Pen, to Paper, to the Polls: Initiative Signature Gathering and Electoral Outcomes

Benjamin Donohoe

Abstract. Does the act of signing a petition to place initiatives on the ballot make a voter more likely to vote for that measure? Previous research suggests that television advertising can influence vote share for ballot propositions and that signature gathering campaigns increase registration and turnout and decrease ballot roll-off but has failed to examine the link between petition drives and vote share. Using heteroskedasticity-robust regression with clustered standard errors and multiple levels of fixed effects, I examine whether either statewide spending on petition circulation or the number of voters a petition campaign reaches in a county has a significant impact on vote share. Though statewide spending analysis proves largely inconclusive, the strength of a petition drive within a county has a significant, positive effect on vote share even after endogenizing campaign strategy and incorporating turnout effects.

I. Introduction

Since 2000, over 700 voter-initiated ballot measures have appeared on ballots in statewide elections (Institute for Public Policy and Social Research [IPPSR], 2016). Because these measures must gain support in the form of voter signatures, an entire industry of political consulting firms focused on signature collecting has formed across the country. Indeed, in California alone, over $150 million has been spent since 2000 on gathering signatures for ballot initiatives (California Secretary of State [CA SOS], 2020a).

To that extent, spending on signature gathering significantly affects whether an initiative gets on the ballot. But do these signature gathering campaigns have an impact on whether the ballot initiative is passed? Can signature gathering campaigns serve a dual purpose—qualify a measure for the ballot, and build electoral support for the measure via connections with voters?

If signature gathering campaigns are effective means of consciousness raising among voters and end up spurring voters to the ballot box, we could see a significant correlation between signature spending and outcomes, suggesting that committees could use these campaigns to strategically motivate voters. If more intense signature gathering campaigns only reach already mobilized supporters, we could see limited effects on electoral outcomes.

The strength of these signature campaigns can be measured in two ways: the total amount of money spent on the campaign, and the actual number of signatures gathered in the campaign. The number of signatures required for an initiative to qualify for the ballot depends on the turnout for the previous gubernatorial election (CA SOS, 2020b).

To analyze spending at the statewide level (with each ballot initiative serving as one observation), using the amount spent per required signature as a metric of signature campaign strength best accounts for changes in the required number of signatures for an initiative. For county-level analysis (with
an individual county’s vote on a single ballot measure serving as one observation), the percent of registered voters that signed the petition in a county best measures the strength of a petition drive.

The effectiveness of signature gathering campaigns as a predictor of success could simply reflect the ability of an organization to fund and organize its campaign. However, if we view ballot measure campaigns as making strategic choices to allocate funds and workers to signature gathering campaigns, we can analyze whether these strategic choices are well-founded—whether a more intense signature gathering campaign implies a higher or lower likelihood of success. These campaigns could be an effective means of consciousness raising and shaping voter opinions.

Furthermore, if little to no correlation is found, firms may be incentivized to spend as little as possible on gathering signatures, focusing their efforts on other campaign outlets. Indeed, this lack of relationship between signature gathering spending and ballot outcomes may suggest some waste on the part of ballot measure committees, who may not be maximizing votes per dollar spent. If more intense signature drives correlate with poor electoral outcomes, signature gathering campaigns could be a sign of difficulty gathering support.

Finally, a lack of correlation between signature spending and electoral success for qualified measures may quell claims of corporate intrusion into one of the purest democratic institutions. Indeed, a lack of correlation could vindicate campaigns with lower cost-per-signatures, who may rely more on motivated volunteers, suggesting that volunteers and paid signature gatherers are essentially the same with regards to informing and motivating voters to support the petitions they sign. If, in the process of gathering the signatures of 5% of registered voters required to place a measure on the ballot (CASOS 2022), signature gatherers speak with individuals already likely to support a ballot measure, then they may simply be facilitating the movement of policy preferences from voters to enactment into law, as opposed to convincing new voters to sign onto measures they would not have otherwise voted for.

Analyzing the effect of statewide spending for signature gathering on vote share and the success of the measure provides inconclusive results, failing to suggest a causal link between gross spending and outcome. However, at the county level, the percent of voters who signed a ballot petition and the vote share exhibit a statistically significant positive correlation, even when controlling for partisanship, television advertising, and county-level fixed effects. A one-percent increase in the percent of voters who signed a petition correlates with a 1.14 percentage point increase in vote share. The results suggest a causal relationship in which ballot petition campaigns positively affect electoral outcomes.

could see a significant correlation between signature spending and outcomes, suggesting that committees could use these firms to strategically motivate voters. If more intense signature gathering campaigns only reach already mobilized supporters, we could see limited effects on electoral outcomes.

II. Literature Review

Ballot measure committees, which can spend unlimited amounts of money on campaigning (Hasen, 2005), play the primary role in gathering signatures for initiatives. Initial research indicated that spending in favor of a ballot measure has “at best only a moderate connection” with its success at the polls (Owens & Wade, 1986 p. 688), though spending against a ballot measure proved more potent (Matsusaka, 2004). But more recent studies reveal a significant impact of both supporting and opposing spending (Stratmann, 2006; Figueiredo et al., 2011). To this extent, interest groups and individual donors can play a significant role in shaping the outcomes of ballot measures. In a county-level analysis, Stratmann (2006) notes that targeted, strategic spending on television advertisements can have significant impacts on outcomes of votes.

Research into the nature of signature gathering campaigns is limited. Boehmke and Alvarez (2005) note that, based on county-level data, more intense signature gathering campaigns correlate with higher levels of voter registration, higher turnout figures, and lower ballot rolloff (in which voters skip down-ballot races and measures). Boehmke and Alvarez, however, fail to address
whether these signature gathering campaigns affect voting for the actual initiatives they support. In general, information campaigns can significantly shape voter opinions on ballot measures (Burnett, 2013), but the effects of signature campaigns on ballot outcomes remains unstudied. Because an entire industry of paid signature gathering firms has risen across the country, the opportunity for strategic spending on signature gathering could prove an important decision for allocating campaign funds. This paper will explore whether signature gathering campaigns have a significant effect on improving a ballot measure’s chances of success.

III. Statewide Initiative Analysis

A. Data

Data regarding signature spending, television spending, and voter registration come from the website of the California Secretary of State. This raw data is provided as a list of all expenditure line-item entries on campaign committee disclosure forms, entries from which we can isolate spending on petition gathering, as well as television advertisements, a helpful additional covariate. The Secretary of State also provides a list of all ballot measure committees and their respective ballot measures, though some committees formed prior to a measure’s placement on the ballot require additional matching to a ballot measure. Because committees frequently amend their past filings, records that are amended entries for the same expense must be removed, along with all sub-itemization entries. I matched expenditure line-item entries to their respective committees, which I, in turn, matched to specific ballot measures.1 Funds raised against a ballot measure are sourced from the National Institute on Money in Politics; data is far less readily available on ballot measure committees opposing propositions, and committees often oppose multiple ballot measures, making matching committees to expenditure line-items significantly less reliable. Data regarding ballot measure outcomes (and the general list of initiatives) come from the National Coalition of State Legislatures. Finally, data on political party endorsements comes from UC Berkeley’s Institute for Governmental Studies.

Voter registration info comes from the Secretary of State’s report on registered voters issued 15 days prior to the relevant election. During the time period studied, California largely did not engage in purging records of inactive voters (Wildermuth, 2019; City News Service, 2019), and same-day voter registration did not exist in California. Registration is partisan in California, though voters can select “No Party Preference.” This NPP designation may cause some decreased clarity about the electorate, but the amount of voters registered Democrat remains a solid measure of partisanship used in other research.

Data regarding funds raised against ballot measures are available dating back to 2003; all initiatives from the 2003 Special Election onward are included, with some exceptions. Almost all ballot initiative committees engaging in signature spending I encountered supported a single party endorsements comes from UC Berkeley’s Institute for Governmental Studies.

Voter registration info comes from the Secretary of State’s report on registered voters issued 15 days prior to the relevant election. During the time period studied, California largely did not engage in purging records of inactive voters (Wildermuth 2019, City News Service 2019), and same-day voter registration did not exist in California. Registration is partisan in California, though voters can select “No Party Preference.” This NPP designation may cause some decreased clarity about the electorate, but the amount of voters registered Democrat remains a solid measure of partisanship used in other research.

Data regarding funds raised against ballot measures are available dating back to 2003; all initiatives from the 2003 Special Election onward

1 Most committees listed the ballot measures for which they advocate on their Form F460, and this information can be collated with spending items using the CVR_Campaign_Disclosure_CD file. However, committees with alternate names formed prior to their respective ballot measures receiving official proposition numbers from the Secretary of State required manual matching of name to ballot measure. Committees supporting initiatives that failed to make the ballot were verified using information from UC Hastings’s database as well as Ballotpedia.
are included, with some exceptions. Almost all ballot initiative committees engaging in signature spending I encountered supported a single ballot initiative. A few committees spent on television advertisements in favor of one measure and opposed to another; these committees were classified based on the respective measures they supported. However, ballot initiatives with overlapping supporting committees are excluded from the data. Appendix A contains a table of all excluded ballot initiatives; in total, I exclude 18 out of 103 initiatives. I study 85 ballot initiatives.

B. Models

The regression model, using vote share as the dependent variable to measure the effects of spending on signature gathering on ballot initiative outcomes, is defined by the following:

\[ Y_{j,k} = \beta_0 + \beta_1 s_j + \beta_2 t_j + \beta_3 a_j + \beta_4 d_k + \beta_5 r_{j,k} + \epsilon_{j,k} \]  

(1)

\( Y \) is the vote share for a proposition \( j \) in election \( k \), \( s \) is the dollars spent on proposition \( j \) per required signature needed to attain ballot status, \( t \) is the amount of money spent on television advertisements for proposition \( j \), \( a \) is the money raised in opposition of the initiative, \( d \) is the percentage of registered voters who are Democrats at the time of election \( k \), and \( r \) is a metric of turnout representing the percentage of registered voters in election \( k \) who voted on the initiative. I hypothesize that the signs for \( s, t, d, \) and \( r \) will be positive and the sign for \( a \) will be negative.

To use an indicator of a measure’s passage or failure, I will use a binary response model; the logit model best allows us to analyze the relationship between a binary dependent variable and a series of independent variables without providing fitted probabilities that are greater than one or less than zero (Wooldridge 2012). The general form of the binary response model is as follows:

\[ P(z = 1|x) = P(z = 1|s_j, a_j, t_j, d_k, r_{j,k}) \]  

(2)

The logit model is as follows:

\[ P(\text{passed} = 1|x) = G(\beta_0 + \beta_1 s_j + \beta_2 t_j + \beta_3 a_j + \beta_4 d_k + \beta_5 r_{j,k}) \]  

(3)

where

\[ G = \frac{\exp(x)}{1 + \exp(x)} \]  

(4)

\( P(\text{passed}=1) \) is the likelihood of the measure passing, and \( G \) is the logistic function (Wooldridge 2012). One particular issue for both models is the presence of clustered observations for the values of the statewide voter registration measure; all measures for the same election year will have the same percentage of registered Democrats. This clustering results in data that violates our assumptions of homoscedasticity (Wooldridge, 2012). Within a cluster (here, defined by a single election), the standard errors will be correlated (Cameron & Miller, 2015). To eliminate this bias, I would normally cluster the standard errors and test statistics and specify each individual election as the defining variable for the clusters. However, only fifteen clusters (at most) exist for the models, so I choose not to cluster here.\(^2\) The results of the clustered model differ somewhat in significance for the percent Democrat and funds raised against variables, but signature spending is not significant in either model.

Television spending is likely to act as a better metric for spending in support of a ballot initiative, given that television advertisements have been shown to be a significant predictor of ballot measure success (Stratmann, 2006). Using all spending in favor of an initiative would incorporate spending on campaign consultants, office workers, and legal costs, expenses which largely lack any hypothetical causal link to electoral outcomes. Spending against a ballot measure is less likely to encounter administrative costs as these committees do not encounter additional non-petition costs in placing measures on the ballot and defending them.

\(^2\) This method is consistent with the explanation in Cameron and Miller (2015), who note that, though clustering should generally be done at the broadest possible level and that no universal guideline exists for determining how few too many clusters, they offer both 20 and 50 as possible numbers for “too few clusters” depending on the nature of the data.
from legal challenges.

**C. Endogeneity**

Correcting for the strategic, endogenous allocation of resources represents a foundational problem to analyzing campaign decisions and electoral outcomes. This endogeneity could lead to overstating the treatment effect and having upward biased coefficients; some unobserved effect of the amount of voters in a county supportive of an initiative prior to the petition process would be positively correlated with the signatures gathered variable and the vote share for the initiative. I use county- and proposition-level fixed effects and control variables for county ideological leanings and turnout to address this endogeneity. Stratmann (2006) suggests that committees engage in strategic spending on television advertisements, with more competitive races likely to draw higher levels of spending. To control for this variable, he introduces county- and proposition-level fixed effects, as well as county-level data regarding partisan lean. I will attempt to address this endogeneity similarly: for the statewide analysis, I will include the percentage of registered Democrats (partisanship) at the time of the election and turnout as explanatory variables. I would include election-year fixed effects, but these effects appear to be colinear with the partisanship explanatory variable. Additionally, Stratmann (2006) separates observations into those representing left- and right-leaning ballot measures, with measures endorsed by the LA Times viewed as left-leaning. I will use the endorsements of the California Democratic Party to determine the ideological lean of a ballot measure. Cases that the party remained neutral by not issuing an endorsement are treated the same as measures the party opposed.3

This model assumes that spending and vote share is distributed homogeneously across the state, ignoring any local political characteristics or localized campaign strategy. Campaign committee strategy is essentially limited to choosing a type of spending in this model; however, donors can implicitly choose which initiatives to support. Additionally, I assume that political party endorsements of ballot initiatives reflect their true policy goals.

**D. Results**

|                      | (1)       | (2)       | (3)       |
|----------------------|-----------|-----------|-----------|
| Signature Spending F | -0.38     | 1.41      | -3.32*    |
| (Millions of Dollars)| (0.91)    | (1.09)    | (1.36)    |
| Percent Reg.         | -4.09*    | 1.77      | 4.91*     |
| Democrats            | (1.80)    | (2.42)    | (2.12)    |
| Money Raised         | -0.07     | 1.11      | -0.06     |
| Against (Millions    | (0.05)    | (0.07)    | (0.06)    |
| of Dollars)          |           |           |           |
| TV Spending (Millions| -0.01     | -0.17     | -0.27     |
| of Dollars)          | (0.18)    | (0.24)    | (0.23)    |
| Turnout              | 7.24      | 23.86     | 7.94      |
| (12.29)              | (14.82)   | (15.87)   |
| Constant             | -132.09   | 6.09      | -166.96   |
| (75.70)              | (101.26)  | (90.69)   |
| R-Squared            | 0.12      | 0.26      | 0.30      |
| N. of cases          | 85        | 42        | 43        |

**Table 1. Models of Spending and the Share of the Vote in Favor of the Ballot Initiative, 2003-2018**

Note: Cell entries are variable coefficients with standard errors in parentheses. Dependent variable is an indicator of the initiative’s passage, with 1 signifying a passed measure and 0 signifying a failed measure. * indicates p < 0.05, ** indicates p < 0.01, and *** indicates p < 0.001. (IPPSR, 2016; California Secretary of State, 2020a, 2020b; UC Berkeley Institute for Governmental Studies, 2020; National Institute on Money in Politics, 2020)

![Figure 1. Vote Share vs. Signature Spending, Statewide, 2003-2018](image)
Table 1 shows the results of the statewide regression models where vote share (from 0 to 100) is the dependent variable. Figure 1 presents a scatterplot of vote share vs. funds spent on gathering signatures. Signature gathering spending does not appear to be statistically significant for the models examining all initiatives (1) and initiatives supported by the Democratic party (2). Signature gathering spending is significant for conservative measures (3); however, it has a negative coefficient in this case. The regression suggests that higher levels of spending correlate with lower vote share for conservative initiative outcomes, with $1 million in spending correlating with a 3.32-percentage point decrease in vote share. The percentage of registered Democrats at the time of the election is significant for the model of all initiatives (1) and for conservative initiatives (3); the coefficients do not behave as predicted, as the coefficient for conservative initiatives is larger than that for the liberal initiatives. Turnout for the measure is not statistically significant. The other spending variables are not statistically significant. The oddities of the statewide model likely flow from its failures to endogenize any committee strategy and broad assumptions about the homogeneity of statewide election results.

Table 2 shows the results of the regression model where an indicator of the initiative’s passage or failure (either 1 or 0) serves as the dependent variable. Signature gathering spending is not found to be statistically significant. The percentage of registered Democrats is not statistically significant. Funds raised in opposition to a measure is significant for all initiatives (4) and liberal initiatives (5). Again, with both television advertisement spending and signature spending, a lack of statistical significance (and coefficients that are negative or equal to zero) may stem from failures to properly endogenize campaign strategy and issues of heterogeneity within the state.

V. County-Level Initiative Analysis

Examining ballot initiatives by looking at their county-level results may be more helpful, allowing us more data points and perhaps more variance.

A. Data

I will examine county-level data on initiatives from four elections: the 2012 Primary and General Elections, the 2014 General Election, and the 2016 General Election. I include 28 initiatives out of the 32 on the ballot over these elections. The Secretary of State provides data that lists the number of signatures collected in each county for each ballot measure. The Secretary of States also provides voter registration data and county-level results for each election. For additional county level effects, I will examine television advertisements based on data from the Wesleyan Media Project; though this data lists advertisements based on Nielsen Designated Market Areas (DMAs), we will assume that advertisements air equally across counties. Additionally, I will consider Riverside county, which falls under both the Palm Springs

---

Table 2: Models of Spending and the Passage of a Ballot Initiative, 2003-2018.

Note: Cell entries are variable coefficients with standard errors in parentheses. Dependent variable is an indicator of the initiative’s passage, with 1 signifying a passed measure and 0 signifying a failed measure. * indicates \( p < 0.05 \), ** indicates \( p < 0.01 \), and *** indicates \( p < 0.001 \). (IPPSR 2016; California Secretary of State 2020a, 2020b; IGS 2020; National Institute on Money in Politics 2020)

---

4 The 2014 and 2016 Primary Elections did not have any initiative statutes. The four excluded initiatives are included in Table A1.
and Los Angeles DMAs, as part of the Los Angeles DMA. I will use the same data on party endorsements as in the statewide model.

**B. Model**

In general, the linear model for county-level regression of ballot initiatives outcomes is

\[
V_{ij} = \beta_0 + \beta_1 g_{ij} + \beta_2 a_{ij} + \beta_3 d_{ij} + \beta_4 t_{ij} + \rho_i + \gamma_j + \epsilon_{ij}
\]  

(5)

where \( V \) is the percentage of votes in favor of ballot initiative \( i \) in county \( j \); \( g \) is the number of signatures gathered in a county divided by its eligible voting population; as is the number of advertisements that aired in support of an initiative in that county; \( a \) is the number of advertisements that aired in opposition to an initiative in that county; \( d \) is the percentage of voters in that county who are registered Democrats at the time of the election; and \( t \) is a turnout metric based on the number of votes for and against initiative \( i \) in county \( j \) divided by the number of registered voters in county \( j \). The two parameters \( \rho_i \) and \( \gamma_j \) represent proposition and county fixed effects, respectively. Achieving a majority vote in a county is not a particularly interesting or relevant dependent variable, as the passage of a ballot measure is determined based on state level vote share; thus, I will not include it in this portion of the analysis.

Though Stratmann (2006) used weighted least-squares to correct for heteroskedasticity, the variance of the residuals does not increase with fitted values, suggesting that this heteroskedasticity concern is unfounded. I will present an unweighted regression model and include a brief description of a model using analytical weights for the numbers of eligible voters.

**C. Endogeneity and Robustness**

To address the issue of endogeneity, I will once again include the percent of a county’s registered Democrats as a variable. Additionally, because turnout increases and rolloff decreases in counties with stronger initiative campaigns, the turnout figure that accounts for how many registered voters voted on the issue endogenizes both turnout and rolloff. I will also incorporate both county and proposition level fixed effects to assess any unmeasured factors impacting the passage of a ballot measure. With regards to clustering standard errors, the fixed effects for counties should control for any within-proposition or within-DMA correlation of standard errors. The fixed effects for propositions may not control for all within-county variation, however, so I will cluster standard errors by county. I also choose to not cluster by proposition due to a low number of clusters. 5 Once again, I will separate ballot initiatives into those supported and opposed by the Democratic party to assess varying effects on liberal and conservative ballot measures. To allow for multiple levels of fixed effects, I will use Correia’s reghdfe (2017), which also uses heteroskedasticity-robust standard errors.

This model allows for committee strategic action via the allocation of funds to specific counties. Using fixed effects as well as the registered Democrats variable, it tries to correct the endogeneity of committees spending more on signature gathering in counties likely to support the measure. These fixed effects allow for variation in the likelihood of an individual proposition’s passage or an individual county’s unmeasured characteristics. Additionally, the percent Democrat variable attempts to further control for this endogeneity.

**D. Results**

|                       | (7)      | (8)      | (9)      |
|-----------------------|----------|----------|----------|
| Percent Signatures    | 1.14***  | 0.69***  | 0.95***  |
|                       | (0.14)   | (0.13)   | (0.22)   |
| Percent Registered Democrat | 0.57**  | 0.78**   | 0.10     |
|                       | (0.21)   | (0.20)   | (0.31)   |
| Turnout               | 0.28***  | 0.26***  | 0.20     |
|                       | (0.06)   | (0.06)   | (0.11)   |
| Number of Supporting Ads Aired (1000s) | 1.34*** | 0.69*    | 1.46***  |
|                       | (0.27)   | (0.27)   | (0.38)   |
| Number of Opposing Ads Aired (1000s)   | -0.99*** | -0.60    | -0.39    |
|                       | (0.26)   | (0.34)   | (0.30)   |
| Constant              | 7.13     | 4.57     | 24.42    |
|                       | (7.94)   | (7.39)   | (12.63)  |
| R-Squared             | 0.77     | 0.89     | 0.77     |
| N. of cases           | 1624     | 986      | 638      |

5 When initiatives are split into liberal and conservative, the resulting analyses include 19 and 9 clusters, respectively; as described in footnote 2, this dataset contains too few clusters to use cluster analysis.
Table 3. Linear Models of County Vote, 2012–2016

Note: Percent values are 0 to 100. Cell entries are variable coefficients with standard errors in parentheses.
* indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$. Test statistics are robust to heteroskedasticity.
(California Secretary of State, 2020a, 2020b; Fowler et al, 2015, 2017, 2019; UC Berkley Institute for Governmental Studies, 2020; J. Kaku, personal communication, July 8, 2020; Sood, 2016)

Table 3 shows the results of the regression models when the vote share (from 0 to 100) for a specific initiative in a specific county acts as the dependent variable. Figure 2 illustrates a binned scatterplot of the vote share for an initiative vs. the percentage of voters in a county who signed the petition. The intensity of a signature gathering campaign has a statistically significant, positive correlation with vote share in all three models. For all examined initiatives from 2012 to 2016, a 1 percent increase in the number of voters who sign a petition within a county correlates with a 1.14 percent increase in vote share. The Democratic registration variable behaves as expected, as it is statistically significant for the models of all initiatives (7) and Democratic party-endorsed initiatives (8), with a larger coefficient for the Democratic party-endorsed initiatives. Advertisements for an initiative are statistically significant for all models (7, 8, and 9), and advertisements against an initiative are significant for the model of all initiatives (7). All coefficients for the advertising have the expected signs.

Unexpectedly, the coefficient for percent signatures is greater than 1 for the model of all ballot initiatives (7), suggesting that a one-percent increase in signatures gathered has a greater-than-one percent increase on vote share. This oddity likely occurs because the percent signatures figure represents the percent of registered voters who signed a petition, whereas the election results are taken out of yes and no votes. The endogenization of turnout helped correct this possible error for the other two models (without the turnout explanatory variable, the general model had a coefficient equal to one for percent signatures). This model reasonably endogenizes both initiative committees contacting likely supporters and petition drives increasing turnout. Given that this study endogenizes county-level fixed effects, I argue that, in combination with the results of Boehmke and Alvarez, voters who sign petitions are more likely to be politically engaged and possibly promote political engagement in others and that this political engagement causes a higher level of support for the ballot initiatives the voters sign.

VI. Discussion and Areas of Future Research

This study found that spending on signature gathering does not have a statistically significant impact on electoral outcomes at the statewide level. However, at the county level, the intensity of a signature gathering campaign, as measured by the percentage of voters in a county who signed an initiative’s petition, has a statistically significant, positive effect on electoral outcomes. The findings are consistent with previous research regarding television advertising, which I included
as a variable. An additional percent of voters who signed a petition is associated with an additional 1.14% in vote share.

The models of state- and countywide ballot initiative outcomes differ significantly, especially with regards to signature gathering campaigns. The statewide model likely endogenizes campaign strategy less effectively than the county model. Statewide spending, even on TV ads, fails to capture actual engagement with voters and could include significant overhead costs. Additionally, the statewide model assumes a homogeneous distribution of all voters and spending, failing to account for the possibilities of strategically targeting counties where petition drives could have a more significant impact.

The study’s limitations stem from the possible endogeneity of signature gathering drives. If committees seek out counties where individuals are more politically engaged or better informed, or if counties have local characteristics that increase their likelihood of supporting a given measure, then the relationship between signature gathering campaigns and electoral outcomes may reflect campaign strategy as opposed to petition drives changing and shaping voter opinions. However, incorporating both the partisanship of a county and unobserved county- and proposition-level fixed effects helps mitigate this endogeneity. Additionally, given that previous research into the effects of petition gathering campaigns has demonstrated their positive effects on voter participation (Boehmke & Alvarez, 2005), that signature drives could increase voter support for a measure appears reasonable.

Further research could examine how campaigns maximize support for ballot initiatives, including whether allocating more funds to paid signature campaigns could further improve electoral outcomes. Attempting to link money spent on campaigns with their outcomes could prove difficult given limited data, but a metric of tradeoffs for various campaign strategies (signature drives, television advertisements, or canvassing) could allow for study of marginal votes per dollar spent on each of these strategies. That the results from the state-level analysis suggest higher spending does not improve initiative performance could conform with normative goals surrounding money in politics—that is, interest groups may not be able to buy votes. The overall normative implications of the study rest on the nature of the signature drives themselves. If these drives educate the voters about issues honestly and help facilitate the flow of policy opinions from voters to law, then the effects we observe in this study could align with our normative democratic goals. But if signature campaigns fail to accurately represent issues to voters, the influence on electoral outcomes may be troublesome. Regardless, this study strongly suggests that the face-to-face interactions between petition gatherers and voters and the action of signing one’s name in support of a measure have an impact when that voter steps into the voting booth months later.

VII. Acknowledgements

I would like to express my sincere gratitude to Prof. John Sides, my advisor for this project, for his consistent feedback and mentorship throughout the research process. I also would like to thank Jordan Kaku from the California Secretary of State’s Office for sharing data on ballot initiatives and Prof. Alan Wiseman, my faculty advisor, for his initial thoughts and mentorship. I’d like to thank the VURJ editorial board for their helpful suggestions.
References
Boehmke, F., & Alvarez, R. (2005). The influence of initiative signature-gathering campaigns on political participation. *Social Science Quarterly*, 95(1), 165–183. https://doi.org/10.1111/ssqu.12053
Burnett, C.M. (2013). Does campaign spending help voters learn about ballot measures? *Electoral Studies*, 32(1), 78–89. https://doi.org/10.1016/j.electstud.2012.11.002
California Secretary of State. (2022). *Statewide Initiative Guide 2022*. https://elections.cdn.sos.ca.gov/ballot-measures/pdf/statewide-initiative-guide.pdf
California Secretary of State. (2020a). *CAL-ACCESS raw data*. California Secretary of State. https://www.sos.ca.gov/campaign-lobbying/cal-access-resources/raw-data-campaign-finance-and-lobbying-activity/
California Secretary of State. (2020b). Statewide Election Results. California Secretary of State. https://www.sos.ca.gov/elections/prior-elections/statewide-election-results/
Cameron, A.C., & Miller, D.L. (2015). A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317
City News Service. (2019). LA county and state to purge 1.5 million inactive voters from rolls. NBC4 Los Angeles. https://www.nbclosangeles.com/news/la-county-and-state-to-purge-15-million-inactive-voters-from-rolls/6866/
Correia, S. (2017). reghdfe: Stata module for linear and instrumental-variable/gmm regression absorbing multiple levels of fixed effects. *Statistical Software Components s457874, Boston College Department of Economics*. https://ideas.repec.org/c/boc/bocode/s457874.html
Figueiredo, J., Chang, H.J., & Kousser, T. (2011). Financing direct democracy: Revisiting the research on campaign spending and citizen initiatives. *Journal of Law, Economics, & Organization*, 27, 485–514. https://doi.org/10.1093/jleo/ewr007
Fowler, E.F., Franz, M., & Ridout, T.N. (2015). *Political Advertising in 2012 (V1.1) [Data set]*. Department of Government at Wesleyan University.
Fowler, E.F., Franz, M., & Ridout, T.N. (2017). *Political advertising in 2014 (V1) [Data set]*. Department of Government at Wesleyan University.
Fowler, E.F., Franz, M., Ridout, T.N., & Baum, L.M. (2019). *Political advertising in 2016 (V1) [Data set]*. Department of Government at Wesleyan University.
Hasen, R. (2005). Rethinking the unconstitutionality of contribution and expenditure limits in ballot measure campaigns. *Southern California Law Review*, 78(4), 885–926. http://dx.doi.org/10.2139/ssrn.621321
Institute for Public Policy and Social Research. (2016). *NCSL ballot measures for correlates project*. http://ippsr.msu.edu/public-policy/correlates-state-policy/ncsl-ballot-measures-correlates
Matsusaka, J.G. (2004). *For the many or the few: The initiative, public policy, and American democracy*. University of Chicago Press.
Owens, J.R., & Wade, L.L. (1986). Campaign spending on California ballot propositions, 1925–1984: Trends and voting effects. *The Western Political Quarterly*, 39(4), 675–689. https://doi.org/10.1177/106591298603900408
Proposed ballot measures that were not on a ballot. (2020). Ballotpedia. Retrieved August 6, 2020 from https://ballotpedia.org/Proposed_ballot_measures_that_were_not_on_a_ballot
Stratmann, T. (2006). Is spending more potent for or against a proposition? Evidence from ballot measures. *American Journal of Political Science*, 50(3), 788–801. https://doi.org/10.1111/j.1540-5907.2006.00216.x
Sood, G. (2016). *Geographic information on designated media markets (V9) [Data set]*. Harvard Dataverse, https://doi.org/10.7910/DVN/IVXEHT
UC Berkley Institute for Governmental Studies (2020). *California ballot measure guides*. https://www.igs.berkeley.edu/library/election-guides/ballot-measures
UC Hastings Law. (2020). *UC Hastings law library – California ballot measures database*. https://www.uchastings.edu/academics/library/ca-ballots/
Wildermuth, J. (2019, January 28). Time’s up for inactive voters: Miss too many elections and you’re out. *San Francisco Chronicle*. https://www.sfchronicle.com/politics/article/Time-s-up-for-voters-who-don-t-vote-Miss-two-13565476.php

Wooldridge, J.M. (2012). *Introductory econometrics: A modern approach* (5th ed.). Cengage Learning.
Appendix

Appendix AI. Excluded Ballot Initiatives

| Prop | Election         | Topic                          | Explanation                                                                 |
|------|------------------|--------------------------------|-----------------------------------------------------------------------------|
| 65   | 2016 General     | Legality and Taxation of Plastic Bags | Overlap. This initiative covers a topic similar to another initiative on the ballot that year and shares the same committees as the other initiative. Discerning spending for this initiative would not be exhaustive. |
| 67   | 2016 General     | Legality and Taxation of Plastic Bags | Overlap                                                                     |
| 58   | 2016 General     | English in Schools              | Though the voter guide lists this as an "Initiative Statue", the SOS website indicates the measure was placed on the ballot by the legislature, and Ballotpedia confirms. |
| 45   | 2014 General     | Health Insurance Regulations    | Irregularities in filing result in a signature spending value of zero. The committee failed to denote spending items with the proper code of PET, and attempting to discern the true amount of spending would not be exhaustive. |
| 40   | 2012 General     | Redistricting                   | The committee opposed to the campaign essentially withdrew opposition following a Supreme Court ruling, rendering the initiative rather irrelevant and eliminating any competition over the measure. Thus, it is unlikely to tell us anything about the relationship we're examining. |
| 20   | 2010 General     | Redistricting                   | Overlap                                                                     |
| 27   | 2010 General     | Redistricting                   | Overlap                                                                     |
| 91   | 2008 Feb Primary | Transit                         | The supporting committee deemed this proposition unnecessary after Prop 1A (a non-initiative passed in 2006). Given that there was no support for this measure, it is unlikely to tell us anything about the relationships we're examining. |
| 94   | 2008 Feb Primary | Indian Gaming Compacts          | Overlap                                                                     |
| 95   | 2008 Feb Primary | Indian Gaming Compacts          | Overlap                                                                     |
| 96   | 2008 Feb Primary | Indian Gaming Compacts          | Overlap                                                                     |
| 97   | 2008 Feb Primary | Indian Gaming Compacts          | Overlap                                                                     |
| 98   | 2008 June Primary| Eminent Domain                  | Overlap                                                                     |
| 99   | 2008 June Primary| Eminent Domain                  | Overlap                                                                     |
| 74   | 2005 Special     | Education                       | Overlap                                                                     |
| 76   | 2005 Special     | Education                       | Overlap                                                                     |
| 78   | 2005 Special     | Education                       | Overlap                                                                     |
| 79   | 2005 Special     | Education                       | Overlap                                                                     |