Polit Behav
DOI 10.1007/s11109-014-9282-4

ORIGINAL PAPER

Are Ballot Initiative Outcomes Influenced by the Campaigns of Independent Groups? A Precinct-Randomized Field Experiment Showing That They Are

Todd Rogers · Joel Middleton

Abstract Ballot initiatives are consequential and common, with total spending on initiative campaigns in the US rivaling that of Presidential campaigns. Past work using observational data has alternately found that initiative campaign spending cannot affect initiative outcomes, can increase the number of votes rejecting (but not approving) initiatives, or can affect outcomes in either direction. We report the first field experiment to evaluate an initiative advocacy campaign with precision. We find that campaigns can influence both rejection and approval of initiatives by changing how citizens vote, as opposed to by influencing turnout or ballot completion. Our experiment (involving around 18 % of Oregon households in 2008) studied a statewide mail program conducted by a Political Action Committee. Results further suggest that two initiatives would have passed if not for the advocacy campaign to reject them. We discuss implications for theories about direct democracy, campaign finance, and campaign effects.

Keywords Direct democracy · Campaign effects · Ballot measures · Field experiments

Ballot initiatives are a consequential, common, and costly aspect of American politics. This form of direct democracy can determine laws that have important economic and social consequences. Moreover, ballot initiatives may have

Electronic supplementary material The online version of this article (doi:10.1007/s11109-014-9282-4) contains supplementary material, which is available to authorized users.

T. Rogers (✉)
Harvard Kennedy School, Cambridge, MA, USA
e-mail: Todd_rogers@hks.harvard.edu

J. Middleton
University of California, Berkeley, Berkeley, CA, USA
e-mail: Joel.Middleton@gmail.com

Published online: 11 July 2014

implications for civic participation. Voters in states with ballot initiatives tend to be more engaged, exhibiting higher turnout (Tolbert and Smith 2005) and greater political awareness (Nicholson 2003). Because the stakes can be high, substantial resources are often devoted to advocating for and against them, and total spending on initiative campaigns can rival that of Presidential campaigns (Gerber 1999; Initiative and Referendum Institute 2000).

There is no clear consensus on whether ballot initiative campaigns affect initiative outcomes. Observational studies have come to a variety of conclusions about the effects of money, interest group involvement, or advertising on initiative outcomes. Some have concluded that initiative campaigns have no effect on initiative outcomes (Gerber 1999; Bowler and Donovan 1998), others suggest they do (De Figueiredo et al. 2011; Stratmann 2006; Smith and Tolbert 2004), and still others say that they are effective but only when advocating for rejection, not passage (Garrett and Gerber 2001; for a descriptive analysis see Broder 2000). Furthermore, this past research is unable to speak directly to whether ballot initiative campaigns influence outcomes by changing *how* citizens vote or by changing *whether* they vote.

In this manuscript, we offer robust evidence that campaigns can have profound effects on initiative outcomes in either direction, by changing *how* citizens vote. To assess the causal impact of initiative advocacy, we report a precinct-randomized experiment conducted during a statewide ballot initiative campaign. In addition to addressing questions about the impact of initiative campaigns, the research speaks to questions of campaign effectiveness, more generally (e.g., Gelman and King 1993; Holbrook 1996; Gerber 2004; Stratmann 2005).

We worked with a left-leaning Political Action Committee (PAC) that sought to sway the outcome of twelve ballot initiatives in the 2008 General Election in Oregon by sending one or two informative ballot guides to most households in the state. Out of twelve initiatives, the organization advocated for the passage of four and the rejection of eight. We find statistically and substantively significant effects on ten of the twelve initiatives. For two initiatives, it appears that effects were large enough to have altered the election outcome; the initiatives were rejected but would have passed were it not for the ballot guides.

Interestingly, the treatment had no effect on overall turnout or on rates of ballot roll-off (i.e., failure to cast a vote on given initiative despite submitting a ballot). Taken together, the pattern of results suggests that the treatment affected vote margins by altering preferences among those who would have voted on the initiatives anyway, but in the other direction.

This manuscript proceeds as follows. We begin with a review of past research on the effect of initiative campaigns. We then discuss the electoral context in which this study takes place and describe the advocacy program. We then discuss the precinct randomized design and the analysis plan of our field experiment involving 320,921 registered voters. Finally, we report the results of the experiment in terms of net vote margin, turnout, ballot roll-off, and effects on “up-ballot” political races. We conclude by discussing the implications of this research.

Theoretical Framing

With little exception, research on ballot initiative campaigns has been based on observational data. Almost no research on this topic has used field experiments to identify causal effects, despite the emergence of field experiments in political science to study voter mobilization effects (Gerber and Green 2000; Gerber and Green 2008), and other campaign activities.[1]

One exception was Keane and Nickerson (2013). Using a precinct-randomized field experiment, they found that a coalition advocating for the rejection of five initiatives decreased support for all five initiatives, but they note several limitations. First, their study was under powered. Power was limited by the number of precincts in the study and the low rates of reaching targeted individuals in the canvass programs they studied. Second, whereas we disentangle the differential effects of persuasion from turnout and ballot roll-off, their lack of statistical power made such analyses impossible. Third, the advocacy campaign they study advocated for rejecting all initiatives. In the present study the political action committee we worked with advocated for passage of four initiatives and rejection of eight others, allowing us to address the question of whether campaigns for and against initiatives are similarly effective (e.g.. Broder 2000; Garrett and Gerber 2001).

Studies using observational data and either regression frameworks or descriptive analyses have come to conflicting conclusions about the impact of initiative campaigns, offering three mutually exclusive claims about their impact. The first claim asserts that ballot initiative campaign spending has no effect on initiative outcomes. Gerber (1999), for example, argues that expenditures by organizations with minimal grassroots support relative to financial influence, called “economic groups”, have no effect on whether an initiative passes, and may even have a negative effect. Groups that have stronger grassroots support appear to have trivial effect as well, they report, though not quite as negative (see also Bowler and Donovan 1998; Garrett and Gerber 2001).

The second claim is that initiative campaigns can be potent when advocating for initiative rejection, but not passage. This is supported by studies showing that when one side has more money than the other, rejection efforts are more effective (Lowenstein 1982; Owens and Wade 1986; Magleby 1984). Bowler and Donovan (1998) find similar results. After arguing that campaigns can affect initiative outcomes when opposing an initiative, they write “money spent by proponents in this arena is largely wasted” [as quoted in Stratmann (2006), Bowler and Donovan (1998, p. 2)]. These authors posit an underlying mechanism that resembles the status quo bias (Samuelson and Zeckhauser 1988): people are biased towards preferring to leave things as they are.

[1] Research on the question of whether initiative campaigns matter could be viewed as a subset of work on the question of whether political campaigns matter, in general. Of course, this question has a long history, with some suggesting that campaigns can matter but only under some circumstances, and others arguing that campaigns have no effect (Jacobson 1978; Erikson and Palfrey 2000; Green and Krasno 1988). Recent evidence from randomized experiments or natural experiments suggests that campaigns can influence voter preferences (Gerber 2004; Arceneaux 2005; Huber and Arceneaux 2007), but possibly for only a fleeting period (Gerber et al. 2011).

The third claim says that initiative spending is effective regardless of whether advocating or opposing an initiative. For example, Stratmann (2006) reaches this conclusion using data on television advertising spending across several California media markets for initiatives from 2000 to 2004. Others make similar arguments, suggesting that interest group spending can influence both passage and rejection of ballot initiatives (De Figueiredo et al. 2011; Ellis 2002; Schrag 1998; Smith 1998; see also Broder 2000 for a descriptive journalistic discussion of this perspective).

Our study yields results that are consistent with this last claim, finding that initiative spending, both for and against, has an effect on initiative outcomes. Further, we do not find an asymmetry in a campaign's impact based on whether it advocates for or against an initiative. Our findings are at odds with the claim that only campaigns advocating for passage of an initiative are effective and also at odds with the claim that initiative campaigns are ineffective, in general.

Likewise, our results are at odds with the conclusion that groups without a widespread base of grassroots support, referred to as "economic groups," have no effects on ballot campaigns (Gerber 1999). The organization behind the advocacy program reported in this manuscript is an economic group because it does not have a broad base of individual financial supporters. In fact, it received 80 % of its 2008 budget from 7 donors who each gave between $600,000 and $5,300,000 (Oregon Secretary of State).

In addition to our contribution to debates around the impact of ballot initiative campaigns, this project speaks more broadly to the impact of political persuasion efforts, in general. Recent randomized experiments and natural experiments suggest that campaigns can influence voter preferences in state and local elections (Gerber 2004), in single issue ballot elections (Arceneaux 2005) and in presidential elections (Arceneaux 2007). The present study has several distinctive features that enhance its theoretical relevance. First, it looks at campaign effects in a particularly low information environment given that the vast majority of electoral attention was being devoted to the US Presidential and US Senate elections occurring on the same ballot. In so doing, it studies a particularly common political decision-making context (Converse 2000). Additionally, by focusing on non-partisan ballot initiatives the present study examines vote choice in a context in which heuristical partisan cues cannot be used to determine preferences. In this context voters are forced to deliberate over the issues at hand, or seek out alternative heuristics like ballot guides. Finally, the present study occurs in a vote-by-mail context, which is increasingly the method, timing, and context of contemporary United States elections.

Our study is also interesting in what it says about direct mail as a medium for political communication. Dozens of randomized field experiments have shown that get-out-the-vote mail tends to have trivial effects on turnout (see Green and Gerber 2008 for a meta-analysis of these studies). Based on those findings, one might speculate that direct mail would have similarly small effects on vote choices but our research shows that direct mail can substantially influence voter preferences.

Electoral Context

This experiment was conducted during the 2008 General Election in the state of Oregon. There were twelve initiatives on the ballot. The exact text that appeared on the ballot is reprinted in Table 1. We collaborated with an advocacy organization that already planned to send ballot guides to voters. The name of the organization (Our Oregon) was constructed for this specific election and so was unfamiliar to recipients of the ballot guides. This organization took a position on each initiative, though, as we discuss below, two of the initiatives (54 and 55) were uncontested and, therefore, deemphasized in the treatment materials.

The initiatives addressed a range of issues and the organization's positions were generally left-leaning. Measures 54, 55, 64 and 65 addressed election issues (eligibility for school board elections, redistricting, campaign financing, and general election nomination processes). Measures 56 and 59 addressed taxes (property taxes and tax deduction). Measures 57, 61 and 62 addressed crime (sentencing rules for a range of crimes and funding for law enforcement). Measures 58 and 60 addressed education issues (English language in schools and teacher promotion rules). Measure 63 addressed building permit rules. As we will discuss in greater detail below, the partner organization sent different ballot guides to different voters. The different guides offered recommendations on subsets of the twelve initiatives on the ballot. In addition to the exact wording of the initiatives, Table 1 also notes the partner organization's position on each initiative, the final statewide results for each initiative, and for which initiatives each ballot guide offered a recommended position.

This election also included several high-profile races, including an election for US President and US Senate. In the Presidential race, Oregonians voted for Barack Obama over John McCain by more than a 16 % point margin. This meant that not much presidential campaign expenditure occurred in the state, though there was substantial discussion and volunteer activity regarding that race. The Senate race was expected to be one of the closest in the country and was also the target of substantial campaign expenditures. According to Federal Election Commission (FEC) records, the Senate race in Oregon between incumbent Republican Gordon Smith and Democrat Jeff Merkley was the most expensive per capita in the country in 2008. The FEC reports that $20 million was spent by campaigns in Oregon for US Senate. An additional $10 million was spent by campaigns in Oregon for US House, and $11 million by registered PAC's in Oregon. FollowTheMoney.org documents that $54 million was spent by state-level candidates and initiatives.

Research Design

Ballot Guides

Our Oregon targeted households in which one or more voters had a predicted probability of turnout greater than 0.3. Turnout probability was estimated for households based on a model score provided by the commercial firm Catalist, LLC

Table 1 Ballot measures, our Oregon's positions and electoral outcomes

Measure	Description as printed on ballot	Position	Progressive	Moderate	Conservative	Education	Young voter	Yes votes	No votes	Result
54	Amends constitution: Standardizes voting eligibility for school board elections with other state and local elections	Pro	X					1,194,173	450,979	Pass
55	Amends constitution: Changes operative date of redistricting plans; allows affected legislators to finish term in original district	Pro	X					1,251,478	364,993	Pass
56	Amends constitution: Provides that May and November property tax elections are decided by majority of voters voting	Pro	X	X	X	X	X	959,118	735,500	Pass
57	Increases sentences for drug trafficking, theft against elderly and specified repeat property and identity theft crimes; requires addiction treatment for certain offenders	Pro	X	X	X		X	1,058,955	665,942	Pass
58	Prohibits teaching public school student in language other than English for more than two years	Anti	X	X	X	X	X	756,903	977,696	Fail
59	Creates an unlimited deduction for federal income taxes on individual taxpayers' Oregon income-tax returns	Anti	X	X	X	X	X	615,894	1,084,422	Fail
60	Teacher "classroom performance," not seniority, determines pay raises; "most qualified" teachers retained, regardless of seniority	Anti	X	X		X	X	673,296	1,070,682	Fail
61	Creates mandatory minimum prison sentences for certain theft, identity theft, forgery, drug, and burglary crimes	Anti	X	X	X		X	848,901	887,165	Fail

Table 1 continued

Measure	Description as printed on ballot	Position	Progressive	Moderate	Conservative	Education	Young voter	Yes votes	No votes	Result
62	Amends constitution: Allocates 15 % of lottery proceeds to public safety fund for crime prevention, investigation, prosecution	Anti	X	X	X	X	X	674,428	1,035,756	Fail
63	Exempts specified property owners from building permit requirements for improvements valued at/under 35,000 dollars	Anti	X	X	X		X	784,376	928,721	Fail
64	Penalizes person, entity for using funds collected with "public resource" (defined) for "political purpose" (defined)	Anti	X	X		X	X	835,563	854,327	Fail
65	Changes general election nomination processes for major/minor party, independent candidates for most partisan offices	Anti		X	X			553,640	1,070,580	Fail

(see Ansolabehere and Hersh 2010). Based on this model score, 78.4 % of all 1.5 million households were targeted, or, equivalently, an estimated 85 % of households with at least one registered voter.

The advocacy organization mailed each targeted household at least one of five possible ballot guides. The guides were mailed less than one month before the election, and the first of the guides arrived around the same time when ballots were mailed to all households (Oregon is an entirely vote-by-mail state). The guides can be divided into two categories, primary and secondary. Primary guides were sent to most targets, while secondary were sent to a smaller, partially overlapping subset.

Table 2 provides the distribution of guides and Supplement 1 shows the actual ballot guides themselves. There were three types of primary guides, targeted to match the predicted ideology of the household as predicted based on data from the Catalist voter file (progressive, moderate or conservative). The progressive guide was four pages long and listed on its first page twelve recognizable progressive sponsor organizations (labor unions, environmental organizations, pro-choice organizations, etc.) under the heading "Recommendations from Organizations that you Trust." In addition to listing sponsors, the guide described that Our Oregon was a "non-profit, non-partisan organization dedicated to promoting economic and tax fairness." Among target households in treatment precincts, 13.8 % were considered progressive and were sent the progressive ballot guide.

The moderate guide was eight pages and reported that it was "Provided as a Public Service by Our Oregon." It did not list any sponsors and described that Our Oregon was a "non-profit, non-partisan organization dedicated to promoting economic and tax fairness." The guide discussed the pros and cons of each ballot initiative, attempting to appear even-handed, then made a recommendation. Households predicted to be politically moderate (51.0 %) received this moderate ballot guide.

The conservative guide was eight pages and it first page included in large bold font "Tax Fairness and Accountability." While discussing each ballot initiative this guide emphasized government waste and taxation, and said that the guide was

Table 2 Relative frequencies of combinations of primary and secondary guides

Primary guides	Secondary guides				Total
	No secondary guide	Education	Young voter	Both	
Progressive	10.4 %	3.3 %	[a]	[a]	13.8 %
Moderate	35.3 %	15.4 %	[a]	[a]	51.0 %
Conservative	25.9 %	5.2 %	[a]	[a]	31.4 %
No primary guide	[a]	[a]	3.6 %	[a]	3.8 %
More than one	[a]	[a]	[a]	[a]	[a]
Total	71.7 %	24.1 %	4.1 %	[a]	100.0 %

Percentages represent relative frequencies of guide combinations over all targeted households in experimental precincts. Percentages lower than 0.3 % are omitted from the table for clarity but are included in the marginal percentages. In total only 0.9 % of households fell into omitted cells

[a] Fewer than 0.3 % of targeted households were sent this combination of guides

"Provided as a Public Service by Our Oregon." It did not list any sponsors and it did not describe anything else about the organization Our Oregon. Households predicted to be conservative (31.4 %) received the conservative guide. In total, 96.2 % of the households that were sent any ballot guide from the advocacy organization received one of these three primary guides.

The other two types of ballot guides we refer to as secondary. The first of these was the education guide. It was sent to households believed to contain children. The education guide was two pages and listed five sponsor organizations associated with education and children (teacher unions, PTA, a children's advocacy organization) under the heading "Facts and Recommendations from Oregon's Leading Advocates for Schools and Children." The education guide did not provide a description of what Our Oregon was or its mission.

The second of these ballot guides targeted households believed to be headed by voters under 35 years old, many of which did not receive a primary guide. The "youth" guide was two pages and showed the logos of eight different Oregon news organizations under the heading "Newspaper Recommendations." The youth guide did not provide a description of what Our Oregon was or its mission. Of all targeted households in treatment precincts, 71.7 % received no secondary guide, 24.1 % received the education guide in addition to a primary guide, and 4.1 % received the youth guide.

The average number of ballot guides sent to a target household was 1.2. A vast majority (97.0 %) of the households that received more than one guide did so because they received the education-oriented guide in addition to a primary guide. In other words, households identified as having children make up nearly all of those who received more than one guide. A negligible amount of duplication occurred due to the vagaries of determining households based on imperfect address information and differences in the ways that "households" can be operationally defined based on this information.

The five ballot guides advocated for positions on different subsets of the 12 ballot measures. In Table 1 we show which measures were addressed in each of the ballot guides.

Randomization

We chose to randomize at the level of the precinct because precincts are the smallest unit at which election outcomes are officially reported. In addition to being publically available, official outcome records have the virtues of being free of sampling error, nonresponse bias and errors of self-report (see Keane and Nickerson 2013; Gerber 2004; Gerber et al. 2011; Arceneaux 2005). The same cannot be said of post-election surveys that are sometimes used to obtain dependent variables in field experiments. As such we assessed that the benefits of randomizing at the precinct level in terms of data quality outweighed the potential for increased statistical precision resulting from a study conducted using individual-level randomization with post-election survey responses used as the dependent variable.

As a condition of their cooperation, Our Oregon constrained the total number of voters in our untreated control condition to 5 % of the Oregon electorate. Given this constraint, we restricted the experimental universe to only the moderately- and small-sized precincts. This approach maximized the total number of control precincts while maintaining the fixed number of voters in the control, thus improving statistical power. The alternative design, randomizing Oregon precincts of all sizes, would have left us with 64 % fewer precincts, and yielded standard errors approximately 20 % larger.

We further excluded precincts for which the reported number of registered voters varied by more than 20 % between two different sources of registered voters per precinct, the National Campaign for an Effective Congress data file and the file from the commercial firm Catalist data file. We reasoned that inconsistent voter counts might be an indication of changes to precinct boundaries or erroneous records, which might have resulted in imperfect treatment implementation.

Once we identified our study universe of 700 small and medium precincts, we randomly assigned 200 of them to the control group using simple random assignment. The total number of voters in the study was 320,921, of which 91,540 were in the control group. According to the Oregon Secretary of State our experimental universe included approximately 18 % of the Oregon electorate. As a formal randomization check, we regressed all covariates in our files on the treatment indicator. None of the covariates were statistically significant, and, likewise, the omibus F-test suggests the treatment and control precincts were balanced ($p = 0.8721$).[2] "Appendix" discusses randomization checks at length, and the slight, but not statistically significant difference in partisanship between treatment and control groups. In short, this small imbalance does not seem to affect results, but future studies may want to block on partisanship in particular to achieve better balance on this covariate in particular since the outcome for a precinct is certainly related to partisanship.

Dependent Variables

Our analysis examines four dependent variables. All are reported on public records of precinct-level results made available by the Oregon Secretary of State. First, we estimate the net change in vote margin for each ballot initiative due to treatment. This is our central outcome of interest because the ballot guides focused on supporting or opposing the initiatives. Second, we estimate the net change in voter turnout due to the treatment. Turnout could be affected if the treatment motivated citizens to cast a ballot, or, alternatively, discouraged citizens from voting. Third, we estimate the net change in the number of voters who "roll-off" or abstain from voting on the ballot measures, even though they cast a ballot to vote on other items. Voters may roll-off if they are uninformed about the initiatives. Because the treatment provided information and recommendations, voters who received the guides may have felt more informed and therefore willing to vote on the measures. Fourth, and finally, we consider whether treatment altered support for up-ballot candidates.

[2] Logistic regression yields similar results indicating balance.

Estimation

Our analysis uses regression to estimate the average change in precinct-level vote margin due to treatment. Precinct-level vote margin is defined as the number of citizens voting for a ballot measure minus the number of citizens voting against that ballot measure, divided by the total number of ballots cast in the precinct. A vote margin between 0 and 1 indicates a majority voted for the measure and margin between 0 and -1 indicates a majority voted against. We used three model specifications. In the first, vote margin was regressed on the treatment indicator only. Model 1, can be written

$$y_j = \beta + t_j\beta_t + \varepsilon_j, \tag{1}$$

where y_j is vote margin, t_j is the treatment indicator, and ε_j is an error term. The subscript j indexes over precincts.

Model 2 controls for partisanship. It can be written

$$y_j = \beta + t_j\beta_t + \partial_j\beta_\partial + \omega_j\beta_\omega + \varepsilon_j, \tag{2}$$

where ∂_j is a predictor of Democratic voting rates in the district (also in percentage terms),[3] and ω_j indicates a missing value of ∂_j.

Model 3 controls for partisanship and vote history from the 2002, 2004 and 2006 elections, including missing dummy variables. It can be written

$$y_j = \beta + t_j\beta_t + \partial_j\beta_\partial + \omega_j\beta_\omega + \sum_k v_k\beta_{v_k} + \sum_k \mu_k\beta_{\mu_k} + \varepsilon_j, \tag{3}$$

where v_k is an indicator of participation in election in year k, where $k = \{2002, 2004, 2006\}$, and μ_k is an indicator of missing v_k.

We estimate the models using ordinary least squares regression. Since ε_j may be homoscedastic due to differing precinct sizes we use "robust" standard errors,[4] though this does not affect the findings.

Per a suggestion from a helpful reviewer, we control family-wise error at $\alpha = 0.05$ (for each model in each table) by using the Holm–Bonferroni method.[5] The method was applied to groups of coefficients within each column in each table. So for example, for model (1) in Table 3, each of 12 coefficients is grouped into a

[3] The Democratic Performance Index (DPI) is a synthetic variable created by the National Center for an Effective Congress, designed to be broadly predictive of Democratic voting across political races. The index is designed to help political organizations make operational decisions about voter outreach campaigns. We use it as a covariate to improve the efficiency of our estimator, but give no causal interpretation to its coefficient. Since we use the DPI created *before* the 2008 election for use by political organizations during the 2008 campaign, we need not worry that this covariate is partially determined by the treatment and, hence, correlated with the treatment indicator.

[4] By "robust" standard errors, we are referring to the so-called Huber Sandwich Estimator (c.f. Freedman 2006).

[5] The Holm–Bonferroni method sorts the p values from a family of hypothesis tests from smallest to largest. The procedure finds the first p value in the list that satisfies $p_{(k)} > \frac{\alpha}{m+1-k}$, where $p_{(k)}$ is the kth ordered p value, and m is the number of tests in the family. Then it rejects all hypotheses up to $k - 1$. The method is less powerful than the Hochberg procedure, but does not require the hypotheses to be independent. In this situation the hypotheses are clearly not independent.

Table 3 Estimated effects on ballot measure vote margin

Measure	Position	(1)	(2)	(3)	Actual statewide vote margin (%)
54	Pro	0.8 (1.3)	0.6 (1.2)	0.8 (1.2)	41.0
55	Pro	−0.4 (1.2)	−0.4 (1.2)	−0.5 (1)	48.9
56	Pro	4.2 (1.5)[a]	3.3 (1.2)[a]	3.2 (1.2)[a]	12.3
57	Pro	4.6 (1.2)[a]	4.2 (1.2)[a]	4.3 (1.2)[a]	21.7
58	Anti	−5.9 (1.8)[a]	−4.9 (1.3)[a]	−5.4 (1.3)[a]	−12.2
59	Anti	−4.9 (1.5)[a]	−4.1 (1.3)[a]	−4.6 (1.2)[a]	−25.9
60	Anti	−4.1 (1.5)[a]	−3.3 (1.2)[a]	−3.5 (1.2)[a]	−21.9
61[b]	Anti	−6.0 (1.7)[a]	−5 (1.3)[a]	−5.4 (1.3)[a]	−2.1[b]
62	Anti	−6.0 (1.5)[a]	−5.1 (1.4)[a]	−5.3 (1.4)[a]	−19.9
63	Anti	−3.6 (2.3)	−2.4 (1.8)	−3.4 (1.7)	−8.0
64[b]	Anti	−5.9 (1.6)[a]	−5.1 (1.4)[a]	−5.5 (1.3)[a]	−1.0[b]
65	Anti	−4.1 (1.0)[a]	−4.5 (1.0)[a]	−4.4 (1.0)[a]	−28.5

[a] Signifies significance using the Holm–Bonferroni adjustment to control family wise error at $\propto = 0.05$. Like the better-known Bonferroni method, this method insures that the probability of incorrectly rejecting one or more hypotheses is less than 0.05, but it is more powerful than the Bonferroni adjustment. For descriptions of the initiatives, see Table 1

[b] Signifies the statewide vote margin is smaller than the effect estimate

family. The Holm–Bonferroni method insures that, if the null hypothesis were true for each of the 12 outcomes, the likelihood of rejecting one or more hypotheses would be no greater than 0.05. The Holm–Bonferroni method is more powerful than the better-known Bonferroni method but also achieves the goal of limiting the family-wise error rate.

We also reran the analysis to see if the effects were moderated by the partisanship of the precinct, presented in the “Appendix” due to a failure to find any statistically significant results. To do this, precincts were divided into three equally sized groupings—liberal, moderate, and conservative—by rank ordering them using the Democratic Performance Index. An indicator for partisanship grouping was then interacted with the treatment indicator. An omnibus test of significance was used to test for a difference in the treatment effect for the three different precinct types and the Holm–Bonferroni procedure was used to determine whether to reject.

Results

This section presents the estimates for the effect of the treatment on ballot measure support, turnout, roll-off and candidate support.

Ballot Measure Vote Margin

Table 3 shows that the treatment had sizable effects on vote margin for nine of the twelve ballot initiatives (56–64) in the intended directions. Among the nine, effects on vote margin ranged from 2.5 % points to 6 % points. It is notable that the treatment affected nine different electoral outcomes, not just one.

There were three initiatives that did not achieve statistical significance. Estimates for one of the three, measure 63, trend in the expected direction but fall short of achieving significance. Estimates for the other two initiatives unaffected by the treatment (54 and 55) were very close to zero. Two unusual features of these initiatives may explain the lack of impact on their election outcomes. First, as described above, the advocacy organization sent treatment households up to two out of five ballot guides. Only one of the five ballot guides made reference to Initiatives 54 or 55 while each of the other ten were referred to in multiple guides; only 20 % of targeted households received a ballot guide that referred to either of these two initiatives. Second, these two initiatives were unique in that they were essentially uncontested in terms of campaign activity and had lopsided vote margins, passing statewide with a 41.0 % point and 48.9 % point vote margin, respectively (i.e., 70.5 and 74.4 % of votes cast for each initiative were cast in favor of the initiatives). This suggests that the vote margin may have reached a ceiling.

Finally, Table 3 shows that the treatment effect for Initiatives 61 and 64 was greater than the statewide vote margin for the initiatives. Initiative 61 (about mandatory drug sentencing) was defeated by 2.2 a vote margin of percentage points statewide, and the treatment decreased support for it in our experiment universe by about 5 or 6 % points in the average precinct; Initiative 64 (about how political funds can be spent) was defeated by a vote margin of 1.0 % point statewide, and the treatment decreased support for it in our experimental universe by about 5 or 6 % points in the average precinct. Were we to conjecture that the treatment effect in our experimental universe generalizes to the larger Oregon precincts that were excluded from this study (but in which the advocacy organization conducted the exact same program), then we can say that if not for the advocacy campaign the electoral outcomes of these two initiatives would have been otherwise.[6] Even if the average effect outside the experimental universe were a fraction of these estimates, the election outcomes would still be attributable to the campaign.

Ballot Roll-Off Rates

Table 4 reports how treatment affected the rates of ballot "roll-off" for each of the initiatives. A voter exhibits roll-off when s/he votes in high profile races, such as President or US Senate, but fails to vote on lower salience items such as ballot initiatives. We calculate roll-off for a given precinct by subtracting the total number of votes cast for a given initiative from the total number of votes cast in the Presidential election. Results suggest that roll-off was not affected for any of the twelve initiatives, suggesting that the mailings did not have a substantive effect on increasing ballot completion. For Initiative 65 there is some evidence that roll-off

[6] We cannot demonstrate, statistically, that the treatment effect in our experimental universe would generalize to the large precincts outside of our experiment universe. That said, we find no relationship in our experimental universe between precinct size and treatment effect. This suggests that the advocacy organization's treatment may have resulted in similar average treatment effects in Oregon's large precincts as we found in our experiment universe of small- and moderately-sized precincts.

Table 4 Estimated effects on ballot roll-off rates

Measure	Position	(1)	(2)	(3)
54	Pro	0.4 (0.3)	0.4 (0.3)	0.3 (0.3)
55	Pro	0.3 (0.4)	0.2 (0.4)	0.2 (0.4)
56	Pro	−0.2 (0.4)	−0.3 (0.4)	−0.4 (0.4)
57	Pro	−0.1 (0.3)	−0.1 (0.3)	−0.2 (0.3)
58	Anti	0.0 (0.2)	0.0 (0.2)	0.0 (0.2)
59	Anti	0.1 (0.3)	0.0 (0.3)	−0.1 (0.3)
60	Anti	0.4 (0.4)	0.3 (0.4)	0.4 (0.3)
61	Anti	0.2 (0.4)	0.2 (0.4)	0.2 (0.4)
62	Anti	−0.1 (0.3)	−0.2 (0.3)	−0.3 (0.3)
63	Anti	0.3 (0.3)	0.2 (0.3)	0.1 (0.3)
64	Anti	0.2 (0.3)	0.1 (0.3)	0.1 (0.3)
65	Anti	−0.7 (0.4)	−0.7 (0.4)	−0.7 (0.3)

Results presented in percentage points. Estimates represent the total number of votes cast for the race with the greatest number of total votes cast (which was always the Presidential election) minus the total number of votes cast for a given initiative in a precinct minus. "Pro" means the communications advocated for the passage of the ballot initiative, while "Anti" means the communications advocated for the passage of the ballot initiative. For descriptions of the initiatives, see Table 1

There are no significant results in this table using the Holm–Bonferroni adjustment to control family wise error at $\propto = 0.05$. For each of the three model specifications, a "family" is the grouping of 12 ballot measures. Like the better-known Bonferroni method, this method insures that the probability of incorrectly rejecting one or more hypotheses is less than 0.05, but it is more powerful than the Bonferroni adjustment

was reduced, but the coefficient is not significant after the Holm–Bonferroni adjustment.

Turnout

Table 5 shows the treatment had no meaningful effect on net turnout. However, since the results are reported at the cluster level (the precincts) one might suggest that the treatment results are averages that mask an increase in turnout among some voters (e.g., Democrats), and a decrease in turnout in equal measure among others (e.g., Republicans). The fact that our treatment did not affect net turnout is consistent with the interpretation that people do not fail to vote because they are uninformed since our treatments were informative about the ballot initiatives.

Effects on Candidate Races

Table 6 presents the results for candidate races. Results are not statistically significant for any of the races. For model 1, however, some of the estimates are rather large and would have achieved statistical significance without the Holm–Bonferroni adjustment. However, the covariate adjusted estimates are greatly diminished, and the results are not replicated using a nonparametric estimation

Table 5 Estimated effects on turnout

	(1)	(2)	(3)
Turnout	−0.1 (0.6)	−0.2 (0.6)	0.4 (0.5)

Results presented in percentage points

Table 6 Estimated effects on candidate races

Number of precincts	Race	(1)	(2)	(3)
696	President	1.8 (2.6)	0.1 (1.5)	0.4 (1.5)
696	Senate	4.2 (2.5)	2.7 (1.4)	2.8 (1.4)
512	House	9.4 (4.8)	3 (3.6)	3.4 (3.4)
696	Secretary of state	2.2 (2.3)	0.4 (1.4)	1.1 (1.4)
696	State treasurer	2.2 (2.1)	0.9 (1.4)	1.1 (1.3)
216	State senate	6.4 (4.5)	1.0 (3.1)	3.4 (3.3)
464	State house	1 (3.3)	−0.3 (2.4)	−0.8 (2.3)

Results presented in percentage points. Estimates represent the total number of votes cast in the treatment precincts for the Democratic candidate minus the total number of votes cast in control precincts for the Republican candidate

There are no significant results in this table using the Holm–Bonferroni adjustment to control family wise error at $\propto = 0.05$. For each of the three model specifications, a "family" is the grouping of 7 races. Like the better-known Bonferroni method, this method insures that the probability of incorrectly rejecting one or more hypotheses is less than 0.05, but it is more powerful than the Bonferroni adjustment

approach, available upon request. These directional results are consistent with the kinds of up-ballot effects one might expect given the generally left-leaning positions advocated for in the mailings, but because of the lack of significance and lack of robustness to specification, the results do not support an interpretation that ballot guides affected up-ballot races.

General Discussion

We present a field experiment examining the impact of an independent organization advocating for and against a suite of ballot initiatives. We find that these organizations can have a consequential and significant effect on the passage and rejection of ballot initiatives. This finding speaks directly to an unsettled debate about the impact these organizations are capable of having on initiative outcomes by providing precise causal evidence in support of the view that independent organizations can affect the outcomes of ballot initiative (Stratmann 2006; Ellis 2002; Smith 1998). The results we report are in contrast with findings suggesting that campaigns do not have an effect on initiative outcomes (Gerber 1999; Bowler and Donovan 1998) and work suggesting that only campaigns *against* initiatives are effective (Broder 2000; Garrett and Gerber 2001). We find that the independent organization's campaign affected the win margin for these initiatives but did not influence turnout[7] or ballot roll-off. This pattern of results suggests that the campaign affected vote margin by altering ballot preferences among those who would have cast their vote on the initiatives anyway, but in the opposite direction.

We can calculate cost per additional vote added to vote margin we use the following equation:

$$(\mathrm{c} \times p)/(\mathrm{e} \times n)$$

The term c is the treatment cost per targeted household, \$0.90 (1.2 guides per household × \$0.75 per guide). The term p is the proportion of households with a registered voter that were targeted (85 %). The term e is the average effect on vote margin for each of the twelve advocated for initiatives. Using the covariate adjusted estimates from Table 3 the average impact on each initiative's vote margin is 3.5 % points. The term n is the number of initiatives for which the treatment advocated. There were twelve initiatives advocated for by this treatment. Calculated this way, the treatment yielded an additional vote in the favored direction at \$1.84.[8]

This is a low cost, in part due to the efficiency derived from advocating multiple initiatives in one mailing. The result is still surprising, however, when you consider that direct mail used in get-out-the-vote efforts can cost around \$400 to \$500 per increased voter (see Green and Gerber 2008; particularly when the mail does not leverage psychological elements shown to motivate voter participation: Rogers et al. 2012). Another persuasion campaign conducted in 2008 in Oregon (Rogers and Nickerson 2013) targeted self-identified pro-choice voters and tried to correct misinformation about the incumbent Senator's stance on abortion. That campaign found that it cost \$40 per additional vote added to the vote margin in the favored direction.

Why might the impact of this unknown independent advocacy organization been so great? We speculate that voters may be especially receptive to the style of the ballot guides: factual, seemingly objective, and providing, for four out of five versions, ostensibly non-political summaries of multiple ballot initiatives. One might still find the effect magnitudes surprising in light of the hotly contested political context in Oregon in 2008. However, inundated with political communications for a variety of contests, recipients may have responded positively to the systematic and clear vote recommendations provided in a compact format. Indeed, theories of bounded attention (Gabaix et al. 2006) suggest that in noisy election environments voters may place a premium on these kinds of efficient communications—so long as they are perceived to be credible. Future research should assess whether treatments studied in this experiment are especially effective in hotly contested elections as opposed to in less cluttered election environments.

[7] Other research has suggested that ballot initiatives have a greater effect on turnout in lower-turnout elections (Childers and Binder 2012). This suggests that the absence of a finding for turnout may be a result of the relatively high turnout election, and that should be taken into account when considering the external validity of these results.

[8] Factoring in the uncertainty of our estimates, we calculate a 95 % confidence interval of (1.07, 6.44).

Another reason for the ballot guide effectiveness may be that they were delivered directly to voters' homes, since all ballots are cast by mail in Oregon. This all vote-by-mail election may have created the conditions whereby voters may have completed their ballots at home with the ballot guide present to inform their decisions. Future research should explore whether vote-by-mail moderates the effectiveness of persuasive ballot guides. This might be testable by studying ballot guide effectiveness in precincts with total populations that are just above and below the number where vote-by-mail becomes mandatory (see Meredith and Malhotra 2008).

Another reason these ballot guides might have been especially effective could be that the independent advocacy organization was specifically unfamiliar. Provocative new research suggests that unfamiliar advocacy groups can sometimes be even more effective advocates than familiar ones (Weber et al. 2011). In the present experiment, the two ballot guides that were most widely distributed included no sponsors of any sort—the moderate guide was sent to 51 % of treated households and the conservative guide was sent to 31 % of households. Perhaps this unfamiliarity enhanced the effectiveness of the guides.

We report some limited evidence of the impact of the ballot guides on up-ballot support for the left-leaning candidates (i.e. the Democrat candidates). The results are not statistically significant after the Holm–Bonferroni adjustment and are also not robust to specification, but they are provocative. Support for up-ballot candidates could be altered if treatments that shift voters to have more left-leaning preferences on ballot measures also prime issues or preferences that change the way the up-ballot candidate choice is construed (Iyengar and Kinder 1987; Krosnick and Kinder 1990).

We might have expected significant effects on up-ballot vote choice in light of research suggesting ballot initiatives make initiative-relevant issues more important factors in up-ballot candidate races (Smith and Tolbert 2010). However, there are two reasons the present findings do not speak directly to that research. First, there were twelve initiatives on the ballot addressing issues ranging from taxes (Measures 56 and 59) to crime (Measures 57, 61 and 62). The diversity of issues raised by these initiatives might have reduced, or even neutralized, the impact of any one initiative. The second reason is the same as that provided above regarding the finding that the ballot guides had no impact on turnout. The argument advanced regarding the impact of ballot initiatives on up-ballot races is about the *introduction* of a ballot initiative into an election (and the ensuing initiative campaign), while the current study focuses on the impact of advocacy efforts after the initiatives had already been introduced into the election in that state.

Our treatment did not increase turnout. Given that the treatment advocated for multiple issues one might have expected it would be especially potent at increasing turnout (and reducing rolloff) since it offered targets a range of potentially motivating issues. The lack of a turnout effect is not consistent with past work suggesting that the presence of ballot initiatives can increase turnout, in general (Tolbert et al. 2001) or in Midterm elections but not Presidential elections (Smith 2001a, b). The absence of a turnout effect in the present experiment only tangentially speaks to that work, though. This is because both lines of that research suggest that the increases in turnout come from the *introduction* of a ballot initiative into an election (and the ensuing initiative campaign), while the current study focuses on changing voters' preferences on ballot initiatives that had already been introduced into the election in that state. Whatever turnout affect arises from the introduction of a given initiative in a state should already have been incorporated into voters' likelihoods of voting before our treatment was administered.

The organization that conducted this mail program was financially supported by a small group of donors. Therefore, our findings are inconsistent with work suggesting that organizations without broad financial support of grassroots donors, so-called economic organizations, have trivial effects on initiative outcomes (Gerber 1999). That economic organizations can influence initiative outcomes has normative implications, particularly because ballot initiatives are often described as giving *the people*, not interest groups, direct democratic power to decide important policy issues. As such, our findings may provide fodder for those who wish to regulate campaign funding.

Finally, the present research involved collaborating with a pre-existing Political Action Committee and studying its multi-issue electoral program. This approach has its strengths and weaknesses. Among the weaknesses is that we cannot isolate which mailing or message caused the overall treatment effects, which individuals were particularly affected by each treatment, and what underlying decision-making processes led to these rather large changes in vote choice. These weaknesses are endemic to field experiments, though. But field experiments offer powerful strengths. Among them is a high degree of external validity in terms of both the dependent variables studied (i.e., actual vote choice and turnout in a consequential election), as well as the independent variable (i.e., an actual political program that targeted the vast majority of voters in a US state). This approach of conducting a field experiment in collaboration with a political organization's voter communication program allows us to make strong causal claims about the impact of the treatment; and, in aggregate, it enables us to estimate that this treatment changed the election outcomes more than previous theory would have predicted.

Acknowledgments We thank Our Oregon for collaborating with us. We thank Josh Berezin and Kevin Looper for cooperation and assistance. We thank Analyst Institute and Catalist LLC for providing data. We thank Don Green, Max Bazerman, David Nickerson, Kevin Collins, Jennifer Green, and the Analyst Group for providing feedback. We thank Julia Kamin, Carly Robinson, and John Ternovski for help with analyses and editing.

Appendix

Table 7 summarizes available covariates for the treatment and control precincts. The reader will see that the treatment and control precincts exhibit good balance. Figure 1 plots on the x-axis the number of registered voters (as reported after the election) against on the y-axis their Democratic Performance Index (a synthetic

Table 7 Descriptive statistics for treatment and control precincts

	Treatment	Control	Excluded
2002 Turnout	69.4 %	70.1 %	68.9 %
	(0.5)	(0.9)	(0.3)
2004 Turnout	85.8 %	87.0 %	86.3 %
	(0.4)	(0.5)	(0.2)
2006 Turnout	70.8 %	71.8 %	70.5 %
	(0.5)	(0.7)	(0.2)
2006 Democratic performance index	43.8 %	42.9 %	49.9 %
	(0.7)	(1.1)	(0.5)
2002 Registration	443	443	1,908
	(15)	(26)	(37)
2004 Registration	488	481	2,192
	(17)	(28)	(44)
2006 Registration	456	454	2,020
	(16)	(28)	(39)
2008 Registration	467	458	2,173
	(14)	(21)	(43)

Background characteristics of control precincts, treatment precincts, and precincts excluded from the experiment. Standard errors are in parentheses. The 2006 Democratic Performance Index is a synthetic variable created by the NCEC to characterize the overall rate of voting for Democrats in 2006 political races

variable produced in advance of the 2008 election that summarizes Democratic voting in previous elections from National Campaign for an Effective Congress). Squares represent Oregon precincts excluded from the study, triangles represent those included. The dividing line between the larger and smaller precincts can be seen clearly with included precincts (triangles) to the left and excluded precincts (squares) to the right. There are exceptions. The figure shows seven large precincts included in the study, represented by stray green triangles on the right. It also shows a number of smaller precincts excluded from the study, represented by blue squares on the left. We explain these two anomalous precinct types in turn.

The unexpectedly large precincts in the experimental universe, triangles on the right, resulted from discrepancies in the two sources we used to assess the number of voters per precinct—National Campaign for an Effective Congress and Catalist, LLC—and the administrative records that provided the outcome measures that we gathered after the election and matched. The values on the x-axis in Fig. 1 come from these post-election administrative records. Less than 1 % of the precincts in the experimental universe (seven) show this discrepancy. What is most striking about Fig. 1 is the sharp discontinuity between included and excluded precincts, providing confidence in our randomization, and our post-election merger of randomization and outcome files. Nonetheless, we reran the analysis described below omitting the seven strays as a robustness check and results were not altered.

Figure 1 also depicts a number of small precincts excluded from the experiment, squares to the left of the dividing line. These precincts were excluded for the reason previously described: the number of registrants reportedly in the precinct differed across the two pre-election sources by more than 20 %.

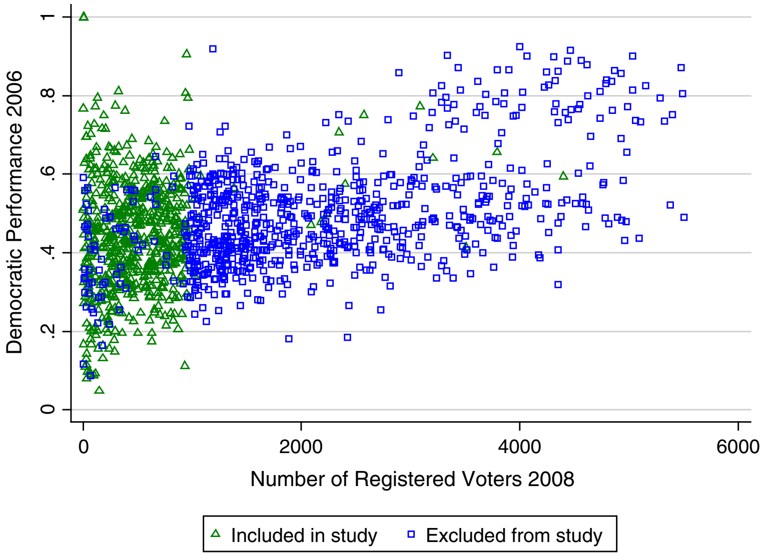

Fig. 1 Democratic performance by 2008 registration

Our choice of experiment universe has little bearing on the internal validity of our study. Moreover, no experiment can represent all precincts, in all elections, and for all time. In that respect, this field experiments is not unique in being limited in its geographic, temporal and demographic extent. Nonetheless, we must account for the sample restrictions when considering external validity, particularly when speculating on the effects of the ballot guide campaign on voters in all of Oregon.

Table 7 presents summary statistics for both included and excluded precincts. The size difference between excluded and included precincts is evident. And while turnout rates are not dramatically different, excluded precincts are more Democratic as measured by the Democratic Performance Index. Figure 1 illustrates why this is true, there is a positive association between precinct size and Democratic voting. The association is no surprise; urban environments tend to have larger precincts and also greater numbers of Democratic-leaning voters. As an illustrative anecdote, consider that the most strongly Democratic county in Oregon, Multnomah County (which contains Oregon's largest city, Portland), also had the highest rate of precinct exclusion in this study due to sizes—90 % excluded compared to 56 % statewide.

Post-election, four (0.6 %) of the 700 precincts included in the study (one control precinct and three treatment precincts) could not be matched to with the reported election returns. Again, we suspect that inevitable imperfections in available records are responsible for this, but by any standard this is a small loss-to-follow-up rate. To address the slight imbalance in the partisanship of treatment and control precincts, a stratified analysis was carried out, creating

Table 8 Democratic performance Index 2006 by partisanship strata

	Treatment	Control
Overall	43.8 %	42.9 %
	(0.7)	(1.1)
Democratic	59.1 %	58.2 %
	(3.9)	(6.2)
Independent	43.2 %	43.2 %
	(3.8)	(6.9)
Republican	28.6 %	28.5 %
	(3.7)	(5.5)

three strata using the Democratic Performance Index; for shorthand. We will refer to these as Democratic, Independent and Republican precincts. The balance in the strata is much better for the Republican and Independent strata, as evidenced in Table 8. The Democratic stratum still shows evidence of imbalance. This imbalance does not concern us, however, because dropping the Democratic stratum from the analysis, estimates become larger, not smaller, suggesting that this residual imbalance in the Democratic stratum cannot explain away the treatment effects.

Table 9 presents coefficients from regression models run with fixed effects for strata, along with interaction terms between strata and treatment assignment. For each row of the table, a p value from a test of the difference between the three effect estimates. (No "main effect" was included in the model so that three interaction terms, one for each of the three strata, could be specified and stratum specific treatment effects computed directly.) There are no significant differences between coefficients for any of the ballot measures regardless of the additional covariates included; in other words, we fail to find evidence of treatment effect heterogeneity. (Again the Holm–Bonferroni adjustment was used to control for family-wise error rate.) Combining these strata-specific estimates into "main effects" by averaging the three returns estimates that look very similar to the original estimates in the main body of the paper. Tables 10, 11 and 12 provide analogous results with fixed effects for strata for roll-off, candidate votes, and turnout, respectively. There are no significant evidence of treatment effect heterogeneity and "main effects" are generally in line with original estimates in the main body of the paper.

Table 9 Ballot measures

	Model 1					Model 2					Model 3				
	Dem	Ind	Rep	p (diff)	Overall effect	Dem	Ind	Rep	p (diff)	Overall effect	Dem	Ind	Rep	p (diff)	Overall effect
Measure 54	−0.5 (2.2)	−0.1 (2.4)	1.6 (2.2)	0.766	0.3 (1.3)	−0.8 (2.2)	−0.1 (2.3)	1.6 (2.2)	0.726	0.2 (1.3)	0.4 (2.1)	0.3 (2.2)	0.8 (2)	0.984	0.5 (1.2)
MeasuRe 55	−0.2 (2.1)	−1.5 (2.2)	−0.4 (2)	0.910	−0.7 (1.2)	−0.3 (2.1)	−1.5 (2.2)	−0.4 (2)	0.917	−0.7 (1.2)	0.9 (1.9)	−1.3 (2)	−1.9 (1.8)	0.525	−0.8 (1.1)
Measure 56	2.4 (2.3)	1.6 (2.5)	5.3 (2.3)	0.490	3.1 (1.4)	1.9 (2.2)	1.7 (2.4)	5.2 (2.2)	0.437	2.9 (1.3)	2.2 (2.1)	1.9 (2.2)	4.3 (2.1)	0.680	2.8 (1.2)
Measure 57	3.7 (2.1)	3.6 (2.3)	5.2 (2.1)	0.831	4.2 (1.2)	3.7 (2.1)	3.6 (2.3)	5.2 (2.1)	0.833	4.2 (1.2)	3.6 (2.1)	3.7 (2.3)	5.4 (2.1)	0.783	4.2 (1.2)
Measure 58	−2.1 (2.6)	−3.4 (2.7)	−9.4 (2.5)	0.097	−4.9 (1.5)	−1.2 (2.4)	−3.4 (2.5)	−9.2 (2.3)	0.045	−4.6 (1.4)	−1.9 (2.3)	−3.5 (2.4)	−10 (2.3)	0.024	−5.2 (1.3)
Measure 59	−2.7 (2.3)	−1.5 (2.5)	−7.8 (2.3)	0.127	−4 (1.4)	−2.2 (2.2)	−1.5 (2.4)	−7.7 (2.2)	0.101	−3.8 (1.3)	−2.2 (2.2)	−2.3 (2.3)	−8.3 (2.1)	0.075	−4.2 (1.3)
Measure 60	−1.8 (2.2)	−2.4 (2.3)	−5.8 (2.2)	0.385	−3.3 (1.3)	−1.1 (2.1)	−2.5 (2.2)	−5.7 (2)	0.278	−3.1 (1.2)	−1.5 (2.1)	−2.8 (2.2)	−5.8 (2)	0.323	−3.4 (1.2)
Measure 61	−2.8 (2.5)	−4.4 (2.7)	−7.4 (2.4)	0.407	−4.9 (1.5)	−2.1 (2.4)	−4.5 (2.5)	−7.3 (2.3)	0.299	−4.6 (1.4)	−3.2 (2.3)	−4.8 (2.5)	−7.2 (2.3)	0.470	−5.1 (1.4)
Measure 62	−0.7 (2.5)	−4.0 (2.6)	−9.4 (2.4)	0.038	−4.7 (1.4)	−0.6 (2.5)	−4.0 (2.6)	−9.3 (2.4)	0.036	−4.6 (1.4)	−1.3 (2.4)	−4.6 (2.5)	−8.6 (2.3)	0.095	−4.8 (1.4)
Measure 63	−1.5 (3.3)	−2.7 (3.4)	−2.2 (3.2)	0.965	−2.1 (1.9)	−0.8 (3.2)	−2.7 (3.3)	−2.1 (3.1)	0.911	−1.9 (1.8)	−1.4 (3)	−2.8 (3.2)	−4.2 (2.9)	0.811	−2.8 (1.8)
Measure 64	−0.5 (2.5)	−4.5 (2.6)	−9.6 (2.4)	0.028	−4.9 (1.4)	−0.1 (2.4)	−4.5 (2.5)	−9.6 (2.3)	0.018	−4.7 (1.4)	−1.1 (2.3)	−4.7 (2.5)	−9.5 (2.3)	0.035	−5.1 (1.4)
Measure 65	−1.5 (1.8)	−6.4 (1.9)	−5 (1.7)	0.143	−4.3 (1)	−1.6 (1.8)	−6.4 (1.9)	−5 (1.7)	0.156	−4.4 (1)	−1.6 (1.7)	−6.6 (1.8)	−4.7 (1.7)	0.132	−4.3 (1)

Table 10 Roll-off

	Model 1					Model 2					Model 3				
	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff
Measure 54	0.5 (0.6)	−0.1 (0.6)	0.5 (0.5)	0.623	0.3 (0.3)	0.5 (0.6)	−0.1 (0.6)	0.5 (0.5)	0.670	0.3 (0.3)	0.3 (0.5)	0 (0.6)	0.3 (0.5)	0.872	0.2 (0.3)
Measure 55	1.4 (0.8)	0.2 (0.8)	−0.7 (0.8)	0.268	0.3 (0.5)	1.3 (0.8)	0.2 (0.8)	−0.7 (0.8)	0.305	0.2 (0.5)	0.9 (0.7)	0 (0.7)	−0.5 (0.7)	0.513	0.1 (0.4)
Measure 56	0.7 (0.7)	0.1 (0.7)	−1.4 (0.7)	0.151	−0.2 (0.4)	0.6 (0.7)	0.1 (0.7)	−1.4 (0.7)	0.154	−0.2 (0.4)	0.4 (0.6)	0.1 (0.7)	−1.5 (0.6)	0.088	−0.3 (0.4)
Measure 57	0.4 (0.5)	−0.2 (0.5)	−0.5 (0.5)	0.554	−0.1 (0.3)	0.4 (0.5)	−0.2 (0.5)	−0.5 (0.5)	0.579	−0.1 (0.3)	0.3 (0.4)	−0.3 (0.5)	−0.6 (0.4)	0.496	−0.2 (0.3)
Measure 58	0.4 (0.4)	−0.1 (0.4)	−0.2 (0.4)	0.737	0 (0.2)	0.4 (0.4)	−0.1 (0.4)	−0.2 (0.4)	0.738	0 (0.2)	0.3 (0.4)	−0.1 (0.4)	−0.3 (0.4)	0.736	0 (0.2)
Measure 59	1 (0.6)	−0.3 (0.6)	−0.5 (0.6)	0.286	0.1 (0.3)	0.9 (0.6)	−0.3 (0.6)	−0.5 (0.5)	0.328	0 (0.3)	0.7 (0.5)	−0.3 (0.6)	−0.6 (0.5)	0.303	−0.1 (0.3)
Measure 60	1.3 (0.7)	−0.1 (0.7)	0 (0.6)	0.272	0.4 (0.4)	1.2 (0.7)	−0.1 (0.7)	0 (0.6)	0.308	0.4 (0.4)	1.2 (0.6)	0 (0.6)	0.1 (0.6)	0.279	0.4 (0.4)
Measure 61	1.7 (0.7)	−0.3 (0.7)	−0.5 (0.7)	0.094	0.3 (0.4)	1.5 (0.7)	−0.3 (0.7)	−0.5 (0.7)	0.117	0.2 (0.4)	1.4 (0.6)	−0.2 (0.7)	−0.4 (0.6)	0.166	0.3 (0.4)
Measure 62	0.9 (0.5)	−0.5 (0.6)	−0.7 (0.5)	0.168	−0.1 (0.3)	0.8 (0.5)	−0.5 (0.6)	−0.7 (0.5)	0.184	−0.1 (0.3)	0.6 (0.5)	−0.5 (0.5)	−0.8 (0.5)	0.131	−0.2 (0.3)
Measure 63	1 (0.6)	0.1 (0.6)	−0.2 (0.5)	0.314	0.3 (0.3)	0.9 (0.5)	0.1 (0.6)	−0.3 (0.5)	0.382	0.2 (0.3)	0.7 (0.5)	0.1 (0.5)	−0.4 (0.5)	0.489	0.1 (0.3)
Measure 64	0.6 (0.5)	−0.4 (0.5)	0.2 (0.5)	0.422	0.1 (0.3)	0.6 (0.5)	−0.4 (0.5)	0.2 (0.5)	0.442	0.1 (0.3)	0.6 (0.5)	−0.4 (0.5)	0.1 (0.4)	0.475	0.1 (0.3)
Measure 65	0.5 (0.6)	−1.1 (0.7)	−1.4 (0.6)	0.041	−0.7 (0.4)	0.5 (0.6)	−1.1 (0.7)	−1.4 (0.6)	0.041	−0.7 (0.4)	0.3 (0.6)	−1.3 (0.6)	−1.1 (0.6)	0.040	−0.7 (0.3)

Table 11 Candidates

	Model 1					Model 2					Model 3				
	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff
pres	−1.8 (3.1)	3.2 (3.3)	−0.5 (3)	0.519	0.3 (1.8)	−3 (2.7)	3.3 (2.9)	−0.7 (2.6)	0.274	−0.3 (1.7)	−3.2 (2.6)	2.8 (2.7)	0.8 (2.5)	0.261	0.1 (1.5)
sen	2.1 (2.9)	3.5 (3.1)	3.4 (2.9)	0.927	2.9 (1.8)	0.8 (2.5)	3.6 (2.6)	3.2 (2.4)	0.694	2.2 (1.7)	0.5 (2.4)	4 (2.5)	3.3 (2.3)	0.562	2.1 (1.7)
cong	1.2 (6.1)	10.5 (6.4)	0.1 (5.7)	0.427	2.6 (4.5)	−1.5 (6)	10.4 (6.3)	−0.4 (5.6)	0.319	1.5 (4.4)	−0.8 (5.7)	9.9 (5.9)	1.9 (5.3)	0.406	2.4 (4.2)
sos	0 (2.8)	0.1 (2.9)	1.5 (2.7)	0.913	1.5 (1.3)	−1 (2.5)	0.1 (2.7)	1.3 (2.4)	0.803	1 (1.2)	−1.3 (2.4)	0.1 (2.5)	3.1 (2.3)	0.403	1.5 (1.1)
st	−0.7 (2.7)	1.8 (2.9)	1.7 (2.6)	0.765	1.3 (1.4)	−1.7 (2.4)	1.9 (2.6)	1.5 (2.4)	0.519	0.8 (1.3)	−2.3 (2.3)	1.6 (2.4)	3 (2.3)	0.244	1 (1.2)
ss	7.1 (7.6)	−1.6 (5.9)	2.9 (7)	0.659	5.3 (3.5)	1 (6.6)	−0.8 (5.1)	1.3 (6)	0.960	2.5 (3)	2 (6.4)	−0.8 (4.9)	−2 (5.9)	0.900	1.6 (2.9)
sh	−0.5 (4.5)	−1.3 (4.2)	2 (4.4)	0.860	1.9 (2.2)	−2.3 (4.4)	−1 (4.1)	2.1 (4.3)	0.762	1.3 (2.1)	−3.5 (4.2)	−1.9 (3.9)	1.7 (4.1)	0.670	0.7 (2.1)

Table 12 Turnout

	Model 1					Model 2					Model 3				
	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff	Dem	Ind	Rep	*p* value	main eff
Turnout	0.6 (1.1)	−0.8 (1.2)	−0.5 (1.1)	0.519	0.4 (0.6)	0.7 (1.1)	−0.8 (1.2)	−0.5 (1.1)	0.274	0.5 (0.6)	1.7 (0.9)	−0.5 (1)	−0.1 (0.9)	0.261	0.9 (0.4)

References

Ansolabehere, S., & Hersh, E. (2010). *The quality of voter registration records: A state-by-state analysis*. Cambridge, MA: Department of Government, Harvard University.

Arceneaux, K. (2005). Using cluster randomized field experiments to study voting behavior. In D. Green and A. Gerber (Eds.), *The science of voter mobilization: The Annals of the American Academy of Political and Social Science* (vol 601, pp. 169–179).

Arceneaux, K. (2007). I'm asking for your support: The effects of personally delivered campaign messages on voting decisions and opinion formation. *Quarterly Journal of Political Science, 2*(1), 43–65.

Bowler, S., & Donovan, T. (1998). *Demanding choices: Opinion and voting in direct democracy*. Ann Arbor: University of Michigan Press.

Broder, D. S. (2000). *Democracy derailed: Initiative campaigns and the power of money*. New York: Harcourt.

Childers, M., & Binder, M. (2012). Engaged by the initiative? How the use of citizen initiatives increases voter turnout. *Political Research Quarterly, 65*(1), 93–103.

Converse, P. E. (2000). Assessing the capacity of mass electorates. *Annual Review of Political Science, 3*(1), 331–353.

De Figueiredo, J. M., Ji, C. H., & Kousser, T. (2011). Financing direct democracy: Revisiting the research on campaign spending and citizen initiatives. *Journal of Law, Economics, and Organization,*. doi:10.1093/jleo/ewr007.

Ellis, R. (2002). *Democratic delusions: The initiative process in America*. Lawrence: University of Kansas.

Freedman, D. A. (2006). On the so-called 'Huber sandwich estimator' and 'robust standard errors'. *The American Statistician, 60*(4), 229–302.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96*(4), 1043–1068.

Garrett, E., & Gerber, E. R. (2001). Money in the initiative and referendum process: Evidence of its effects and prospects for reform. In D. Waters (Ed.), *The battle over citizen lawmaking* (pp. 73–95). Durham, N.C: Carolina Academic Press.

Gelman, A., & King, G. (1993). Why are American presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science, 23*(4), 409–451.

Gerber, E. (1999). *The populist paradox: Interest group influence and the promise of direct legislation*. Princeton, NJ: Princeton University Press.

Gerber, A. (2004). Does campaign spending work? Field experiments provide evidence and suggest new theory. *American Behavioral Scientist, 47*(5), 541–574.

Gerber, A., & Green, D. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review, 94*(3), 653–663.

Gerber, A. S., Kessler, D. P., & Meredith, M. (2011). The persuasive effects of direct mail: A regression discontinuity based approach. *The Journal of Politics, 73*(1), 140–155.

Green, D. P., & Gerber, A. S. (2008). *Get out the vote: How to increase voter turnout*. Washington, DC: Brookings Institution Press.

Holbrook, T. M. (1996). *Do campaigns matter?*. Thousand Oaks, CA: Sage.

Initiative and Referendum Institute. (2000). http://www.iandrinstitute.org.

Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American opinion*. Chicago: University of Chicago.

Keane, L., Nickerson, D. (2013). A field experiment on nonpartisan mobilization and persuasion down-ballot. Unpublished manuscript.

Krosnick, J. A., & Kinder, D. R. (1990). Altering the foundations of support for the president through priming. *The American Political Science Review*, 497–512.

Lowenstein, D. H. (1982). Campaign spending and ballot propositions: Recent experience, public choice theory and the first amendment. *UCLA Law Review, 29*(3), 505–641.

Magleby, D. (1984). *Direct legislation: Voting on ballot propositions in the United States*. Baltimore, MD: Johns Hopkins University Press.

Meredith, M., Malhotra, N. (2008). *Can October surprise: A natural experiment assessing late campaign effects*." Stanford Graduate School of Business Research Paper Number 2002.

Nicholson, S. P. (2003). The political environment and ballot proposition awareness. *American Journal of Political Science, 47*, 403–410.

Owens, J. R., & Wade, L. L. (1986). Campaign spending on California ballot propositions, trends and effects, 1924–1984. *Western Political Quarterly, 39*(4), 675–689.

Rogers, T., Gerber, A. S., & Fox, C. R. (2012). Rethinking why people vote: Voting as dynamic social expression. In E. Shafir (Ed.), *Behavioral foundations of policy*. New York: Russell Sage.

Rogers, T., & Nickerson, T. W. (2013). Can inaccurate beliefs about incumbents be changed? And can reframing change votes?

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1*(1), 7–59.

Schrag, P. (1998). *Paradise lost: California's experience, America's future*. New York: New Press.

Smith, D. A. (1998). *Tax crusaders and the politics of direct democracy*. New York: Routledge.

Smith, D. A. (2001a). Homeward bound?: Micro-level legislative responsiveness to ballot initiatives. *State Politics and Policy Quarterly, 1*(Spring), 50–61.

Smith, M. (2001b). The contingent effects of ballot initiatives and candidate races on turnout. *American Journal of Political Science, 45*(3), 700–706.

Smith, D. A., & Tolbert, C. J. (2004). Educated by initiative: The effects of direct democracy on citizens and political organizations in the American states. University of Michigan Press.

Smith, D. A., & Tolbert, C. J. (2010). Direct democracy, public opinion, and candidate choice. *Public Opinion Quarterly, 74*(1), 85–108.

Stratmann, T. (2005). Some talk: Money in politics. A (partial) review of the literature. *Public Choice, 124*(1–2), 135–156.

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.

Tolbert, C. J., & Smith, D. A. (2005). The educative effects of ballot initiatives on voter turnout. *American Politics Research, 33*(2), 283–309.

Tolbert, C. J., Grummel, J. A., & Smith, D. A. (2001). The effect of ballot initiatives on voter turnout in the American states. *American Politics Research, 29*, 625–648.

Weber, C., Dunaway, J., & Johnson, T. (2011). It's all in the name: Source cue ambiguity and the persuasive appeal of campaign ads. *Political Behavior*. doi:10.1007/s11109-011-9172-y.