

Jumping on the Bandwagon after the Election? Testing Alternative Theories of Vote Share Overestimation for California Ballot Initiatives

David Crow
Survey Research Center
University of California, Riverside

Shaun Bowler
Department of Political Science
University of California, Riverside

Martin Johnson
Department of Political Science
University of California, Riverside

Abstract

Post-election poll results typically overstate the proportion of people who voted for winning candidates at all levels of government—and, now, citizen ballot initiatives. In polls on the May 2009 California special election, the percentage of respondents claiming in a post-election poll to have voted for the winning side was invariably higher than both the pre-election poll and the actual vote. Using original data, I test four alternative explanations of this “post-electoral survey bandwagon” effect. First, respondents may misrepresent how they voted to save face (“social desirability”). Second, they may genuinely forget how they voted (“memory lapse”). Third, opinion may have shifted dramatically between the last poll taken before the election and the moment of polling (“last minute opinion shift”). Fourth, inflated vote percentages may occur because a greater proportion of people who voted for the winning side took than the survey than did people who voted for the losing side (“non-response bias”). This paper devises empirical tests to choose between these hypotheses. We find that, rather than misrepresenting their votes to poll-takers, many people who voted for the losing side of the proposition contests simply do not take polls.

Paper prepared for presentation and the 2010 annual meeting of the Midwest Political Science Association, Palmer House Hilton Hotel, Chicago, Illinois, April 22-25, 2010.

Introduction

On May 19, 2009, California voters took to the polls in a special election to vote on a package of six ballot propositions that sought to redress the \$23 billion fiscal deficit scourging the state. Californian electors roundly rejected five of the proposals (which contained, collectively, a mixture of spending cuts and fee hikes) and even more overwhelmingly approved the sixth, forbidding legislative pay raises in years of budget deficits. The margins separating the “yes” from the “no” votes on the first five propositions were formidable, ranging from 24 to 33 percentage points. The margin for the sixth measure was nearly 49%.

Wide though these margins of victory were, an original post-election poll the “2009 California Special Election Survey,” carried out by the University of California, Riverside, Survey Research Center) inflated them further. Poll respondents appear to have jumped on the electoral bandwagon after the election results were in. That is, they appear to have exaggerated considerably the extent to which they voted for the winning side of the proposition contests—by as much 17%. Why?

We investigate the overestimation¹ of vote shares for electoral contest winners focusing on ballot initiatives. Post-election poll results routinely overestimate the proportion of respondents who both report having voted, and who report having voted for the winner. Explanations for this typically focus on misreporting—that is, on respondents’ inaccurately answering survey questions. People claim they voted when they really did not, and they claim to have voted for the winner when they really voted for someone else. Faulty memory may explain

¹ We prefer the term vote “overestimation” to the more widely used “misreporting” or “overreporting.” “Misreporting” implies that respondents supply inaccurate information when taking a survey. But overestimation can also result from overrepresentation of citizens who voted for the winner in the sample. We therefore use “overestimation” in the general case where survey estimates are higher than actual vote share, and “overreporting” in the specific case where overestimation results from respondents’ inaccurately claiming to have voted for the winner.

some inaccurate reporting, but most scholars attribute inaccurate reports to respondents' deliberate dissembling of voting behavior to conform to perceived norms of "social desirability."

We explore alternative explanations for vote overestimation in which people report their vote honestly. One such explanation is a massive change in public sentiment that occurs too late for detection by the last pre-election poll taken, but is reflected in post-election polls. We find a likelier explanation, however, in non-response bias: people who voted for the winning side of an election are likelier to take a post-election poll than people who voted for the losing side. Thus, respondents are telling the truth, but the post-election poll winds up overestimating the winning side's vote share anyway.

The heart of our strategy for choosing between competing hypotheses—particularly between social desirability and non-response bias—is to estimate a model of vote preference for pre-election survey respondents. Then, we plug the post-election data into the pre-election model coefficients to estimate probabilities that a post-election respondent will report having voted for the winning side. Comparing predicted to reported voting behavior will allow us to choose between the two main hypotheses. If the predicted vote differs significantly from the reported vote and there are many classification errors, respondents on the whole exaggerated the extent to which they voted for the winning side—consistent with the "social desirability" hypothesis. On the other hand, few classification errors resulting from consonance between predicted and reported voting behavior militates in favor of non-response bias, in which respondents report voting behavior honestly, but people who voted for the losing side opt not to take the survey.

We pursue a second strategy for adjudicating between the hypotheses: estimating a Heckman selection model, in which coefficients in a model of vote preference are adjusted by a

model of survey response. If non-response bias is to blame for vote share overestimation, the Heckman-adjusted coefficients of the vote share model should differ significantly from the non-adjusted coefficients. If, on the other hand, social desirability is the culprit the Heckman selection model should have no perceptible effect on the coefficients for the vote share model.

This research is important because the mechanism underlying vote overestimation has important implications for choosing the best post hoc statistical adjustment to remedy the problem. Moreover, we hope to shed light on the psychological motivations of potential survey respondents not only in choosing whether to answer truthfully or not, but in choosing whether or not to take the survey in the first place. Rather than unjustly maligning survey respondents for giving dishonest but “socially desirable” responses, survey researchers should turn a critical eye on our own failure to ensure that all respondents are well represented in our sample.

The May 19 Special Election: The Initiatives, Election Results, and Vote Overestimation

The Ballot Initiatives

Facing a budget shortfall of tens of billions of dollars, the California Legislature approved a series of stopgap budget measures, signed into law by Governor Arnold Schwarzenegger, in February, 2009. Since the measures proposed amending the constitution or modifying laws previously enacted through referenda, the California State Constitution requires that the legislature submit them directly to voters in new referenda (Article II, Sections 8 and 10). So, the bill scheduled a special election for May 19, 2009, and called on California electors to vote on six ballot initiatives.

The first five propositions contained a highly complex, confusing mix of fiscal provisions that hiked some taxes and fees, and cut spending in some areas while maintaining it in others.

These measures also threw in some accounting sleight of hand that reallocated funds among budget categories (Props 1B and 1D) and took some unfunded liabilities off the books by shifting them out of the general budget (Prop 1E). The sixth proposition was a sop to the rampant anti-politician mood that had settled over the state; it curtailed legislative pay raises.

Proposition 1A sought to extend income tax and vehicle fee increases approved the previous year through 2013, and channel revenue into the Budget Stabilization (or “rainy day”) Fund to hedge against future crises. Proposition 1B (which depended on passage of Prop 1A to take effect, though enactment of Prop 1A was not conditioned on approval of Prop 1B) would have released \$9.3 billion dollars from the Budget Stabilization Fund to schools over six years to close the education spending gap between statutorily mandated expenditures and actual allocations. Proposition 1C authorized the state to borrow \$5 billion dollars against future lottery revenues to reduce the budget deficit. Proposition 1D would have diverted \$608 million dollars in cigarette tax revenues from early child development programs into the general budget in 2009. Proposition 1E proposed shifting \$226.7 million dollars from state mental health services to a the federally-mandated mental health initiative (the Early Periodic Screening, Diagnosis and Treatment program), offsetting liabilities that would otherwise be paid for out of the general fund. Finally, Proposition 1F prevented state authorities in charge of determining compensation for public servants from raising state officials’ salaries in deficit years.

The Special Election: Context and Results

Deep citizen disgust with politicians framed the May special election. Arduous legislative negotiations, complicated by California’s two-thirds supermajority requirement to approve budgets, had delayed approval of the 2008 budget by a record three months, to September. No sooner had the budget passed than further projections of revenue shortfalls led

the state comptroller to pay state debt with IOUs in January, 2009. Governor Schwarzenegger declared a fiscal emergency and called a special legislative session in February, which produced the series of ballot propositions put before the voters in May. Little wonder, then, that citizen approval of politicians in May, 2009, was around 14%, where it had hovered at an all-time low since September, 2008. The UC Riverside Survey Research Center's "2009 California Special Election Study" survey concurred. The mean approval for Governor Schwarzenegger was 4.13 (on a scale of 1 to 10, with 18% of respondents giving him a positive rating, above the scale's midpoint) and for the legislature, 3.0 (with 11% giving the legislature a positive rating).²

Voters roundly rejected the first five propositions, which addressed the budget deficit. The "Yes" vote for Prop 1A was 34.6% (with a "No" vote of 65.4% for a difference of 30.8%); for Prop 1B, 38.1% ("No", 61.9%; difference of 23.8%); for Prop 1C, 35.6% ("No", 64.4%; difference of 28.8%); for Prop 1D, 34.0% ("No", 66.0%, difference of 32.0%); and for Prop 1E, 33.5% ("No", 66.5%; difference of 33.0%). On the other hand, the anti-politician climate led voters to approve overwhelmingly the sixth ballot initiative, Prop 1F, which forbids raises for legislators in deficit years. It was approved 74.3% to 25.7% (a difference of 48.6%).

The Data: The 2009 California Special Election Survey

The UC Riverside Survey Research Center fielded the "2009 California Special Election Survey" from May 11 to May 24, 2009—before and after the May 19 election—as part of a project for the "Mass Media and Public Opinion" class that one of the authors taught. Under the supervision of an instructor, students designed the questionnaire, collected data in the Survey Research Center's call center, and produced a database that they analyzed for the final paper.

² Figures taken from the April 1, 2009 (<http://field.com/fieldpollonline/subscribers/RIs2307.pdf> [accessed on March 30, 2010]) and September 12, 2008 (<http://field.com/fieldpollonline/subscribers/RIs2284.pdf> [accessed on March 30, 2010]) Field Polls.

Survey participants constituted a simple random sample (SRS) drawn from California voting registration records. Between May 11 and 18, in the pre-election portion of the survey, 169 registered voters took the survey; and between May 20 and 24, the post-election portion, 107. (We suspended data collection the day of the election, May 19.)

The survey asked about intention to vote—or, in the post-election poll, reported vote—for four of the propositions, 1A, 1B, 1D, and 1F. It also asked respondents to evaluate Governor Schwarzenegger’s and the Legislature’s performance. It elicited opinions on salient political issues, including California’s two-thirds supermajority requirement for passing budgets, use of citizen ballot initiatives for budgeting, cuts in education spending, and legalization of marijuana. And it gathered some basic sociodemographic information.

The sample records provided by the on-line sample supplier, Aristotle, also included a number of background variables such as prior voting history, geographical information (including county and election districts), income, occupation, ethnicity, education, religion, and others. These variables are culled both from voting registration records and other sources, such as other government records, credit bureau reports, and commercial information compilers. Having information on both respondent and non-respondent characteristics allows us to develop a model of survey response, which we use to assess the possibility that non-response bias at least partly explains vote preference overestimation.

Vote Share Overestimation

Post-election polls typically overestimate the proportion of voters who voted for the winning candidate in an election. The 2009 California Special Election Survey is no exception. Table 1 compares the percentage of respondents who reported voting “Yes” on the propositions

after the election (column labelled “Post”, followed by the number of post-election poll respondents, “N-Post”) with the percentage of respondents who stated, before the election, that they intended to vote “Yes” (“Pre”, followed by the number of respondents in pre-election poll, “N-Pre”) and with the actual vote share (“Actual”) for propositions 1A, 1B, and 1D.³ In all three cases, the post-election poll reports higher vote shares for the winning side than the pre-election poll does. The post-election poll overestimated the winning vote share by 11.2% (relative to the pre-election poll) for Prop 1A, 8.5% for Prop 1B, and 6.9% for Prop 1D. In two cases, Prop 1A and Prop 1B, the differences were significant at $p < .10$ (one-tailed test, reported in the column “p^a”, to the right of “N-Post”), and in a third case, Prop 1D, the difference approached statistical significance at $p = .10$.

Comparing the post-election poll to the actual election results reveals even more dramatic, significant differences. The post-election poll overestimated winning vote share (relative to actual results) by 14.6% for Prop 1A, 13.1% for Prop 1B, and 16.9% for Prop 1D. All these differences are statistically significant at $p < 0.01$ (as reported in the last column, “p^b”). In short, the 2009 California Special Election Survey results suggest more people voted for the winning of ballot initiatives than actually did.

Explaining Vote Overestimation: Misreporting or Non-Response Bias?

The Extent of the Problem

Studies that seek to explain why people choose to vote or not have long noted that post-election polls routinely overestimate the percentage of people who report having voted (Belli *et al.* 1999). Wolfinger and Rosenstone found that in the American National Election Study

³ We omit Prop 1F from this analysis. Though the post-election portion of the 2009 California Special Election Survey overestimates the Prop 1F winning vote share relative to the pre-election portion by 5.5%, it *underestimates* winning vote share by 9.8% ($p = 0.27$) relative to the actual vote.

(ANES), taken after every presidential and mid-term election since 1948, the percentage of respondents who report having voted is always between 5% and 20% higher than official turnout figures provided by the Federal Electoral Commission (1980: 115). The gap in the 1984 presidential election was 18.3% and in 1988, 19.1% (Deufel and Kedar 2000: 24). Based on “validated vote data,” which compare self-reported voting behavior on post-electoral surveys to actual voting records maintained by county registrars, Silver *et al.* note that reported turnout exceeded actual turnout by 27.4% in 1964, 31.4% in 1976, 22.6% in 1978, and 27.4% in 1980 (1986: 613). In the post-electoral portion of the 2009 California Special Election Survey, 73.9% of respondents claimed to have voted; statewide turnout was 28.4%.

Similarly, studies that seek to explain why people vote as they do have also noted that post-election polls overestimate the percentage of people who report having voted for the winning candidate (Wright 1990). Averaging over ANES studies since 1952, Wright found that the “pro-winner” bias was 4.0% in U.S. Senate races, 4.7% in gubernatorial contests, and (between 1978 and 1988) 7.0% in races for the U.S. House of Representatives (1993: 295). Also using ANES data, Eubank and Gao demonstrated a disparity of 14.0% between the average survey-reported vote share for incumbents in House races, 78.8%, and their average share on ballot returns, 64.8% (1984: 224). Atkeson shows that post-election survey vote overestimation also obtained in presidential primary races between 1972 and 1992, where overestimation for the eventual nominees averaged 15.2% for Democrats (reaching as high as 27.1%)

Both turnout and vote share overestimation are problematic. When turnout and vote share are dependent variables in a regression analysis, their overestimation biases point estimates if the determinants of overestimation overlap with those of turnout and vote choice. For example, turnout studies consistently highlight the link between educational attainment and

voting. But Silver *et al.* found that “high-status” respondents people, including the well educated, who *should* vote but don’t are likelier to overreport voting than their “low-status” counterparts (1986: 615). Since education is correlated with both voting and vote overreporting, it is possible that studies have overestimated the effect of education on electoral turnout.

Where turnout and vote preference are independent variables, their overestimation biases effect estimates upward. Atkeson points out, for example, that pro-winner bias in primary election polls may overstate the extent to which primary vote choice predicts vote choice in the general election (1999: 209). Studies on the “divisive primary effect” (see, e.g., Cantor 1996, Southwell 1986, 1994), in which supporters of losing primary candidates vote defect by voting for candidates from another party in the general election (or abstaining), may exaggerate this effect’s magnitude. Some respondents who report voting for eventual primary winners in both the primary and general elections, in fact, voted for a losing candidate in the primary. Therefore, “divisive primary” studies may underestimate the degree to which voters for losing primary candidates eventually rally behind the party nominee in the general election.

Overreporting and Memory Lapse

Scholars seeking to account for post-election surveys’ overestimation of both turnout and winning candidate vote share have focused overwhelmingly on *misreporting*—or, more precisely, *overreporting*—of electoral behavior. Survey respondents *overreport* when they inaccurately claim to have voted (when they did not) and to have voted for the winning candidate (when they voted for someone else). Overreporting has its roots in social psychology. It occurs either because respondents *misremember* whether and how they voted, or because they *deliberately misrepresent* having voted and who they voted for to a survey interviewer.⁴

⁴ Some studies have argued that overreporting is partly an artifact of the survey instrument itself. Eubank and Gow (1983) point to question order effects as the culprit for overreporting. The ANES’s placement of the question on

“Source-monitoring” psychological theory traces remembered events back to the perceptual and cognitive sources give rise to them. It posits that “source confusion” results from real and imagined events’ sharing overlapping characteristics (Johnson *et al.* 1993: 4-5). In survey research, source confusion could cause respondents to recall their voting behavior inaccurately if, for example, they conflate the intention to vote with actually voting, or of weighing the merits of a given candidate with casting a ballot for that candidate. Respondents may also “forward telescope a remote voting experience”, transforming prior votes into a vote cast in the last election (Belli 1999: 91).

Wright argues that memory failure afflicts less sophisticated voters more. They are more susceptible than sophisticated voters to shifts in public perceptions in the time that intervenes between the election and the survey. When they cannot reconstruct how they voted accurately, they substitute judgments at the time of the survey for those at the time of the election, and may claim to have voted for someone other than the candidate for whom they really voted (Wright 1993: 293).

Studies have also established that misreporting increases the later the post-election survey is taken after the election. Using validated voting data in an Oregon study, Belli found that an experimental question wording designed to prod respondents’ memories increased reliability of self-reported voting data for surveys carried out later in the data collection period. The authors inferred that misreporting increased the more time had elapsed between the election and the survey (1999: 99). For her part, Atkeson noted that memory failure was an especially important

vote choice after questions that identified incumbents slanted responses in favor of more recognized, incumbent candidates. For their part, Jacobson and Rivers (1993) argued that the form of the question itself induces respondents to report voting for winning candidates and incumbents. The switch in 1978 from an open-ended question about whom respondents voted for to one that provided name and party cues boosted reported vote shares for incumbents. Explanations of vote share overestimation based on the survey instrument rely on name recognition of winning candidates. They therefore do not obtain in this study, which deals with ballot propositions.

explanation for vote misreporting given the large number of days between the primary election and the National Election Survey, taken after the general election (1999: 205). We do not believe memory lapse to be a likely explanation for overestimation of winning vote shares in the 2009 California Special Election Survey, simply because not enough time elapsed for respondents to misremember their vote.

Overreporting and Socially Desirable Responses

If some survey respondents cannot recall their voting behavior accurately, others may dissemble it on purpose. “Social desirability” theory postulates that respondents are loath to admit attitudes or behaviors contrary to those sanctioned by society (Noelle-Neuman 1993). Most of us partake of a natural desire to please others—even interlocutors with whom we have only fleeting contact—and are concerned with appearing to be responsible citizens. These behavioral imperatives inhere, virtually unmodified, even in the context of the survey interview. Motivated by “reasons of self-presentation,” survey respondents conform to norms they perceive as socially desirable and give the answer they think survey respondents want to hear (Presser 1990). So, post-election poll respondents who recollect their voting behavior accurately rightly may nonetheless claim to have voted though they did not, and to have voted for a winning candidate though they voted for a loser, since the electoral outcome invalidates their personal preference (Atkeson 1999, Belli 1999). “The result,” writes Atkeson, “is a bandwagon effect for the winner after the election” (Atkeson 1999: 203).

Non-Response Bias as an Alternative Explanation

Voting studies find in misreporting (whether accidental or deliberate) for their leading explanation for overestimation of turnout and winning candidate vote share. But another explanation, almost completely overlooked in the voting literature, is possible: *non-response*

bias. That is, overestimation may occur not because of respondents' overreporting but because citizens who abstained (or voted for losing candidates) refuse to take the survey in greater proportion than voters (and voters for winning candidates). Disproportionately high non-response among abstainers and voters for losing candidates results in overrepresentation of voters and voters for winning candidates and, consequently, overestimation of the percentages of citizens who voted, or voted for winning candidates. Abstainers are, as a rule, less interested in politics—and, presumably, in participating in political surveys—than voters. Citizens who cast their ballots for losing candidates, momentarily dispirited, may be disinclined to take a survey about an election whose outcome they find disagreeable. It is therefore possible, in theory, for overestimation to occur even when all respondents report their voting behavior truthfully—although it is likelier that misreporting and non-response bias contribute to overestimation in tandem.

Research on sampling and survey methodology has long grappled with the potential biases produced by non-response (see, e.g., Berinsky 2004, Cochran 1977: 359-64, Groves and Couper 1998, Groves 2002, Kish 1965). Practitioners distinguish between *item non-response*, in which respondents neglect to answer some (but not all) questions on a survey, and *unit non-response*, in which some people selected into the sample fail to take the survey altogether. We are concerned here with unit non-response—that is, with the possibility that voters for the losing side of an election decline to participate in a survey more frequently than voters for the winning side. Non-response bias occurs when the probability of responding to a survey is different for different segments of a population *and* there are significant differences between segments: “For the bias to be important, a large nonresponse must coincide with large differences between the

means of the two segments” (Kish 1965: 535). No bias will occur if non-response is random (i.e., distributed equally among subpopulations) or if subpopulations have the same means.

A number of techniques have been proposed to ameliorate non-response bias.⁵ The best way of minimizing non-response bias is to minimize non-response itself during data collection by persisting in attempts to locate difficult-to-reach respondents. Failing that, however, there are several *post hoc* remedies. One is reweighting the sample, giving more weight to underrepresented segments and less to overrepresented segments. Raghunathan notes that sample weighting has long been used to compensate for unequal probabilities of selection (2004: 105). When individual-level information is available on non-respondents, Heckman (1979) suggested developing a model for response propensities and using estimates from the response model to adjust coefficient estimates of the models for the substantive behavior that interests us. Brehm (1999) adapted the Heckman “sample selection” model to cases where the response probabilities have to be estimated from aggregate data.

Given the prominence that survey methodology accords treating non-response bias, it is surprising that voting studies fail to consider this potential explanation of overestimation of vote choice for the winning candidate. Some studies mention non-response bias *en passant*, but then pull back from a more thorough explanation of the possibility. For example, Atkeson shows that there was considerable pro-winner bias in the ANES 1988 Super Tuesday post-election poll. African American voters were underrepresented in the sample. Correspondingly, the survey results understated support for Jesse Jackson and overstated support for the eventual winner, Michael Dukakis. Post hoc weighting adjustments brought vote share estimates in line with actual results, which raises the possibility that non-response bias drove overestimation of

⁵ Techniques for dealing with item non-response, such as multiple imputation (Little and Rubin 2002), are generally inapplicable to unit non-response since they depend on respondents providing sufficient information to support reasonable guesses about the content of missing information.

Dukakis's vote share—and, conceivably, other results as well (1999: 207). Unfortunately, Atkeson failed to examine this suggestive result more systematically. In his study of presidential and congressional races, Wright avers that “hostile” respondents are not likely to misreport vote choice intentionally, but rather “would generally refuse to be interviewed in the first place” (1993: 293). However, Wright, too, neglected to develop this aside into a full-blown consideration of non-response bias as a possible cause of vote preference overestimation.

Late Opinion Shift

A second alternative explanation for overestimation that does not involve misreporting is the possibility that some voters had an eleventh-hour about face in voting intentions. If large numbers of people who supported the propositions changed their minds after the last pre-election interviews were carried out and voted against them, this could account for at least some of the sharp discrepancies observed between the pre- and post-election polls (and between the pre-election poll and actual voting results). Of course, the longer the pre-election poll is taken before the election, the less accurately it will predict election results (Crespi 1988). In this case, interviews were conducted right up to the day before the elections.

Hypotheses

Summarizing, we have identified four explanations for the 2009 California Special Election survey's overestimation of winning vote share:

- 1) **Memory lapse**, in which survey respondents, unable to recall how they voted, misreport their electoral preference.
- 2) **Social desirability**, in which respondents do recall how they voted but deliberately misreport their electoral preference because they are embarrassed to admit voting for the losing side.

- 3) **Non-response bias**, in which a survey sample overrepresents citizens who voted for the winning side because those who voted for the losing side fail to participate in the survey.
- 4) **Late opinion shift**, in which large numbers of voters change their minds too late for pre-election polls do not capture the shift, which is registered in the post-election survey.

Research on survey overestimation of support for the winner informs four hypotheses we test here. We anticipate that surveys overestimate support for winners as a function of intentional misrepresentation on the part of survey respondents to appear to be on the winning side of an election, respondents forgetting, genuine last minute shifts in the preferences of voters yielding seemingly inflated post-election support for the victor, and non-response bias.

Detecting evidence of genuine late shifts in voter preferences as well as memory lapses involves investigating the time course of opinion on the electoral question over the days of interviewing. If late shifts of opinion affect survey overestimation of winner support, we should see a trend across days leading up to the election in support for the winner, controlling for other attributes of survey respondents. Similarly, if people forget for whom they voted and then systematically misremember that they voted for the winner when they did not, then we should observe a trend in self-reported voting for the winning side of an election after Election Day.

Here, then, are formal statements of our hypotheses for the two hypotheses (one and four) involving expectations about reported vote choice over time, before and after the election:

Late Opinion Shift Hypothesis: Support for the winning side of an election will increase over time before Election Day, *ceteris paribus*. So,

$$\Pr(\text{Vote} = \text{Yes} | t \geq T) > \Pr(\text{Vote} = \text{Yes} | t < T) \forall t < 0$$

where t is the day on which the survey was taken, T is an arbitrarily fixed reference day, and 0 is Election Day.

Memory Lapse Hypothesis: Support for the winning side of an election will increase over time after Election Day, *ceteris paribus*.

$$\Pr(\text{Vote} = \text{Yes} | t \geq T) > \Pr(\text{Vote} = \text{Yes} | t < T) \forall t > 0$$

where t , T , and 0 are as above.

For its part, detecting evidence of post-election overreporting for the winning side—or overreporting’s flipside, non-response bias favoring the winner—implied, for us, a two-pronged approach. The first prong entails predicting vote preference in the post-election portion of the survey and comparing reported to predicted voting behavior. We base these predictions on a baseline, pre-election model of voting, assumed to be the true model of vote choice (or close to it),⁶ which we use to project votes in the post-election survey. Consonance between predicted and reported votes constitutes evidence of non-response bias. Divergence between these two—specifically, where reported “no” votes exceed predicted “no” votes—constitutes evidence of overreporting.

The invariance of logged odds ratios (the coefficients in the logistic regressions used to predict voting) to changes in marginal distributions implies that even when the composition of the post-election sample differs from that of the pre-election sample, the model will predict voting equally well in the pre- and post-election samples. For example, if the post-election survey sample contains a greater percentage of Republicans or of people who voted against a ballot proposition, the pre-election survey the reported vote share may change but the difference

⁶ We note that the pre-election survey is within the margin of error for Propositions 1A and 1B, and not too far outside it for Proposition 1D.

between the reported and predicted vote should not differ across the two surveys.⁷ That is, in a cross-classification of predicted vote choice by reported vote, the classification error produced by predicting a winning, “no” vote when a “yes” is recorded should not be greater for the post- than for the pre-election model.

On the other hand, greater classification error in the post-election survey than in the pre-election survey constitutes evidence of overreporting. A proportion of votes for the winning side in excess of that predicted by the baseline model (taking into account that the baseline model does not predict votes with complete accuracy) indicates that survey respondents are, on average, exaggerating the extent to which they voted for the winning side of survey, the higher the likelihood of reporting a winning, “no” vote.

The second prong of our approach to ferreting out overreporting and non-response bias is to estimate a sample selection model of the type proposed by Heckman (1979). Here, we estimate two models, an “outcome” model of the determinants of support for the ballot propositions and, using background information we have for both respondents and non-respondents, a “selection” model of the determinants of survey participation. Under well-known results of simultaneous equation theory, if the errors of the outcome and selection models are correlated the unequal probabilities of survey participation across subgroups biases estimates of the outcome model’s coefficients and, correspondingly, predicted voting probabilities.

In our case if a greater percentage of people who voted “no” on the ballot propositions participated in the survey, the survey results would overestimate voting for the winning side. Specifically, there would be a negative correlation between the selection equation (which gauges the probability of taking a survey) and the outcome equation (which gauges the probability of voting “yes” on the propositions). That is, the higher the likelihood of participating in the

⁷ See, e.g., Powers and Xie, 2000: 97-99.

survey, the higher the likelihood of reporting a winning, “no” vote. A by-product of estimating a Heckman selection model is estimation of the parameter ρ (“rho”), a measure of the correlation between the selection and outcome equations. A negative, significant ρ provides evidence for non-response bias.

In our case if a greater percentage of people who voted “no” on the ballot propositions participated in the survey, the survey results would overestimate voting for the winning side. Specifically, there would be a negative correlation between the selection equation (which gauges the probability of taking a survey) and the outcome equation (which gauges the probability of voting “yes” on the propositions). That is, the higher the likelihood of participating in the survey, the higher the likelihood of reporting a winning, “no” vote. A by-product of estimating a Heckman selection model is estimation of the parameter ρ (“rho”), a measure of the correlation between the selection and outcome equations. A negative, significant ρ provides evidence for non-response bias.

Formalizing:

Social Desirability Hypothesis: In a cross-classification of predicted by reported vote for both the pre- and post-election samples, classification error produced by predicting a “yes” vote and observing a “no” vote will be higher in the post- than in the pre-election sample:

$$\Pr(PV = Yes, OV = No|T = 2) = \Pr(PV = Yes, OV = No|T = 1)$$

where PV = predicted vote, OV is the observed, or reported, vote, and T is the survey period (1=pre-election, 2=post-election).

Non-Response Bias Hypothesis I: In a cross-classification of predicted by reported vote for both the pre- and post-election samples, classification error produced by predicting a “yes” vote and observing a “no” vote will be the same in the pre- and post-election sample:

$$\Pr(PV = Yes, OV = No|T = 2) = \Pr(PV = Yes, OV = No|T = 1)$$

where the notation is as before.

Non-Response Bias Hypothesis II: The error terms of the two component models of a sample selection model (i.e., the selection and outcome models) will be negatively correlated:

$$\rho < 0$$

where ρ is the correlation of errors between the selection and outcome models.

Methods

We conduct a series of empirical investigations to test the hypotheses set forth above. Because these are not mutually exclusive hypotheses (e.g., it is possible that some voters could move in the direction of the winner prior to the election and others could experience memory lapses after the election, we conduct complementary analyses to investigate the various potential sources of survey winner overestimation. Both the *Late Opinion Shift Hypothesis* and the *Memory Lapse Hypothesis* revolve on patterns of self-reported voting behavior over time. Consequently, we are able to test these hypotheses in the same regression framework. We model support for Propositions 1A, 1B, and 1D over time as a function of a counter for the day of interview, starting with May 11, 2009 (-8, for eight days before Election Day), running through May 24, 2009 (five days after Election Day). We also interact this day counter with a dichotomous indicator for post-election survey responses. This allows us to examine a different slope for the relationship between time and initiative vote in the pre-election and post election periods. Given that each of these propositions lost, we should see negative relationships between day of interview and support for each Proposition before the Election Day (indicating *Late Opinion Shift*) and after Election Day (indicating *Memory Lapse*).

We also control for a variety of potential correlates of voting for or against these propositions. We include party identification, using a dichotomous indicator for Republican respondents, anticipating that Republicans were more opposed to each of these ballot propositions than Democrats or other voters. We also include indicators of approval of the job performance of California Gov. Arnold Schwarzenegger and the California Legislature, anticipating that supporters of California government would be more inclined to vote for each proposition given their generation from deals struck among Schwarzenegger and legislative leadership. We also include a measure of each voter's electoral context too, the percentage vote in support of each proposition in the respondent's county, anticipating that voters in counties supporting the propositions would be more likely to vote for them. We also include indicators for political knowledge, using a measure of the respondent's self-reported knowledge of the party identification of her representative in the California Assembly, and respondent age.

To test the *Social Desirability* and *Non-Response Bias* hypotheses, we first run three separate logistic models of the binary vote variable for Propositions 1A, 1B, and 1D on a common set of explanatory variables for all the pre-election (T=1) respondents. The explanatory variables are education, age, income, identification with the Democratic party, approval of the governor (on a scale of 1 to 10), approval of the legislature (also on a scale of 1 to 10), agreement with citizen budgeting through the initiative process, the county-wide percentage of citizens voting for the proposition, and reported votes on the other two propositions (i.e., Propositions 1B and 1D if the dependent variable is Proposition 1A, etc.). We then use the coefficients from this model to predict a linear score for the post-election (T2) respondents, from which we recover individual-level probabilities of a "yes" vote, using the logit transformation, and predicted votes ("no" if the probability is 0.5 or less, "yes" if it is over 0.5).

We then generate two cross-classifications of predicted vote versus reported vote, one for the pre-election sample (T=1) and the second for the post-election sample (T=2). Each of the two cross-classifications take the form:

[Table 2 about here]

In cells on the diagonal, predicted “no” and “yes” votes correspond to observed votes. So, cell (1,1) contains respondents accurately predicted to vote for (or intend to vote for, in the pre-election sample) the winning “no” side of the propositions, and cell (2,2) contains respondents accurately predicted to vote (or intend to vote for) the losing, “yes” side. The off-diagonal cells (1,2) and (2,1) contain misclassified respondents; total classification error is the sum of their proportions. The cell (2,1), a predicted vote for the losing side, but a reported vote for the winning side, particularly concerns us. The (2,1) cell proportion in the pre-election sample represents classification error; in the post-election sample, it represents classification error *plus* overreporting. That is, the percentage of respondents in cell (2,1) of the post-election cross-classification in *excess* of the percentage in cell (2,1) of the pre-election cross-classification is attributable to overreporting.

In our second test of the *Non-Response Bias Hypothesis*, we estimated three Heckman models (using full information maximum likelihood), one for each of the three ballot propositions we examine. The binary, dependent variable in the selection model is survey response, and the independent variables are county-wide turnout rate, number of elections the voter had voted in previously, education, age, and the square of age. In the outcome model, the dependent variable (also binary) is vote preference on the ballot propositions, and the independent variables are education, age, income, identification with the Democratic party, approval of the governor, approval of the legislature, county vote share on the proposition, and

reported votes on the other two propositions. For our purposes, the crucial parameter in the Heckman model is ρ , the measure of correlation between the two equations.

Results

We find little evidence of *Late Opinion Shifts* or *Memory Lapses* in self-reported voting in these ballot contests. Table 3 reports models of voting for Proposition 1A, 1B, and 1D as a function of a day-of-interview counter before and after Election Day, controlling for a variety of correlates of proposition voting. Negative slopes would indicate shifts toward the winning side (“No”) before the election, and biases in recalled votes for the winner. In the model for Proposition 1A, we actually identify a shift toward the losing side of the ballot initiative, with increased support for Proposition 1A closer to the election. This is the only statistically significant, time-oriented finding across these three models. Prior to the election, there is no meaningful shift in support for Proposition 1B or 1D.

[Table 3 about here]

We graph the predicted probability of support for each of these three ballot initiatives over time in Figure 1, computed using Clarify (Tomz, Wittenberg, and King 2001). The graph demonstrates the shift toward support for Proposition 1A prior to the election, at least in our pre-election data, with its substantial limitations. However, it also shows a suggestive pattern in post-election support for Proposition 1D. At a minimum, the post-election support for 1D is signed in the direction anticipated by our *Memory Lapse* Hypothesis: over time, people are more likely to say they voted against Proposition 1D. This change does not reach conventional levels of statistical significance, although that might also be due to lack of power (i.e., limited post-

election observations). Nonetheless, votes in support of Propositions 1A and 1B are flat across the days following the election.

[Figure 2 about here]

Our tests of the *Social Desirability* and *Non-Response Bias Hypotheses* incline the balance of the evidence in favor of *Non-Response Bias*. Table 4 presents the results of the vote choice models for the pre-election sample. We note that the models are reasonably predictive of voting at T=1, as indicated by the high pseudo-R2 for the Prop 1A (0.46) and Prop 1B (0.43) models and, to a lesser extent, the Prop 1D (0.20) model. Our assumption that these models approximate a true voting model appears to be reasonable. Significant predictors of vote choice at T=1 included education, identification with the Democratic party, legislative job approval, agreement with citizen budgeting via ballot initiatives, and voting on other propositions.

[Table 4 about here]

Tables 5a through 5f present cross-classifications of predicted vote with observed vote in for the T=1 pre-election sample (left column) and the T=2 post-election sample. The lower left-hand cell (2,1) represents classification error where respondents were predicted to have voted “yes” but report having voted (or intending to vote) “no”. In the post-election sample, this cell represents a combination of classification error and overreporting. If there were significant overreporting of votes for the winning side, motivated by reluctance to admit behavior that contravenes social norms, we would expect to see a higher percentage of respondents in this cell for the post-election sample than the pre-election sample.

We observe the opposite. In every case, the cell proportion is *lower*. For Prop 1A, the pre-election sample cell percentage is 5.6%, with 5.5% for the post-election sample. The -0.1% difference is insignificant at $p=0.97$. For Prop 1B, the pre-election percentage is 10.1%, and the

post-election, 7.3% (difference of -2.8% insignificant at $p=.409$). Finally, for Prop 1D the pre-election percentage is 7.9%, with no (0%) respondents in the cell for the post-election sample (difference of -7.9%, significant at $p=0.03$ but in the opposite direction of that predicted by the *Social Desirability Hypothesis*).

[Tables 5a through 5f about here]

Finally, Table 6 presents the results of the Heckman sample selection models for Propositions 1A and 1B.⁸ The selection and outcome models are negatively correlated for Prop 1A ($\rho = -.265$) and significant ($p=.004$). So, the greater the propensity to take part in the survey, the likelier a respondent reported a vote for the winning, “no” side of Prop 1A. The correlation parameter between the Prop 1B selection and outcome equations was in the expected direction ($\rho = -.265$), but did not come close to achieving statistical significance.

Conclusions

The problem of post-election polls’ overestimating winning vote share pervades survey research, extending even to (as we show here) dry, technical ballot proposition contests—an electoral context hitherto unexplored in voting behavior research. Explanations for overestimation in contests between candidates center on psychological factors that cause survey respondents to *misreport* their votes. Respondents either remember their votes inaccurately or, because they wish to present themselves as having engaged in socially sanctioned behavior, deliberately misrepresent how they voted. Here, we propose (and find some evidence in favor of) an alternative hypothesis: voters for the losing side may not lie about how they voted, but rather choose not to participate in a post-election survey in the first place. Despite a plethora of

⁸ Models were estimated using the Stata “heckprob” routine. We omit the model for Proposition 1D, since it did not converge.

research on survey non-response and the reasons for it, scholars have not taken it into account in explaining overestimation of winning vote share.

Our findings, preliminary though they are, suggest that survey researchers need to revise our understanding of survey psychology and respondents' motivations in taking surveys (or not). We know from vote validation turnout studies that survey participants will prevaricate when responding truthfully is embarrassing. Scholars assumed that this explanation also accounted for survey overestimation of winning candidates' vote shares: respondents would say they voted for the winner, because they are ashamed to say they voted for the loser. This study, however, raises the possibility that overestimation occurs not because voters for the losing candidate find it awkward to admit they did so, but because they are simply less interested in taking a survey in the first place. The fault for overestimation, then, may not lie with deceitful survey respondents, but with survey research techniques' inability to get a sample that truly resembles the electorate.

Bibliography

- Atkeson, Lonna Rae (1999). "‘Sure I Voted for the Winner!’ Overreport of the Primary Vote for the Party Nominee in the National Elections", *Political Behavior*, Vol. 21, No. 3 (September), pp. 197-215.
- Belli, Robert F., Michael W. Traugott, Margaret Young, and Katherine A. McGonagle (1999). "Reducing Vote Overreporting in Surveys: Social Desirability, Memory Failure, and Source Monitoring", *Public Opinion Quarterly*, Vol. 63, No. 1 (Spring), pp. 90-108.
- Berinsky, Adam (2004). *Silent Voices: Public Opinion and Political Participation in America* (Princeton NJ: Princeton University Press).
- Brehm, John (1999). "Alternative Corrections for Sample Truncation: Applications to the 1988, 1990, and 1992 Senate Election Studies", *Political Analysis*, Vol. 8, No. 2 (December), pp. 183-199.
- Cochran, William G. (1977). *Sampling Techniques* (New York: John Wiley & Sons).
- Crespi, Irving J. (1988). *Pre-election Polling. Sources of Accuracy and Error* (New York: Russell Sage Foundation).
- Groves, Robert M., and Mick P. Couper (1998). *Nonresponse in Household Interview Surveys* (New York: John Wiley & Sons).
- Groves, Robert M. (2002). *Survey Nonresponse* (New York: John Wiley & Sons).
- Heckman, James J. (1979). "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, No. 1, pp. 153-161.
- Johnson, Marcia K., Shahin Hashtroudi, and D. Stephen Lindsay (1993). "Source Monitoring," *Psychological Bulletin*, Vol. 114, pp. 3-28.
- Little, Roderick J.A., and Donald B. Rubin, 2002. *Statistical Analysis with Missing Data*, 2nd ed. (New York: John Wiley & Sons).
- Kish, Leslie (1965). *Survey Sampling* (New York: John Wiley & Sons).
- Noelle-Neumann, Elisabeth (1993). *The Spiral of Silence: Public Opinion—Our Social Skin* (Chicago: University of Chicago Press).
- Powers, Daniel, and Yu Xie (2000), *Statistical Methods for Categorical Data Analysis* (San Diego, California: Academic Press).
- Presser, Stanley (1990). "Can Context Changes Reduce Vote Overreporting?" *Public Opinion Quarterly*, Vol. 54, pp. 586-593.

- Raghunathan, Trivillore E. (2004). "What Do We Do with Missing Data? Some Options for Analysis of Incomplete Data", *Annual Review of Public Health*, Vol. 25, pp. 99-117.
- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson (1986). "Who Overreports Voting?" *American Political Science Review*, Vol. 80, No. 2 (June), pp. 613-624.
- Southwell, Priscilla L. (1986). "The Politics of Disgruntlement: Nonvoting and Defection among Supporters of Nomination Losers, 1964-1984," *Political Behavior*, Vol. 8, pp. 81-95.
- Tomz, Michael, Jason Wittenberg, and Gary King. 2001 *CLARIFY: Software for Interpreting and Presenting Statistical Results* (Version 2.0). Cambridge, MA: Harvard University.
- Wolfinger, Raymond, and Steven J. Rosenstone (1980). *Who Votes?* (New Haven: Yale University Press).
- Wright, Gerald C. (1990). "Misreports of Vote Choice in the 1988 NES Senate Election Study", *Legislative Studies Quarterly*, Vol. 15, No. 4, pp. 543-563.

Table 1: Comparison of Pre- and Post-Election Surveys and Actual Voting for Three Ballot Initiatives in the May 2009 California Special Election
(Cells are percentage that voted “Yes” on each proposition)

	Pre	N-Pre	Post	N-Post	p ^a	Actual	p ^b
Prop 1A	31.2%	154	20.0%	80	.035	34.6%	.001
Prop 1B	34.4%	151	25.9%	81	.092	38.1%	.008
Prop 1D	23.2%	151	16.3%	80	.109	34.0%	.000

^a p-value of independent samples t-test for difference of means (one-sided) between pre- and post-election survey respondents

^b p-value for single-sample t-test for difference of means (one-sided) between post-election vote share against actual vote share

Table 2: Typology of Survey Respondents in Cross-Classification of Predicted and Reported Voting for Ballot Propositions 1A, 1B, and 1D

		Intended/Reported Vote	
		No	Yes
Predicted Vote	No	Vote for Winning Side	Classification Error (Underreport)
	Yes	Classification Error (Overreport)	Vote for Losing Side

Table 3. Support for Propositions 1A, 1B and 1D, before and after Election Day

	Proposition 1A	Proposition 1B	Proposition 1D
	β (robust s.e.)	β (robust s.e.)	β (robust s.e.)
Day counter	0.123* (0.062)	-0.014 (0.043)	0.025 (0.045)
Post-election survey respondent	-0.833‡ (0.540)	-0.300 (0.424)	-0.014 (0.403)
Post-election respondent × Day counter	-0.088 (0.148)	0.077 (0.142)	-0.127 (0.122)
Republican	-0.610** (0.191)	-0.671*** (0.140)	-0.176 (0.206)
Approval of California Governor	0.165*** (0.046)	0.063† (0.036)	0.092† (0.051)
Approval of California Legislature	0.124* (0.042)	0.171*** (0.044)	0.092* (0.045)
Knows party of representative	-0.353† (0.202)	-0.211‡ (0.149)	-0.139 (0.244)
Age	0.002 (0.007)	-0.003 (0.007)	0.011* (0.005)
County vote for Proposition	0.025† (0.014)	0.021* (0.010)	0.020 (0.020)
Constant	-1.819** (0.822)	-1.564* (0.763)	-2.598** (0.764)
N	233	222	223.000
χ^2_9	72.49***	115.78***	74.21***
pseudo R ²	0.26	0.21	0.11

***p<.001, **p<.01, *p<.05, †p<.10 (two-tailed test); ‡p<.10 (one-tailed test). Respondents clustered by county of residence.

Figure 1. Time Trend in Probability of Respondent Support for Proposition 1A, 1B, and 1D

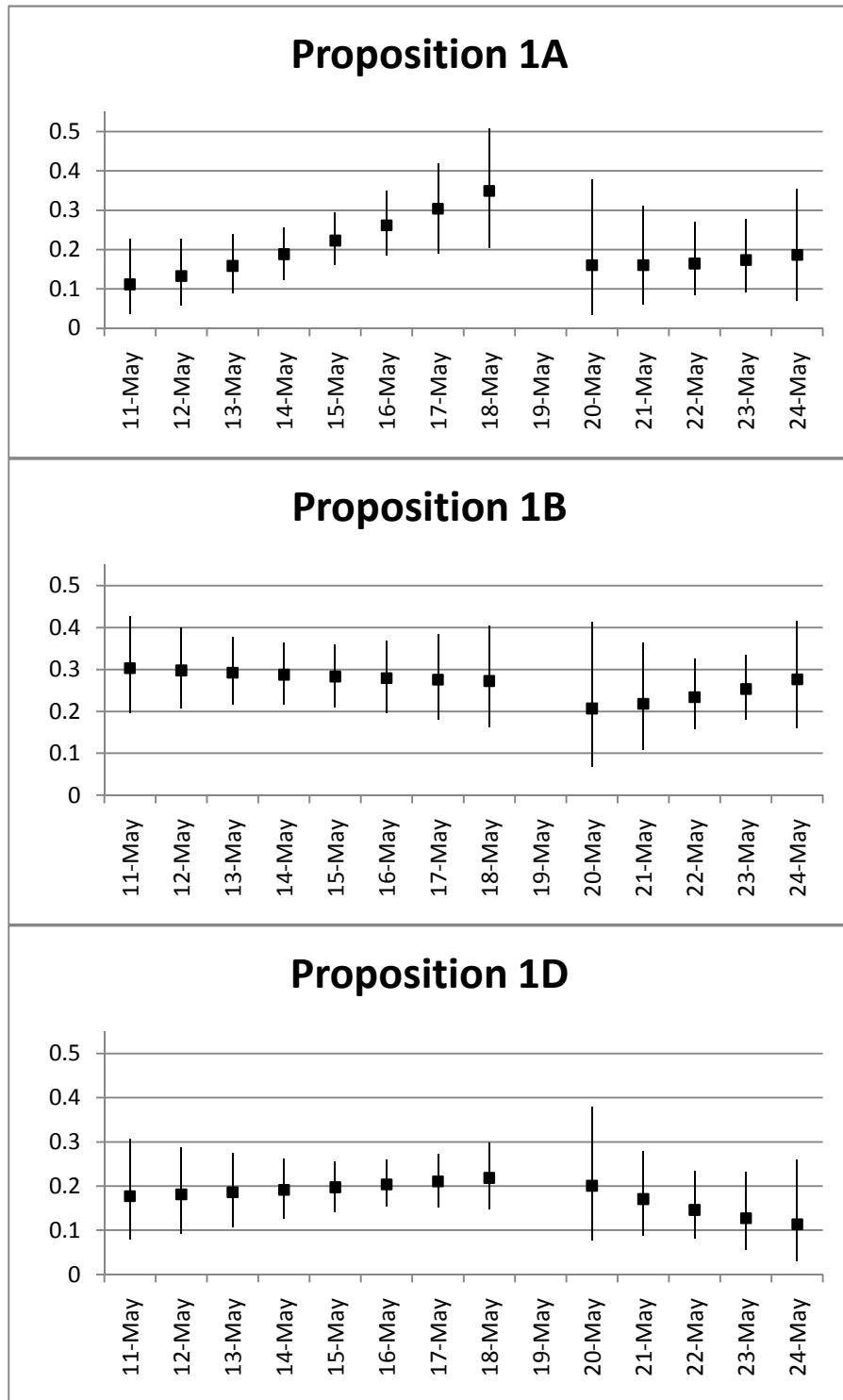


Table 4: Logistic Regression Models for Determinants of Vote Choice on Propositions 1A, 1B, and 1D (Pre-Election Sample)

	Proposition 1A	Proposition 1B	Proposition 1D
	β	β	β
	(robust s.e.)	(robust s.e.)	(robust s.e.)
Education	0.422 (0.409)	0.074 (0.461)	0.42 ‡ (0.306)
Age	0.002 (0.035)	-0.032 (0.029)	-0.001 (0.020)
Income	0.054 (0.048)	-0.062 (0.068)	-0.005 (0.056)
Democrat (1=Yes)	-0.3 (1.298)	1.551 † (0.819)	-0.718 (0.776)
Approval Governor	-0.013 (0.228)	0.039 (0.171)	0.134 (0.154)
Approval Legislature	0.316 * (0.16)	0.01 (0.153)	0.059 (0.126)
Agree with Budgeting by Initiative	-1.299 * (0.645)	0.792 (0.822)	0.333 (0.685)
County Vote %	-0.032 (0.108)	0.091 ‡ (0.068)	0.107 (0.086)
Prop 1B /1A/1A Vote (1=Yes)	3.26 *** (0.974)	3.528 *** (1.245)	1.274 (1.015)
Prop 1D /1D/1B Vote (1=Yes)	1.439 ‡ (1.058)	0.356 (1.153)	0.425 (0.950)
Constant	-4.355 ‡ (3.124)	-4.417 ‡ (3.009)	-7.534 ** (3.390)
N	69	69	69
χ^2_{10}	26.21***	27.84***	13.59
pseudo R ²	0.46	0.43	0.20

***p<.001, **p<.01, *p<.05, †p<.10 (two-tailed test); ‡p<.10 (one-tailed test).

Respondents clustered by county of residence.

Tables 5a-5f: Cross-Classification of Predicted Vote with Reported Vote for Pre- and Post-Election Samples, Propositions 1A, 1B, and 1D

Pre-Election Sample

Post-Election Sample

		Intended Vote on Prop1A	
		No	Yes
Predicted Vote	No	66.3% (59)	10.1% (9)
	Yes	5.6% (5)	18.0% (16)

		Reported Vote on Prop1A	
		No	Yes
Predicted Vote	No	80.0% (44)	1.8% (1)
	Yes	5.5% (3)	12.7% (7)

		Intended Vote on Prop1B	
		No	Yes
Predicted Vote	No	52.8% (47)	12.4% (11)
	Yes	10.1% (9)	24.7% (22)

		Reported Vote on Prop1B	
		No	Yes
Predicted Vote	No	69.1% (38)	9.1% (5)
	Yes	7.3% (4)	14.6% (8)

		Intended Vote on Prop1D	
		No	Yes
Predicted Vote	No	68.5% (61)	13.4% (12)
	Yes	7.9% (7)	10.1% (9)

		Reported Vote on Prop1D	
		No	Yes
Predicted Vote	No	84.4% (47)	10.9% (6)
	Yes	0.0% (0)	3.6% (2)

Table 5: Heckman Sample Selection Models for Propositions 1A, 1B, and 1D

	Proposition 1A	Proposition 1B
	β	β
	(robust s.e.)	(robust s.e.)
Outcome Model		
Education	-0.13 (0.203)	0.04 (1.584)
Age	0.004 (0.015)	-0.012 (0.01)
Income	0.07 (0.025)	-0.087 (0.244)
Democrat (1=Yes)	-0.107 (0.41)	0.544 (1.577)
Approval Governor	0.054 (0.085)	0.041 (0.167)
Approval Legislature	0.146 (0.073)	0.017 (0.079)
County Vote %	-0.001 (0.03)	0.026 (0.107)
Prop 1B Vote (1=Yes)	1.635 (0.377)	1.581 (4.474)
Prop 1A Vote (1=Yes)	1.096 (0.378)	0.568 (1.38)
Constant	-2.337 (1.427)	-0.13 (22.389)
Response Model		
County Turnout %	0.009 (0.007)	0.01 (0.012)
No. of Elections	0.008 (0.006)	0.008 (0.01)
Education	0.253 (0.035)	0.253 (0.035)
Age	0.041 (0.021)	0.04 -0.027
Age ²	0.000 (0.000)	0.000 (0.000)
Constant	-4.387 (0.729)	-4.382 (0.776)
Rho	-0.265	-0.577
p-value	0.004	0.993