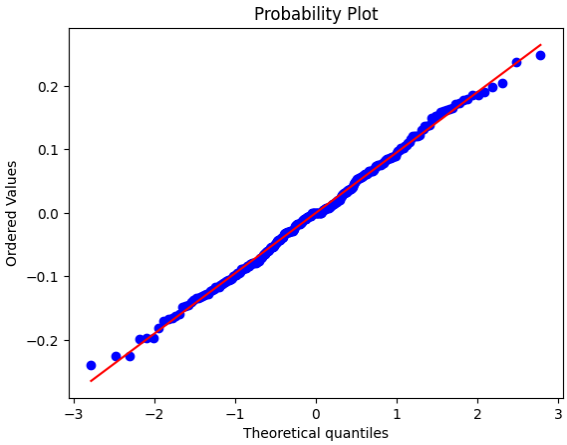
Prediction **residual plot** to verify **linearity**:



Breusch-Pagan test for **homoskedasticity** p-value: .8758 (> .05 indicates homoskedasticity)

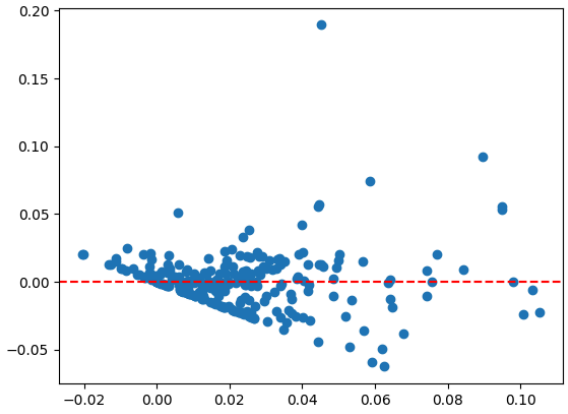
Durbin-Watson test for **independence**: 2.2092 (~2 indicates independence)

Quantile-Quantile plot to verify residuals follow **normal distribution**:



B.2. Ballot Measure Roll-off

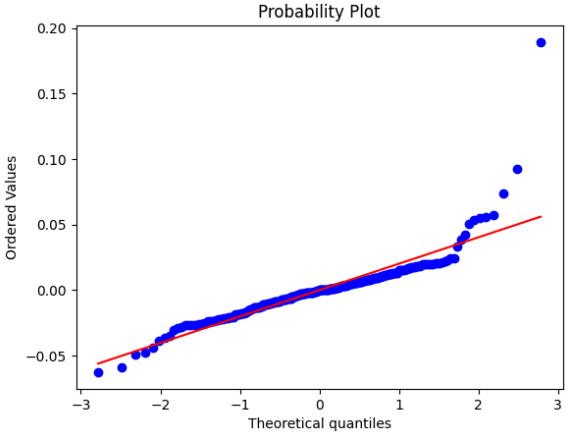
Prediction **residual plot** to verify **linearity**:



Breusch-Pagan test for **homoskedasticity** p-value: .9988 (> .05 indicates homoskedasticity)

Durbin-Watson test for **independence**: 1.9976 (~2 indicates independence)

Quantile-Quantile plot to verify residuals follow **normal distribution**:



C. Linear Regression Summary Tables

C.1. Ballot Measure Support

	Dependent variable: Measure support		
	Full	Reduced	High-Importance
(Intercept)	142.987*	17.398	-4.875
	(78.282)	(30.049)	(4.449)
Flesch-Kincaid grade	0.000		0.000
	(0.000)		(0.000)
Is question	-0.045	-0.022	
	(0.049)	(0.023)	
Multiple sentences	-0.059	-0.033**	
	(0.052)	(0.017)	
Fiscal statement	-0.019	-0.007	
	(0.039)	(0.020)	
Number of words	-0.000	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)
Races on ballot	0.002	0.001	0.001
	(0.006)	(0.002)	(0.003)
Support dollars per capita	0.023***	0.009*	0.010
	(0.008)	(0.005)	(0.007)
Oppose dollars per capita	-0.025***	-0.009**	-0.026***
	(0.007)	(0.004)	(0.006)
Total dollars per capita	-0.002	0.000	
	(0.003)	(0.002)	
Partisanship	0.073	0.038	
	(0.054)	(0.024)	
COVI	0.061	-0.012	-0.006
	(0.052)	(0.025)	(0.018)
Turnout	0.224	0.233***	0.034
	(0.194)	(0.089)	(0.091)
Rolloff	-0.391	-0.027	-0.173
	(0.332)	(0.193)	(0.249)
Bachelor's degree or higher	0.109***	0.025*	
	(0.032)	(0.014)	
Black or African American	0.139	0.105***	
	(0.100)	(0.038)	
Civil rights	-0.078	-0.082**	
	(0.060)	(0.038)	
Judicial	-0.051	-0.066**	
	(0.051)	(0.031)	
Kentucky	0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)
Bond Issue	0.140	0.248**	0.299**
	(0.203)	(0.104)	(0.150)
Leg. Bond Measure	0.000***	0.327***	0.000***
	(0.000)	(0.124)	(0.000)
Leg. Referred Con. Amend.	0.118	0.211**	0.261**
	(0.152)	(0.094)	(0.132)
Leg. Referred State Statute	0.095	0.194**	0.264*
	(0.156)	(0.097)	(0.136)
R^2	0.543	0.384	0.405
Residual Std. Error	0.119	0.122	0.119
F Statistic	2.017***	3.454***	3.143***
Note:			*p<0.1; **p<0.05; ***p<0.01

C.2. Ballot Measure Roll-off

	Dependent variable: Measure roll-off		
	Full	Reduced	High-Importance
(Intercept)	-17.751	-5.791	-19.758
	(18.655)	(6.481)	(12.609)
Flesch-Kincaid grade	-0.000		0.000
	(0.000)		(0.000)
Is question	0.008	0.004	
	(0.012)	(0.005)	
Multiple sentences	-0.030**	-0.003	
	(0.012)	(0.004)	
Fiscal statement	0.009	0.003	
	(0.009)	(0.004)	
Number of words	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Races on ballot	0.002	0.003***	0.002*
	(0.001)	(0.001)	(0.001)
Support dollars per capita	0.002	-0.002	0.000
	(0.002)	(0.001)	(0.002)
Oppose dollars per capita	-0.003	0.001	-0.001
	(0.002)	(0.001)	(0.002)
Total dollars per capita	-0.001	-0.001**	-0.001
	(0.001)	(0.000)	(0.001)
COVI	-0.002	0.012**	-0.013
	(0.012)	(0.005)	(0.010)
Turnout	0.010	0.068***	0.026
	(0.046)	(0.019)	(0.037)
Partisanship	0.016	-0.005	0.018*
	(0.013)	(0.005)	(0.009)
Finance	-0.004	0.012**	
	(0.010)	(0.006)	
Language other than English	0.005	0.007*	
	(0.010)	(0.004)	
Speak English less than "very well"	0.004	-0.013**	
	(0.016)	(0.006)	
Total population	0.000	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)
R^2	0.509	0.504	0.445
Residual Std. Error	0.028	0.026	0.028
F Statistic	1.756***	5.638***	1.935***
Note:			*p<0.1; **p<0.05; ***p<0.01

D. Data Analysis Notebook

The full Python notebook in which I collected, cleaned, and analyzed all data for this research project can be found at the following link:

https://github.com/andrewdkim7/portfolio/blob/d438403f90c91adc2a03586bba8398e974798b44/Python/BallotMeasures.ipynb