Ballot Order Effects in Direct Democracy Elections

John G. Matsusaka[*]

University of Southern California

Many political practitioners believe that voters are more likely to approve propositions listed at the top than the bottom of the ballot, and this belief influences election laws across the country. A large body of research shows that ballot structure matters for candidate elections, potentially distorting democratic decisions, but there is little comparable evidence for ballot propositions. This paper offers two strategies for identifying the causal effect of ballot order in proposition elections, and applies the methods to data from California during 1958-2014 and Texas during 1986-2015. There is little evidence that being listed at the top compared to the bottom of the ballot has an effect on approval. Approval rates are lower with more propositions on the ballot.

Keywords: Direct democracy; initiative; referendum; ballot proposition; ballot order; causality

December 2015

[*] Marshall School of Business, Department of Political Science, and Gould School of Law at the University of Southern California. Comments welcome: please contact the author at matsusak@usc.edu. For helpful comments and suggestions, I thank Odilon Camara, Dan Klerman, and participants at the Initiatives and Referendums Conference at USC in November 2012 and the CLASS workshop at USC. I thank USC for financial support.

Ballot Order Effects in Direct Democracy Elections

1. Introduction

In the summer of 2012, allies of California Governor Jerry Brown persuaded the legislature to amend the state's elections code so that the governor's tax-raising initiative would be listed first among 11 propositions on the ballot. Although the change was officially motivated by a desire to ensure that voters were able to "carefully weigh the consequences of important measures" on the ballot, it was widely believed that the real purpose was to increase the initiative's chance of passing.[1] Opponents of the initiative argued that the governor's allies had cynically manipulated the elections code to secure the most favorable position for the governor's proposal. The implicit assumption in the debate was that ballot position matters for direct democracy elections, and specifically, that the first position confers an advantage.

The purpose of this paper is to assess the premise that ballot position influences the outcome of direct democracy elections. The idea that the top position is best is not new: writing almost a half century ago, Mueller (1969, p. 1208) observed:

> "The state legislature devoutly believes in the existence of a body of citizens who start out voting affirmatively on bond issues but turn to negativism as they move down the ballot viewing with mounting horror the extent of the proposed

[1] The findings and declarations in the new law (AB 1499) stated: "bond measures and constitutional amendments should have priority on the ballot because of the profound and lasting impact these measures can have on our state. . . . In recognition of their significance, bond measures and constitutional amendments should be placed at the top of the ballot to ensure that the voters can carefully weigh the consequences of these important measures."

expenditures. Part of the reason for placing state bond issues at the top of the ballot is to catch the affirmative votes of these citizens before they turn sour."

Theoretically, the top position may be advantageous if voters become fatigued moving down the ballot, and if decision fatigue causes a status quo bias that leads to rejection of new proposals.[2] Empirically, there is a healthy literature on order effects in candidate elections, but little evidence on order effects in proposition elections. Given the apparent existence of order effects in candidate elections, the widespread belief of order effects among political practitioners, and the role of this belief in framing election law, it seems worthwhile to try to estimate the extent to which ballot structure actually matters for direct democracy elections.

Our current knowledge of order effects in direct democracy elections is limited by a dearth of evidence that is convincingly causal in nature. The main contribution of this paper is to offer evidence that addresses common challenges to causal inference. First, since 1986 Texas has assigned ballot positions for propositions by lottery, producing randomized experimental data. The mean observed approval rates can be compared across ballot positions to provide direct estimates of ballot order effects. Second, in California, the Field Poll routinely surveys likely voters about their voting intentions on select ballot propositions in a way that is not closely linked to the order in which the propositions will appear in the ballot. These survey responses capture voter preferences about a proposition independent of the proposition's position on the ballot. Ballot

[2] For discussion and variants on this idea, see Miller and Krosnick (1998), Bowler and Donovan (1998), Levav et al. (2010), and Augenblick and Nicholson (forthcoming).

position effects can then be inferred by comparing each proposition's approval rate when "treated" with its actual ballot position to its expressed pre-election Field Poll approval rate (the "control").

The main finding is a consistent absence of evidence that the top (or any) position on the ballot is particularly favorable. Election data for the 240 Texas propositions during 1986-2015 show a correlation of -0.01 between ballot position and approval rates, and parametric estimates controlling for other factors also fail to reveal a meaningful connection. Similarly, an examination of the 242 California propositions during 1958-2014 for which Field Poll data are available fails to reveal a robust effect of ballot position on approval rates after controlling for pre-election opinion.

The evidence from Texas and California points in the same direction and is thus complementary given the dissimilarity in the electoral contexts of the two states. Texas proposition elections typically take place in odd-numbered years in which there are no major candidate races on the ballot and feature somewhat technical amendments to the constitution proposed by the legislature, while California elections often feature controversial voter initiatives that attract significant public attention and appear on the same ballot as high-profile candidate elections. One could argue that ballot order effects are more likely to occur in low turnout, low information elections (Texas) or in high turnout, high information elections (California); the absence of an effect in both cases using different methods suggests the finding may be fairly general.

I also explore the related issue of ballot length. Practitioners and scholars argue that the information requirements associated with long ballots can overwhelm voters, causing the status quo bias to kick in and leading to more "no" votes. It is not difficult to find examples of elections that voters must have found challenging, such as the 1914 California general election in which voters had to decide 48 propositions that were on the ballot. The danger of overloading voters has

led some states to establish limits on the number of propositions that can appear on a ballot, for example, Arkansas and Illinois limit the number of legislative constitutional amendments to three, and Mississippi has a cap of five initiatives per ballot. Previous research suggests that voters are more likely to reject a proposition when it appears on a ballot with many other propositions (Bowler et al., 1992), and more generally, that the size of the choice set affects decision making (Selb, 2008; Iyengar and Kamenica, 2010). This study's estimates on ballot length are less compelling in terms of causal inference than the order estimates, but reveal a consistently lower approval rate for propositions on long compared to short ballots.

The initial motivation for this paper was to evaluate the premise underlying a live policy issue. However, the evidence also speaks to broader issues. In terms of voter behavior, some evidence suggests that decision making is cognitively costly (e.g. Baumeister et al. (1998), Danziger et al. (2011)); if voters deplete their mental resources when faced with numerous decisions, their choices may not be rational. The evidence reported here suggests that decision fatigue does not cause order effects in ballot proposition elections, but it may play a role when ballots become too long. At an even broader level, direct democracy continues to play a leading role in policy making in the United States. Ballot propositions have been a central arena for emerging social issues such as same-sex marriage and marijuana legalization, and continue to be a vehicle through which substantial financial decisions are made, for example, voters have decided whether to authorize over $195 billion worth of bond propositions since 2000.[3] Direct democracy is motivated by the belief that laws passed by the voters are more likely to reflect their preferences than laws passed by legislatures, so it would be of concern if non-preference factors such as ballot design turned out to

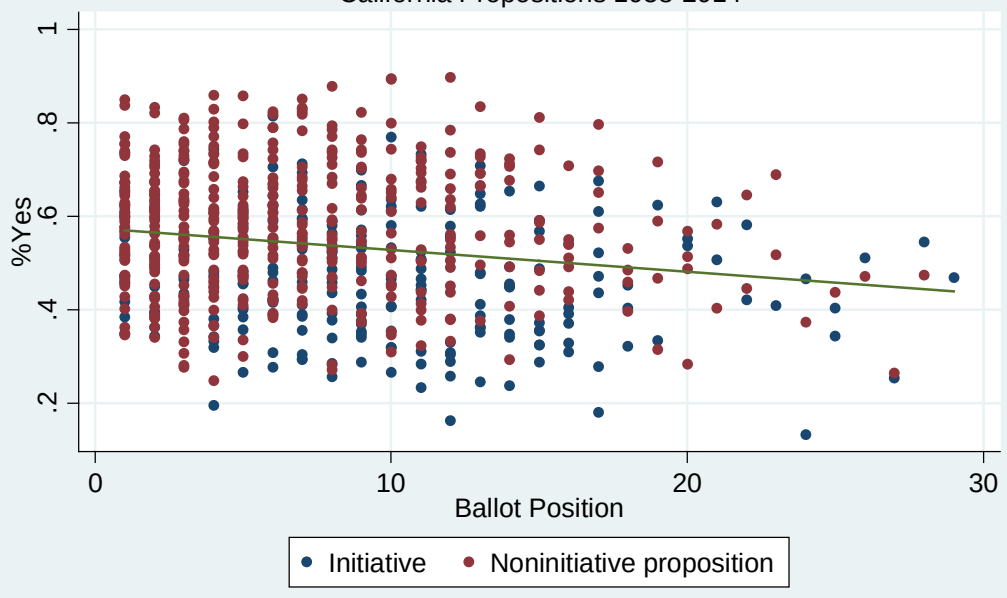
[3] Author's calculation.

have a big effect on outcomes. The evidence in this paper suggests that on average, ballot position is unlikely to have a large effect on election outcomes, allaying that concern to some extent.

2. Institutional Context

As a simple correlation, propositions listed at the top of the ballot do better than propositions listed at the bottom of the ballot. Figure 1 plots the approval rate $\left(= \frac{\text{yes votes}}{\text{yes votes + no votes}}\right)$ on each California ballot proposition during the period 1958-2014 against the proposition's ballot position, where #1 indicates that the measure was listed first. The solid line, from a linear regression, shows that there is indeed a negative relation between votes in favor and ballot position; approval falls approximately 0.5 percent with each additional position.

While Figure 1 shows that historical approval rates decline moving down the ballot, it does not follow that ballot position *causes* the declining approval rates. It could be that more popular



Note. The figure plots all 678 propositions that appeared on the California ballot during 1958-2014. Data source: Initiative and Referendum Institute.

measures are more likely to be placed at the top of the ballot. The recent California episode illustrates how this could happen. Before it was modified in June 2012, the election law read:[4]

> The order in which all state measures that are to be submitted to the voters shall appear upon the ballot is as follows:
>
> (a) Bond measures in the order in which they qualify.
>
> (b) Constitutional amendments in the order in which they qualify.
>
> (c) Other Legislative measures in the order in which they are approved by the Legislature.
>
> (d) Initiative measures in the order in which they qualify.
>
> (e) Referendum measures, in the order in which they qualify.

To define terms: in California, (a) a bond measure is a proposal to authorize the issuance of bonds; (b) a constitutional amendment is a proposal to amend the state constitution; (c) "other legislative measures" are proposals to modify previously approved initiative statutes; (d) an initiative is a new law – bond measure, constitutional amendment, or statute – that is proposed by citizens and qualifies for the ballot by petition; and (e) a referendum is a proposal, qualified by petition, to repeal a law recently passed by the legislature.[5] As can be seen, the original code placed legislative proposals (bond issues, constitutional amendments, statutes) first followed by citizen

[4] California Elections Code 13115, enacted by Stats. 1994, Ch. 920, Sec. 2, SB 1547.

[5] Ballot proposition terminology varies by state and country. In the California election code, a "referendum" is a proposal to veto a law passed by the legislature; in other jurisdictions it refers more generally to any popular vote on a law, whether proposed by citizens, the legislature, or other means. See Lupia and Matsusaka (2004) for more details.

proposals (initiatives and referendums). Within each category, propositions were ordered by the date at which they qualified for the ballot.[6] After June 2012, the election code became:[7]

> The order in which all state measures that are to be submitted to the voters shall appear upon the ballot is as follows:
>
> (a) Bond measures, including those proposed by initiative, in the order in which they qualify.
>
> (b) Constitutional amendments, including those proposed by initiative, in the order in which they qualify.
>
> (c) ~~Other~~ Legislative measures, other than those described in subdivision (a) or (b), in the order in which they are approved by the Legislature.
>
> (d) Initiative measures, other than those described in subdivision (a) or (b), in the order in which they qualify.
>
> (e) Referendum measures, in the order in which they qualify.

The new code blurs the distinction between legislative and citizen-initiated proposals. Now bond measures are listed first, regardless of whether they originate from the legislature or citizen petition, followed by constitutional amendments, regardless of whether they originate from the

[6] The pre-2012 code is actually somewhat ambiguous. One could read the law to mean that a bond measure proposed by initiative is to be included in subdivision (a) and a constitutional amendment proposed by initiative is to be included in subdivision (b). Under such an interpretation, subdivision (d) would apply only to non-bond, non-amendment initiatives. The text describes how the law was implemented in practice.

[7] An underline is new text; a strikethrough is deleted text. The code was modified by Stats. 2012, Ch. 30, Sec. 2.

legislature or petition. For non-bond statutory proposals, the ordering stays the same: legislative proposals followed by citizen initiatives. Referendums remain at the bottom of the ballot.[8]

The California elections code introduces several potential selection effects. First, prior to 2012, it placed proposals from the legislature ahead of citizen proposals. Historically, legislative measures have a much higher rate of passage than citizen measures; during the period 1958-2014, 72 percent of legislative proposals were approved compared to 37 percent of citizen-initiated proposals. This is probably due to the fact that legislative proposals must garner majority support in both chambers – supermajority support in the case of constitutional amendments – so they are likely to have broader appeal than initiatives and referendums, which require only signatures of a small percentage of the electorate.[9] Second, bond proposals must pass a different screening process than constitutional amendments (see footnote 9), which could cause voters to view them differently, and voters may be more hesitant to amend the constitution than to approve a bond measure. Historically, during the period 1958-2014, voters approved 78 percent of legislative bond measures, 69 percent of legislative constitutional amendments, and 76 percent of legislative statutes. Third, measures that qualify at an earlier date appear toward the top of the ballot.

---

[8] Governor Brown's Proposition 30 was an initiative that proposed to amend the constitution. As an initiative, it was originally included in subdivision (d), and because it qualified later in the cycle than other initiatives, it was slated to appear near the bottom of the ballot. By giving precedence to constitutional amendments, whatever the source, the revised code moved the governor's proposal to the top of the ballot; there were no bond propositions in that election and Proposition 30 was the only constitutional amendment.

[9] To reach the ballot, a bond proposal requires a majority vote in both the Assembly and Senate and signature of the governor; a constitutional amendment requires a two-thirds vote in both chambers but does not require the governor's signature; and a statute that amends an initiative requires a majority vote in both chambers and signature of the governor. The initiative signature requirement, expressed as a percent of the votes cast for governor in the previous election, is 8 percent for constitutional amendments and 5 percent for statutes (since 1966); the requirement for referendums is 5 percent.

Proposals that are inherently more popular may qualify earlier because it is easier to achieve a legislative consensus on them and easier to collect the requisite signatures.

California's practice of arranging the ballot by grouping issues and placing them in a predetermined order is common. For example, most states give priority to issues according to the time they are qualified for the ballot. Arkansas, Arizona, Colorado, and North Dakota place constitutional amendments before statutes. Maine places bond measures at the bottom of the ballot. New Mexico and Rhode Island place constitutional amendments at the top and bond measures at the bottom. Washington places advisory measures at the bottom of the ballot. Because in most states we expect to see different approval rates for propositions at the top compared to the bottom of the ballot for reasons having nothing to do with ballot order, we cannot infer that that a correlation between approval and ballot position is causal. [10]

3. Theory and Existing Literature

The literature on ballot position effects in candidate elections is extensive. Miller and Krosnick (1998) in a well-known survey observe that while much research concludes that candidates benefit from being listed first, often the estimated effects are small and research designs do not separate causation from correlation. The more recent literature that employs stronger research designs generally finds that the first position is advantageous (see Meredith and Yuval (2013)), but some

[10] We can get a rough sense of the importance of the ballot ordering rules in California by estimating the relation between approval rates and ballot position with and without controls for type of proposition. The coefficient of -0.46 in Figure 1 becomes -0.21 when dummy variables for initiatives, referendums, and legislative measures are added to the regression, suggesting that about half of the negative relation is due to selection by type of measure. A negative relation (of smaller magnitude) appears if the sample is restricted to only initiatives, only referendums, or only legislative measures.

studies find small or nonexistent order effects (Alvarez et al., 2006; Ho and Imai, 2008). Such ballot order effects as do exist are typically attributed to voters losing interest or ceasing to seek favorable information about candidates as they move down the ballot (satisficing). The purpose of the present study is not to develop a new theory of ballot order effects, but rather to provide an empirical assessment of existing theories. To do that it is useful to review the main ideas, with an eye toward their testable implications.

Theoretically, the logic for order effects in candidate elections does not easily carry over to proposition elections. In candidate elections, voters might be more inclined to select a name at the top of the list, perhaps because they lose interest or stop moving down a list once they find an acceptable option, but this line of reasoning applied to propositions would imply roll-off (abstention) moving down the ballot, not a proclivity to vote "no" on propositions at the bottom of the ballot. In a candidate election, voters face a problem like the following:

Choose one:

□ T. Butler

□ A. Iommi

□ J. Osborne

□ W. Ward

Voters can select one and only one name from the list. If voters satisfice – stopping once they find a "good enough" option – or become tired moving down a list, then appearing at the top of the ballot in a candidate election would confer an advantage.

The problem facing voters in a referendum election is different:

Proposition 1 Choose one: □ Yes □ No

Proposition 2 Choose one: □ Yes □ No

Proposition 3 Choose one: □ Yes □ No

Proposition 4 Choose one: □ Yes □ No

If voters become fatigued when moving down the list of propositions, we might expect to see more abstention moving down the ballot, but it is less obvious why voters would be more inclined to check the "No" box as they move down the ballot. Evidence of order effects from candidate elections, then, does not generalize naturally to direct democracy elections; to understand whether position matters in proposition elections, we need to evaluate evidence from proposition elections.

The existing literature on order effects using data from direct democracy elections is modest. Early statistical evidence was compiled and published by the California Secretary of State (1981). That study, entirely descriptive, reports the mean percentage of votes in favor by ballot position for all California propositions during the period 1884-1980.[11] The data show an irregular pattern, with approval rates not obviously dropping when moving down the ballot.

Bowler et al. (1992) examine a subset of these data, 190 California propositions during 1974-1988, in a more systematic way. The study reports regressions of the percentage of votes against a proposition on its ballot position and several control variables, including type of measure (initiative, bond measure, constitutional amendment), type of election (presidential, general,

[11] The data are reported in an unnumbered table with the heading, "Success Rate of Each Ballot Position."

primary), number of words in a proposition, and campaign spending (see their Table 1; reproduced as Table 5 in Bowler and Donovan (1998)). The regression includes first and second order terms for ballot position, and the coefficient estimates imply a U-shaped relation that bottoms out at position #8.[12] That is, votes against a proposition decline over the first eight ballot positions, and then increase over the subsequent ballot positions. There is no theoretical reason to expect ballot order effects to reverse at position #8; these correlations may not be causal.

Matsusaka (2013) examines 637 California propositions during 1960-2010. The study documents an overall negative relation between approval rates and ballot position, but shows that this relation is mainly due to the fact that voter initiatives, the least popular type of proposition, typically appear at the bottom of the ballot. When initiatives, bond measures, and legislative constitutional amendments are considered as separate groups, the correlation between approval and ballot position vanishes, except in the group of bond proposals. The study also reports non-California evidence on ballot position from all 1,058 state-level propositions that appeared in the other states during the period 2003-2012. A negative relation between approval and ballot position appears in this sample as well, but again, appears to be due to legal rules that place inherently unpopular propositions at the bottom of the ballot. The study does not offer evidence that can support strong causal inference.

The one existing study that employs plausibly random assignment to identify effects is Augenblick and Nicholson (forthcoming). That study, which uses precinct-level voting data from San Diego County during 1992-2002, exploits the fact that a typical ballot includes a set of federal,

[12] The coefficient on ballot-position is -2.12 and the coefficient on ballot-position-squared is 0.13, so the turning point is $\frac{2.12}{2\times 0.13} = 8.2$.

state, and local candidate races that are listed before the state propositions, and the set of candidate races varies by precinct. Because of variation in the number of state and local candidate races, voters in different precincts may find the ballot propositions preceded by a different number of races. For example, if voters in one precinct face a state senate race while voters in another precinct do not face such a race, the propositions will appear one position farther down the ballot in the first precinct than the second precinct. Using this variation, the study finds that proposition approval rates are lower when they are listed farther down the ballot; specifically each position farther down the ballot results in 0.12 percent fewer votes in favor. The Augenblick and Nicholson study offers the clearest evidence to date on order effects, however, the variation exploited by the study – moving the *entire* block of propositions lower on the ballot – is different from the exercise of moving one proposition to another position *within* the block, which is the situation of concern in recent debates.[13]

Finally, Binder and Kousser (2014) study experimental evidence from a survey. They ask a sample of Florida voters in 2012 their opinion on three Florida propositions appearing on the 2012 ballot, as well as two hypothetical propositions related to contemporary California propositions, varying the order in which questions are asked. The findings are mixed; some propositions do better when asked about first while others do better when asked about last.

[13] It is also possible that including another candidate race affects voting on ballot propositions independent of the fact that the candidate race is placed before the propositions. For example, having a presidential election on the ballot pushes the propositions one spot lower, but the effects of a presidential election might be significant regardless of whether that race is listed before or after the propositions.

4. Evidence from Texas

A. *Methods and Data*

The Texas data are analyzed assuming a model of the following form:

$$(1) \qquad V_{it}^{ELECT} = V_{it}^{*} + \alpha + \beta \cdot POS_{it} + \gamma \cdot LENGTH_{it} + e_{it},$$

where V_{it}^{ELECT} is the percentage of votes cast in favor of proposition i in election t, V_{it}^{*} is the (unobserved) "true" preferences of voters in the hypothetical situation where voting is uninfluenced by ballot position, POS_{it} is the measure's ballot position (#1 is the top of the ballot, #2 is the second proposition, etc.), $LENGTH_{it}$ is the number of propositions on the ballot, e_{it} is an error term, and α, β, γ are parameters to be estimated. If voters are less inclined to approve as they move down the ballot, then $\beta < 0$.

The estimation challenge is that V_i^* is not observable; if we omit V_i^* and simply regress votes on ballot position, we have a textbook omitted variables problem, and the estimates of β will be biased if V_i^* is correlated with ballot position. One way to avoid this problem is to randomize ballot position, as Texas has done since 1986.[14] With positions assigned randomly, there is no reason to expect the underlying popularity of a measure to be related to its ballot position; therefore, omitting V_i^* from estimates of (1) should not introduce a bias in the estimate of β. The strategy then is simply to investigate whether propositions at the top of the ballot attract more favorable votes than those at the bottom of the ballot.

[14] Texas Election Code, Title 16, Chapter 274, Subchapter A, Section 274.002. The relevant text is: "If more than one proposed constitutional amendment is to be submitted in an election, the order of the propositions submitting the amendments shall be determined by a drawing . . . "

The data are drawn from official election results published by the Texas Secretary of State. Summary information on the 240 Texas propositions that appeared during 1986-2015 are reported in Panel A of Table 1. Texas does not allow initiatives or referendums, and the legislature does not place statutes on the ballot; therefore, all propositions are constitutional amendments proposed by the legislature. The main variable of interest V_{it}^{ELECT} is operationalized as the approval rate, or "%Yes", defined to be $\%Yes = 100 \times \frac{\text{yes votes}}{\text{yes votes + no votes}}$. Abstentions are ignored.[15]

B. Findings

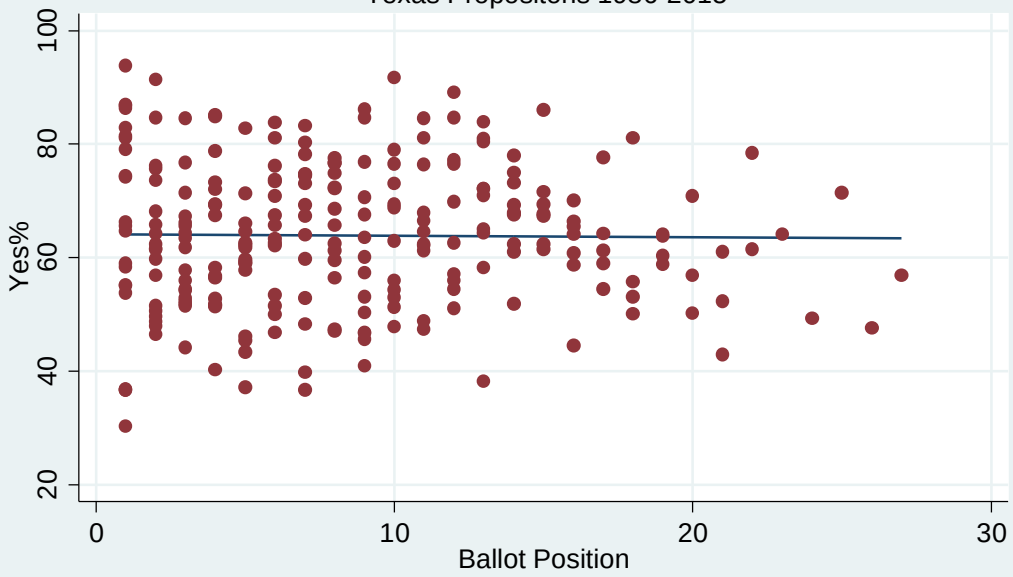
Figure 2 plots the approval rate for Texas propositions against their ballot position. The solid line is a regression of approval on position. The regression line is almost completely flat, indicating that there is essentially no connection on average between ballot position and approval rates. This is initial evidence against the hypothesis that position has an important effect on approval.

Table 2 extends the investigation by reporting regressions of the approval rate on ballot position. Each column in the table reports results from a regression. Column (1) reports a regression representing the solid line in Figure 2. Taken at face value the coefficient of -0.03 on ballot position indicates that each position further down the ballot is associated with 0.03 percent fewer votes in favor, a small number that cannot be distinguished from zero statistically.

[15] For reasons of space, this study does not consider the interesting phenomenon of abstention or "roll off." Ignoring this consideration should not affect the estimates of how ballot structure influences approval rates since approval rates are net of participation issues.

Figure 2. %Yes by Ballot Position

Texas Propositons 1986-2015



Note. The figure plots all 240 propositions that appeared on the Texas ballot during 1986-2015. Ballot positions in a given election were assigned randomly by election officials. Data were collected from official election reports published by the Texas Secretary of State.

The remaining regressions of Table 2 include a variable equal to the number of propositions on the ballot. Column (2) shows a positive relation between ballot position and votes in favor, again statistically insignificant, and a negative relation between ballot length and votes in favor. The coefficient -0.40 on ballot length implies that each additional proposition on the ballot is associated with 0.40 percent fewer votes in favor. This coefficient is different from zero at the 1 percent level of statistical significance. The regression in column (3) is the same as the regression in column (2) except that extreme values of the dependent variable are Winsorized at the 99th percentile; the coefficients are essentially unchanged.[16] The regression in column (4) explores sensitivity to a different outlier concern by establishing a maximum ballot position of #16.

[16] The particular cutoff is not important. For example, the reported coefficient on ballot position of 0.20 becomes 0.16 when Winsorizing at 90th percentile and 0.17 when Winsorizing at 95th percentile, neither statistically significant.

As Figure 2 shows, the number of propositions with positions greater than #15 is rare, so the column (4) specification reduces the chance that the few extreme positions are driving the result. The coefficient on ballot position remains positive and is now statistically different from zero at the 10 percent level, suggesting an advantage to appearing near the bottom of the ballot. The coefficient on ballot length is essentially the same.[17]

One concern with the regressions in columns (2)-(4) is that ballot position and ballot length are to some extent positively correlated for mechanical reasons. This could cause the ballot position coefficient to absorb ballot length effects, and vice versa. The regression in column (5) avoids this problem by including election-specific fixed effects; in this case the ballot position coefficient is estimated based on within-ballot variation, and thus is free from ballot length effects. As can be seen, the coefficient on ballot position remains positive, with similar magnitude as in other specifications, and without statistical significance.[18]

Following Augenblick and Nicholson (forthcoming), we can estimate how many elections outcomes would have come out differently if every proposition had appeared in the first position, that is, if no proposition suffered the consequences of being listed other than first. To do this, we calculate the implied change in each proposition's approval rate based on the coefficient on ballot position and the proposition's actual ballot position, and compare this to its margin of victory or defeat. Using the estimate in column (1), where appearing down the ballot is disadvantageous, only

[17] The positive coefficient is robust to alternative caps. For example, setting the maximum position at #10 gives a coefficient of .51 on ballot position; capping at position #20 gives a coefficient of 0.24.

[18] The regressions assume a linear relation between approval rate and ballot position, but equation (1) allows for any sort of nonlinearity. I estimated a variety of models with alternative specifications – for example, including second order terms and allowing for differential effects in the first position – and did not find robust evidence of order effects with these more complicated specifications either.

one of one of the 240 propositions would have gone from fail to approve if listed at the top of the ballot. Using the estimate in column (5) where appearing down the ballot is advantageous, no proposition would have gone from approval to failure if listed at the top instead of its current position. In contrast to Augenblick and Nicholson (forthcoming), which concludes that 6 percent of elections would have been different without ballot order effects, I find that ballot order was a determining factor in virtually no election outcomes in Texas.[19]

The statistically insignificant coefficients on ballot position do not mean that there are no order effects, only that we are unable to distinguish any potential effects from noise. To get a rough sense of what size effects are plausible with these data, we can add or subtract multiples of the standard error to the coefficients. For example, in regression (5) of Table 2, we can say that with 95 percent confidence, the true coefficient is between $0.20 \pm 1.96 \times 0.15 = -0.094$, given a range of -0.09 to 0.49. Thus, even if being listed down the ballot costs votes, it is highly likely that the cost is extremely low. In contrast, there appears to be a reliably negative relation between approval rates and ballot length.

[19] Given existing evidence, we can only speculate on the reasons for the discrepancy. It could be that elections are more competitive in San Diego County than Texas, so small effects are more likely to swing an election. Another possibility is that the Augenblick and Nicholson estimates overstate the consequences by considering not the effect of switching position within the block of propositions, but of jumping a proposition to the top of the entire ballot, ahead of all candidate elections; no propositions appear in such a position in their data – projecting outside the support of their data may give unreliable predictions.

5. California

A. *Methods and Data*

Recall that the core problem in estimating (1) is that V_i^* typically is not observable. The research strategy for the California data is to use pre-election opinion surveys to proxy for V_{it}^*. If survey responses do not depend on ballot position (more on this below), then we can assume they are generated according to:

$$(2) \qquad V_{it}^{SURVEY} = V_{it}^* + \lambda + \mu X_{it} + u_{it},$$

where V_{it}^{SURVEY} is the percentage of respondents who express support for proposition i, X_{it} is factors that cause survey responses to differ from underlying preferences, λ is a fixed survey "bias" (for example, pre-election polls in California systematically overstate support for propositions), and u_{it} is an error term that is independent across propositions.

The difference between election returns and pre-election survey results is denoted Δ and from (1) and (2) can be expressed as:

$$(3) \qquad \Delta_{it} = V_{it}^{ELECT} - V_{it}^{SURVEY} = \alpha - \lambda + \beta \cdot POS_{it} + \gamma Z_{it} - \mu X_{it} + e_{it} - u_{it}.$$

Then we can regress Δ on ballot position to recover the position effects without needing to know the electorate's underlying preferences. The estimate of β will be unbiased even if ballot position is determined by underlying preferences rather than being randomly assigned. Another advantage of

specification (3) is that it is not necessary to control for determinants of the vote choice itself because they are subsumed in the V_{it}'s.

A less formal way to think about this empirical strategy is that it uses pre-election survey information to reveal the "untreated" preferences on a proposition. This expressed preference is compared to the actual election outcome that has been "treated" with the position effect, and the difference is used to infer the treatment effect. A potential limitation of using pre-election survey data as a control is the possibility that preferences change between the time of the poll and the election, or that voters express different preferences in an opinion survey than they truly believe. However, to the extent that there are systematic biases in the survey, they will be absorbed into the intercept term, and will not confound inferences as long as they are not correlated with ballot position.

The core data consist of election returns, taken from *Statement of Vote*, published by the California Secretary of State, and pre-election survey data taken from the Field Poll, available at www.field.com, and the Field Research Data at UC-Berkeley at ucdata.berkeley.edu/data.php. If the Field Poll conducted multiple surveys on a particular proposition, I use data from the final survey, that is, the survey that was closest to the election. The Field Poll runs from 1958 to 2014. Of the 678 propositions that went before the voters during that time, Field Poll data are available for 242 of them.

The key variables V_{it}^{ELECT} and V_{it}^{SURVEY} are operationalized as yes votes as a percentage of all votes. Abstainers or, in the case of a survey, individuals who decline to state or otherwise fail to give an opinion in favor or against are ignored. The gap between the election outcome and pre-election survey is defined as $\Delta = \%Yes_{ELECT} - \%Yes_{Field\ Poll}$.

Summary statistics for California propositions are reported in Panel B of Table 1. The propositions in the sample are not representative of all propositions that appeared on the ballot because the Field Poll focuses on high profile or controversial propositions. Field Poll propositions are less popular than other propositions, with a mean vote in favor of 48.6 percent compared to 53.8 percent for all propositions that reach the ballot. Field Poll propositions are much more likely than the full set of propositions to be initiatives (70 percent versus 34 percent), and much less likely to be legislative proposals (27 percent versus 65 percent).

Table 1 shows that the final Field Poll before the election exceeds the percentage of favorable votes in the election by 6.0 percent on average. This indicates a systematic "bias" in the Field Poll, or put differently, a consistent tendency for support to deteriorate leading up to an election. Many election observers have noted that support for propositions tends to deteriorate over time; Table 1 provides a large-sample quantification of the effect. The deterioration may occur because proponents are usually the first to mobilize – they have to secure legislative approval or collect signatures – and their arguments are the first to reach the voters. As the campaign progresses, opponents become active and some initial support deteriorates in the face of counterarguments. This deterioration in support between the last survey and the election is not a problem for the identification exercise as long as deterioration is uncorrelated with ballot position.[20]

[20] While plausible, the absence of such a correlation is not self-evident. For example, one might suppose that propositions near the top of the ballot were organized earlier than propositions at the bottom of the ballot, and better organized propositions suffer less deterioration. I am not aware of any evidence on this; the absence of correlation in deterioration by position should be taken as a maintained assumption.

The empirical analysis assumes that Field Poll responses are not influenced by ballot position. This assumption would be questionable if the Field Poll asked voters about all propositions on the ballot in the exact same order that the propositions appeared on the ballot. That is not the case. As noted above, the Field Poll only examined 36 percent of the propositions that appeared on the ballot. Furthermore, only 14 percent of the surveys included all of the questions, and in 37 percent of the surveys the questions were not asked in the order in which they appeared on the ballot. For example, the 2002 general election featured seven ballot measures; the Field Poll asked about four of them in this order: Proposition 47 – 50 – 49 – 52.[21] The order on the ballot was 46-47-48-49-50-51-52. The survey contains omissions as well as reorderings and does not simply reproduce the order on the ballot.[22]

Finally, it is worth emphasizing that the Field Poll does not select issues at random. Rather, it chooses issues that are likely to be of interest to the public or policymakers. This should not create a bias in the coefficients of interest, but it does affect external validity. The California results should be seen as applying to relatively high profile issues. The Texas data in some sense fill in the picture by providing evidence on lower profile issues.

B. Findings

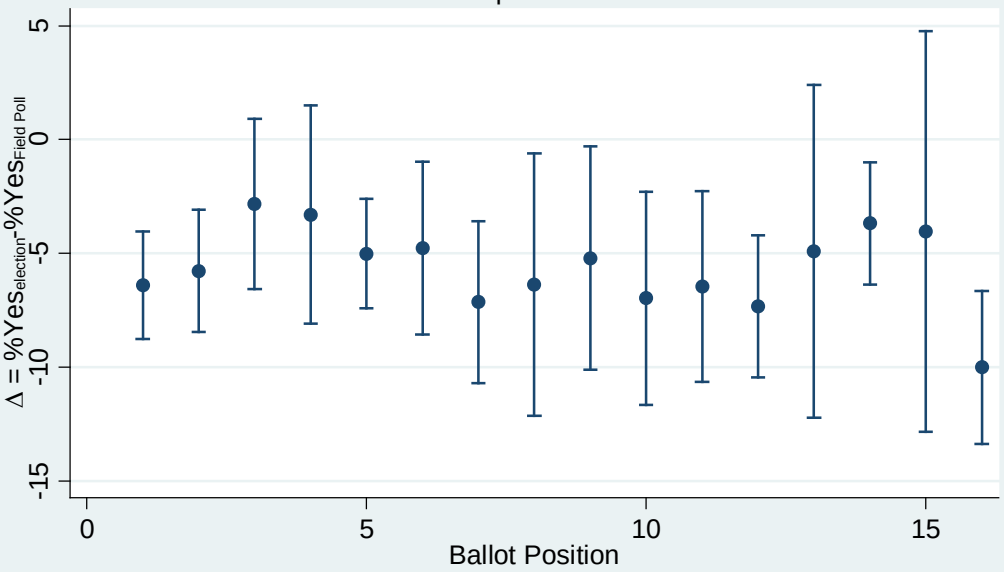
Figure 3 provides a characterization of the California data by plotting the mean gap by position, with 95 percent confidence intervals indicated. Positions greater than #15 are collapsed into a

[21] *The Field (California) Poll: Codebook 02-05*, questions Q19 to Q26.

[22] In 1994, the Field Poll conducted surveys on four ballot propositions in a randomized order. Those results are summarized in the appendix.

Figure 3. Mean Gap (Δ) by Ballot Position

California Propositions 1958-2014



Note. The figure plots the mean difference between the election approval rate and the pre-election Field Poll approval rate, by position. The bars indicate plus and minus two standard errors. Positions 16 and larger are combined into the position #16 category. The sample includes 242 propositions.

single group because of the scarcity of observations. The means do not show a consistent downward (or upward) pattern.

Table 3 reports statistical evidence from the California data: each column reports a regression of the gap, Δ, on ballot position, following equation (3). The coefficient on ballot position in column (1) is -0.07 (meaning that each position down the ballot reduces approval by 0.07 percent), quite small and far from statistical significance. Using the standard error, the 95 percent confidence interval for the true effect is -0.31 percent to 0.17 percent, which allows the possibility of nontrivial negative effect at the boundary. A similar pattern appears for the Winsorized specification in column (2) and the capped position specification in column (3). The regression in column (4) includes election-specific fixed effects. The coefficient is even smaller, -0.03, and again far from statistical significance. The California data, like the Texas data, offer little reason to believe that propositions benefit from being listed first.

The California data also produce the pattern in the Texas data that propositions on longer ballots receive fewer votes in favor, independent of the proposition's own ballot position. In column (1) of Table 3, each additional proposition on the ballot reduces the approval rate by 0.30 percent on average, a relation that is statistically significant at the 1 percent level, and similar in magnitude to what appears in the Texas sample. The coefficient on ballot length is negative and statistically different from zero in regressions (2) and (3) as well.

Another control variable is a dummy equal to one if the proposition was an initiative, as opposed to a proposal from the legislature or a referendum. Initiatives might be expected to attract more attention before the election, and therefore show less of a gap between election approval and pre-election approval. This turns out not to be the case: the coefficient on the initiative dummy suggests a larger gap for initiatives, although the coefficient is not distinguishable from zero at conventional levels of significance in any of the regressions. The final control variable is also related to information conditions, a dummy equal to one for presidential election years. One could argue that voters pay more attention to politics in presidential election years, and thus are more informed, or conversely, that a presidential election draws voters to the polls who are uninformed about ballot propositions. The data show a significantly higher gap in presidential election years, indicating that the election approval rates are 2.3 to 2.4 percent higher than indicated by opinion surveys in presidential election years.

I also estimated but do not report regressions under a variety of alternative specifications in order to assess robustness of the findings. Alternatives included: allowing a separate effect for the first position and for the last position; including higher order terms for ballot position; including time dummies; including dummies for general as opposed to primary elections; including

dummies for bond propositions and for referendums; including controls for the fraction of undecided voters; and alternative Winsorization cutoffs. For all of these alternatives, it continued to be the case that there was no reliable relation between approval rates and ballot position.

6. Discussion and Conclusion

State and local governments in the United States, and increasingly abroad, rely on ballot propositions to resolve important public policy issues. More than 1,800 state-level propositions have come before voters in the 21st century alone, addressing high-profile and high-impact issues such as same-sex marriage, marijuana legalization, taxes, and spending. The number of issues appearing in counties, cities, and towns is at least an order of magnitude larger, and equally diverse. With citizen lawmaking playing a central role in American democracy, it is important to assess the process by which these decisions are made, and identify mechanisms that might lead to distortions in the decisions. One potential distortion – the order in which issues are presented to the voters – has long concerned practicing politicians, many of whom believe that being listed at the top of the ballot is advantageous, and this belief has influenced the design of state election laws. And one recent study focusing on a county in California suggests that such distortions often lead to the wrong side winning (Augenblick and Nicholson, forthcoming). Yet research on the effect of ballot structure in proposition elections remains scarce, and seldom allows causal inference.

This paper contributes to the discussion by proposing and implementing two different empirical strategies, both of which are designed to separate causality from correlation. One strategy is to examine Texas propositions since 1986 when the state began to place propositions on the

ballot in a random order. The other strategy is to use pre-election survey data from California, which has a long history of polling on ballot measures, to control for public opinion independent of ballot position. Both approaches fail to turn up robust evidence in support of the idea that propositions attract more favorable votes when listed at the top of the ballot (or any other position) than they would have if listed elsewhere on the ballot. Because the evidence comes from two rather different states and two different information environments – low profile measures in Texas in off-year elections versus high profile issues in California – the absence of evidence for order effects suggests the findings may have some generality. While it is difficult to prove a negative, and variance around the estimates leaves some room for the possibility of order modest order effects, the most natural conclusion from the dearth of evidences across the different environments is that ballot order effects are at best rather small.

I do find a robust negative relation between approval rates and ballot length. Each additional proposition on the ballot is associated with about 0.3 percent lower approval for all propositions on that ballot. These estimates are similar across various specifications, and for both California and Texas. The research design does not allow strong claims about causality with respect to ballot length; we cannot rule out the possibility that long ballots have more inherently unpopular propositions. The reason for lower approval on longer ballots is an open question. Ballot fatigue is a natural candidate, but failure to find order effects casts some doubt on this explanation. Another possibility is that voters simply dislike long ballots, and adopt a negative orientation when asked to resolve a large number of issues. Formulating tests that allow stronger causal inference seems to be a useful direction for future research on ballot length.

The policy implications of these findings are nuanced. In terms of providing a level playing

field, it appears one should not be overly concerned with order manipulation because the top of the ballot is not demonstrably better than the bottom. Even so, there is no obvious downside to randomizing ballot position, so it seems a useful precaution. The evidence of lower approval rates on long ballots, if interpreted as a causal effect, calls for some attention concerning long ballots. From the perspective of quality of public decisions, one would like voters to cast their ballot when informed. However, there is no evidence whether approval rates are too high or too low for ballot propositions; it could be that a status quo bias with to regard to new proposals is healthy. Moreover, an attempt to shorten ballots would mean that fewer issues reach the voters. Any benefits from higher approval rates on short ballots would have to be balanced against the downside of curtailing the number of public issues that voters are allowed to decide.

Appendix

The Field Poll conducted a randomized controlled experiment in its survey for the June 7, 1994 primary election. I report this experiment here to bring it to the attention of researchers in the area. The Field Poll exercise is similar to the experiment reported in Binder and Kousser (2014). Four bond propositions were on the ballot, Propositions 1A, 1B, 1C, and 180. Each proposition authorized a bond issue for a different purpose (seismic retrofit, K-12 schools, higher education, or parklands). The Field Poll asked respondents if they expected to vote for or against each proposition. Half of the respondents were asked the questions in a random order, and half were asked the questions in the order they were to appear on the ballot (1A-1B-1C-180). This experiment presents an interesting opportunity to check for order effects because of the availability of a clear "order-free" benchmark; its main limitation is that it does not involve actual election votes and cognitive processes might be different when speaking to a pollster than when in the voting booth.

Table A summarizes the responses. In the randomized sample, column (1), the highest pre-election approval rate was 72.9 percent for Proposition 1A and the lowest was 59.3 percent for Proposition 180. Column (2) reports responses when questions were asked in the order they were to appear in the ballot. If the top of the ballot is a favored position, the gap (Δ) between the approval rate with the actual order and with the randomized order should decline moving down the ballot. There is no evidence for such a pattern. Column (3) reports approval rates in the actual election. As is common, overall support eroded substantially between the survey date (early April) and the actual election (early June). The Δ between the approval rates in the election and the randomized order survey does not show a convincing pattern of declining moving down the ballot.

References

Alvarez, R. Michael, Betsy Sinclair, and Richard L. Hasen, "How Much Is Enough? The 'Ballot Order Effect' and the Use of Social Science Research in Election Law Disputes," *Election Law Journal*, 2006, Vol. 5(1), 40-56.

Augenblick, Ned and Scott Nicholson, "Ballot Position, Choice Fatigue, and Voter Behavior," *Review of Economic Studies*, forthcoming.

Baumeister, Roy F., Ellen Bratslavsky, Mark Muraven, and Dianne M. Tice, "Ego Depletion: Is the Active Self a Limited Resource?," *Journal of Personality and Social Psychology*, 1998, Vol. 74(5), 1252-1265.

Binder, Michael and Thad Kousser, "First Come, First Served? Experiments on Ballot Order in Direct Democracy Elections," Working Paper, University of North Florida and UC-San Diego, 2014.

Bowler, Shaun and Todd Donovan, *Demanding Choices: Opinion, Voting, and Direct Democracy*, Ann Arbor: University of Michigan Press, 1998.

Bower, Shaun, Todd Donovan, and Trudi Happ, "Ballot Propositions and Information Costs: Direct Democracy and the Fatigued Voter," *Western Political Quarterly*, June 1992, Vol. 45(2), 559-568.

California Secretary of State, *A Study of Ballot Measures: 1884-1980*, Sacramento, CA: 1981.

Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso, "Extraneous Factors in Judicial Decisions," *Proceedings of the National Academy of Sciences*, 011, Vol. 108(17), 6889-6892.

Ho, Daniel E. and Kosuke Imai, "Estimating Causal Effects of Ballot Order from a Randomized Natural Experiment," *Public Opinion Quarterly*, Summer 2008, Vol. 72(2), 216-240.

Iyengar, Sheena S. and Emir Kamenica, "Choice Proliferation, Simplicity Seeking, and Asset Allocation," *Journal of Public Economics*, 2010, Vol. 94(7-8), 530-539.

Levav, Jonathan, Mark Heitmann, Andreas Herrman, Sheena S. Iyengar, "Order in Product Customization Decisions: Evidence from Field Experiments," *Journal of Political Economy*, 2010, Vol. 118(2), 274-299.

Lupia, Arthur and John G. Matsusaka, "Direct Democracy: New Approaches to Old Questions," *Annual Review of Political Science*, 2004, Vol. 7, 463-482.

Matsusaka, John G., "In Search of Ballot Order Effects in Proposition Elections," Working Paper, University of Southern California, August 2013.

Meredith, Marc and Yuval Salant, "On the Causes and Consequences of Ballot Order Effects," *Political Behavior*, 2013, Vol. 35(1), 175-197.

Miller, Joanne M. and Jon A. Krosnick, "The Impact of Candidate Name Order on Election Outcomes," *Public Opinion Quarterly*, Autumn 1998, Vo. 63(2), 291-300.

Mueller, John E., "Voting on Propositions: Ballot Patterns and Historical Trends in California," *American Political Science Review*, December 1969, Vol. 63(4), 1197-1212.

Selb, Peter, "Supersized Votes: Ballot Length, Uncertainty, and Choice in Direct Legislation Election," *Public Choice*, June 2008, Vol. 135(3/4), 319-336.

Table 1. Summary Statistics for Texas and California Propositions

Panel A. Texas Propositions, 1986-2015

Variable	Mean	S.D.	Min	Max
%Yes (election)	63.4	12.6	30.2	93.8
Position	8.5	5.9	1	27
Number of propositions on ballot	16.0	6.4	1	27

Panel B. California Propositions, 1958-2014

	Field Poll sample ($N = 242$)				All props ($N = 678$)
Variable	Mean	S.D.	Min	Max	Mean
%Yes (election)	48.6	12.5	13.3	74.2	53.8
%Yes (Field Poll)	54.5	13.5	19.1	89.0	NA
%Yes (election) − %Yes (Field Poll)	-6.0	8.1	-30.4	14.3	NA
Position	8.5	5.8	1	29	7.8
Number of propositions on ballot	14.1	7.1	1	29	14.6
Type = legislative bond measure	0.23	0.42	0	1	0.22
Type = initiative	0.70	0.46	0	1	0.34
Dummy = 1 presidential election year	0.51	0.50	0	1	0.47

Note. This table reports summary statistics for the Texas (Panel A) and California (Panel B) data. Texas data include 240 propositions. All Texas propositions were constitutional amendments placed on the ballot by the legislature. California data include 678 propositions, 242 of which were surveyed in the Field Poll. Panel B reports the mean approval rate for Field Poll propositions and for all propositions (those included and excluded from Field Poll) during the period.

Table 2. Regressions of %Yes on Ballot Position, Texas Propositions 1986-2015

	(1)	(2)	%Yes Winsorized (3)	Position Capped at #16 (4)	Election Fixed Effects (5)
Position	-0.03 (0.14)	0.20 (0.16)	0.20 (0.16)	0.32* (0.19)	0.20 (0.15)
Number of propositions on ballot	…	-0.40*** (0.15)	-0.39*** (0.15)	-0.42** (0.15)	…
Intercept	64.1*** (1.4)	68.5*** (2.2)	68.4*** (2.2)	68.1*** (2.2)	…
R^2	.001	.028	.028	.034	.272

Notes. Each column reports estimates from a regression in which the dependent variable is the percentage of votes in favor of a proposition. In column (2), the dependent variable is Winsorized at the 99th percentile. In column (3), ballot positions above 16 are restated as 16. The regression in column (4) includes election-specific fixed effects. The data include all 240 Texas ballot propositions during 1986-2015. Significance levels are indicated: * = 10 percent, ** = 5 percent, *** = 1 percent.

Table 3. Regressions of Gap Between Election and Survey %Yes, California Propositions 1958-2014

	(1)	Δ Winsorized (2)	Position Capped at #16 (3)	Election Fixed Effects (4)
Position	-0.07 (0.12)	-0.06 (0.11)	-0.07 (0.14)	-0.03 (0.12)
Number of props on ballot	-0.30*** (0.09)	-0.30*** (0.09)	-0.31*** (0.08)	…
Dummy = 1 if initiative	1.44 (1.20)	1.31 (1.15)	1.41 (1.21)	-1.19 (1.35)
Dummy = 1 if presidential election year	2.44** (1.01)	2.31** (0.97)	2.42** (1.01)	…
Intercept	-3.40** (1.37)	-3.32** (1.32)	-3.29** (1.36)	…
R^2	.098	.102	.097	.464

Note. Each column reports a regression in which the dependent variable is Δ (or some variant thereof as indicated at the top of each column), where $\Delta \equiv \%Yes_{ELECT} - \%Yes_{Field\ Poll}$. In column (2), the dependent variable is Winsorized at the 95th percentile; in column (3), ballot positions above 16 are restated at position 16. The regression in column (4) includes election fixed effects. The data include 242 California propositions during 1958-2014. Significance levels are indicated: * = 10 percent, ** = 5 percent, *** = 1 percent.

Table A. %Yes from Field Poll with Randomized Question Order, California 1994

Proposition	Ballot Position	Question Order = Randomized (1)	Question Order = 1A-1B-1C-180 (2)	Election Results (3)
1A (retrofit bonds)	#1	72.7	72.7 (Δ= −0.2)	45.7 (Δ= −27.2)
1B (K-12 bonds)	#2	68.0	69.8 (Δ= 1.8)	49.6 (Δ= −18.4)
1C (higher education bonds)	#3	60.5	58.8 (Δ= −1.7)	47.4 (Δ= −13.1)
180 (parkland bonds)	#9	59.3	64.2 (Δ= 4.8)	43.3 (Δ= −16.0)
Respondents/Voters		416	416	4,966,827

Note. The election took place June 7, 1994. Field Poll data are taken from *The Field Institute/The California Poll – 9403*, administered April 1 to April 9, 1994. Columns (1)-(3) report approval rates, defined as votes in favor as a percentage of votes in favor plus votes against. Δ in column (2) is the column (2) approval rate minus the column (1) approval rate; Δ in column (3) is the column (3) approval rate minus the column (1) approval rate.