

Lies, Damn Lies, and Pre-Election Polling

Elias Walsh

University of Michigan, Ann Arbor,

`fgelias@umich.edu`,

Sarah Dolfin

Mathematica Research Inc.,

`SDolfin@mathematica-mpr.com`,

and John DiNardo,

University of Michigan, Ann Arbor and NBER

`jdinardo@umich.edu`

Corresponding Author: John DiNardo

Session Title: Polls, Votes, and Elections

Session Chair: JOHN DINARDO, University of Michigan

Discussants: MIREILLE JACOBSON, University of California-Irvine

HEATHER ROYER, University of California, Santa Barbara

March 11, 2009

Lies, Damn Lies, and Pre-Election Polling

By ELIAS WALSH, SARAH DOLFIN AND JOHN DINARDO*

In this paper we ask the question: how well do pre-election polls forecast the actual results of elections in the U.S.?¹ The question is interesting for a number of reasons. First, even polling data suggests about 1/3 of polling respondents do not believe that polls work in “the best interests of the general public.”² The situation is such that even many national governments have undertaken to restrict some aspect of pre-election polling. A 1997 international survey of governments, for example, found 30 of 78 surveyed nations had some kind of ban on publication of poll results. Second, there is a strong presumption in the literature on professional forecasting in other contexts (such as interest rate forecasting), which do not rely on sampling per se, that forecasts will be biased. There are a variety of explanations for why forecasts will be biased; one “honest” motivation is that pollsters may avoid reporting results from the unavoidable “atypical” polls. Third, in the literature in economics it is sometimes assumed that polls are unbiased forecasts of (potentially time-varying) underlying preferences for candidates. Fourth, unlike much “opinion” polling, it is possible (albeit imperfectly) to verify the accuracy of the poll. It is therefore possible, with certain caveats, to compare the behavior of polls to what might be expected from probability sampling.

Although the art of polling has become considerably more sophisticated in some respects, the *practice* of polling is a far cry from a textbook description of the power of random sampling and the central limit theorem. Indeed, our analysis of pre-election polling in presidential races suggests some reason for skepticism.

To illustrate the possible problem, consider the 42 “last-minute” national horse race polls from pollingreport.com³ for the 2000 U.S. Presidential Election. This election is particularly well-suited for illustration of the problem since the actual vote was a virtual “tie” (with Al Gore actually winning the popular vote) and the predictions were generally for a close election. Only 3 of the 42 polls predicted either a tie or Gore ahead in

* DiNardo: University of Michigan, School of Public policy, jdinaro@umich.edu. Dolfin: Mathematica, Mathematica Policy Research Inc., Princeton, NJ, SDolfin@mathematica-mpr.com. Walsh: University of Michigan, Department of Economics and School of Public Policy, fgelias@umich.edu. We would like to thank Carolina Krawiec, Amy Kandilov, and Il Myoung Hwang for excellent research assistance.

¹Due to space constraints, the present paper is a shortened version of the original. Along with the Web appendix, the interested reader will find complete references, discussion, and analysis available at <http://www.aeaweb.org/articles.php?doi=10.1257/aer.99.2.XX>.

²More than two thirds of the respondents to the same poll doubted that a random sample of 1,500 people can “accurately reflect the views” of the American public. This, of course, could reflect skepticism about the central limit theorem as well as issues such as non-response!

³See Web appendix Table 1.

the national race. The pollsters themselves appear to have felt that they did “well” in projecting this election. Michael W. Traugott (2001), for example, observes that the performance of the 2000 pre-election presidential polls stands in stark (favorable) contrast to the their performance in the 1996 Presidential election.

For our purpose, what is of immediate import is how *unlikely* it is that these polls – conducted by well-regarded polling agencies – are generated by an unbiased procedure. Consultation of the tables for the binomial distribution reveals that the probability of 39 or more predictions for George W. Bush out of the 42 is less than 5×10^{-7} percent.⁴

I. Background

Our chief argument is that pre-election presidential polling is an activity more akin to forecasting next year’s GDP or the winner of a sporting match than to scientific probability sampling.

Unlike forecasts of economic outcomes which routinely point to a “model” that is generally expected to be different for different forecasters, pre-election polls (and opinion polls in general) routinely characterize themselves as involved in *sampling*. Reports from polls are routinely accompanied by a “margin of error” which is a variant of the confidence interval.

One problem for our analysis which we can not evade is that it is possible that the intent of pollsters is not to forecast an election result, but to correctly sample the current “state of opinion”. Since the current state of opinion can’t be observed, maintaining this view requires maintaining a view that can’t be rejected or accepted by any research design of which we are aware.

Nonetheless, it seems clear to us that a primary reason why pre-election polls (particularly those close to an actual election) are interesting to many is because they are viewed as forecasts of election results. This is the view of some analysts as well: Irving Crespi (1988) observes, ”concluding that even if a poll were conducted immediately before an election, one cannot hope to measure voter preferences accurately enough to approximate election results closely is to impugn the meaningfulness of all polls. If polls cannot achieve such accurate predictability, why should we accept any poll results as having meaning relevant to real life? In fact, using the deviation of pre-election polls conducted close to election day from election results as a measure of accuracy does provide an objective criterion when evaluating alternative methodologies for measuring voting preferences.”

⁴In making this calculation we use the assumption that Gore and Bush received exactly the same number of votes, and the polls were independent samples.

Our approach to assessing bias in pre-election polls is to treat polls as reporting the sample means resulting from random sampling of voters. We find that polls do not fare well by this standard. We also observe that it is impossible to explain “why” polls are biased: there are too many different reasons.

II. Some Basic Problems With Polls

The polls we analyze are largely conducted by profit-making private firms who do not disclose key details of how they arrive at their estimates. Nonetheless, the most reputable pollsters readily acknowledge potential departures from probability sampling.

A. Non-response

Non-response in polls is a critical concern for pollsters. The 2004 National Elections Study had a non-response rate of 24 percent, which varied with the time of year and level of media coverage. Non-response in telephone surveys can be more than 10 percentage points higher. The case for pre-election horse race polls is probably much worse. For example, take this snippet from an interview⁵ with the highly respected pollster John Zogby:

Stewart: “How many people do you have to call... to get 1,300 [responses]?”
Zogby: “Oh boy, figure about 10,000 telephone numbers.”
Stewart: “Really?”
Zogby: “Yeah, really. A lot of people are not home, and about 2 out of 3 people refuse.”
Stewart: “So why isn’t the margin of error 70%?”

In fact, ignoring sampling error and assessing the worst-case bounds arising *only* from non-response bias produces in this case an interval that ranges from $\max(0, \mu - 66)$ to $\min(100, \mu + 66)$, where μ is the population mean.⁶ In one study Pew Research Center found significant differences between “amenable respondents” and “reluctant respondents” in a poll that was likely far more rigorous and expensive to conduct than the best of the pre-election presidential polls we study.⁷ Add the uncertainty involved in estimating (not sampling) voter participation to the above worst-case bound, and almost any estimate can be obtained.

⁵Transcribed from a televised interview with John Stewart on The Daily Show (Zogby, 2004).

⁶ Thus if the population mean is between 44 and 66 percent, then the poll is utterly uninformative.

⁷The two groups differed in the amount of effort that was spent in trying to procure a response. The additional effort included providing a small cash incentive, unlimited and staggered attempts to initiate an interview, and up to three additional attempts to complete unfinished or refused interviews

B. Uncertain Turnout, Uncertain Preferences

In the simplest case, where all voters are certain of their intentions and whether or not they will vote, a suitable probability sample would be sufficient to get an accurate prediction of an election outcome. With certain intentions but uncertainty about whether someone will actually vote or not, estimating the election outcome requires, at a minimum, an estimator of the form, $\bar{Y} = \sum_{i=1}^N P_i X_i$, where P_i is the probability a person will vote and X_i indicates their certain intention. To the extent that P_i is not 1 or zero, an estimate of the election outcome requires a *model* of participation since mere sampling cannot produce a valid estimate of participation even if it could produce a valid estimate of “opinion.”

The problem is exacerbated by the possibility that some important fraction of voters are uncertain about which candidate they support (Charles F. Manski, 1990). Since pollsters generally ask respondents to express their intentions of voting for one candidate or the other as a binary variable, the poll could be biased as a forecast of the election result even if there was ready information on P_i and a proper probability sample was possible.

A simple example will make this clear. Imagine that people can express their preference as a probability from 0 to 1, and that no “surprises” or new information occurs between the time of a poll and the election. Furthermore, for simplicity, imagine voters are identical, are all (correctly) certain that they will vote, and have a 51 percent probability of voting for candidate *A*. Suppose further that they respond to the pollster by saying they would vote for candidate *A* if their underlying probability is greater than 0.5. In this example, the poll would record 100 percent of the vote for candidate *A*, but the election result would be 51 percent.⁸

III. Polling Data

In Web appendix Table 2 we present descriptive information on the polling results we collected from `pollingreport.com`.⁹ We focus on state level presidential polls completed on or after the first day of June in the relevant election year because these tend to be the most consistently well-reported and conducted. Our sample from the 2000, 2004, and 2008 elections contains 1,761 polls with an average of about 12 polls per statewide race, although some races had as few as one poll and some as many as 80. Polling organizations sometimes distinguish between polls of “likely voters” and “all voters” and roughly 83 percent of our polls

⁸Indeed, it is simple to construct examples where, over time, the poll and the underlying preferences of the electorate go in separate directions. See the Web appendix for such an example.

⁹See the Web appendix for details on the data and sample selection criteria.

are from likely voters. The mean reported size of a poll in our sample is 702.

There are several problems with the data that deserve mention and some of these are summarized in Web appendix Table 3 and Table 4. Of particular importance is the fact that some polls report “undecided” voters, and other polls simply drop some fraction of respondents. For virtually all of the analysis we assume that the missing data are “strongly ignorable” – that is, we assume that the “missing” or “undecided” individuals share preferences in the same proportion as those who announce a preference.¹⁰ If a poll reports 40 percent for candidate A, 40 percent for candidate B, 20 percent undecided, and no other candidates, our “adjusted” measure would assign both candidates 50 percent.¹¹ Web appendix Table 4 displays a tabulation of such cases. Nearly all of the polls in our sample require this adjustment. In all of the analysis that follows we focus on the adjusted shares.¹²

A. Results from Analyzing Pre-Election Polls

Tables 1 and 2 summarize several key aspects of the polls we analyze as forecasts of election results. We consider separately all polls, and polls which restrict themselves to “likely voters” only.¹³

Taken as a whole, the polls, on the most favorable terms we can devise, do not behave as would be suggested by simple random (probability) sampling and are biased. To assess departures from what might be expected under random sampling (with certain and unchanging intentions, and certainty about participation), we construct “standardized” prediction errors by subtracting the “true” state election result μ from the corresponding poll prediction and dividing by $\sqrt{\frac{\mu(1-\mu)}{N}}$, where N is the sample size of the poll. Under the null of random sampling, the usual Central Limit Theorem argument suggests that these standardized prediction errors should have a variance of 1. As is evident from the estimates in Table 2, the actual variance of the prediction errors is much larger in magnitude than implied by sampling theory.

Another view of bias and dispersion of the standardized poll errors is provided by a simple kernel density

¹⁰Slightly more formally, if we let r_c denote the percentage point reported in the poll for candidate c among the C candidates reported, our adjusted measure p_i^{Adj} is given by $p_i^{\text{Adj}} = r_c / (\sum_{i=1}^C r_i)$.

¹¹In a related problem, the poll results are virtually always reported rounded to the nearest percentage point. This creates rounding error so that in some cases, the poll results do not sum to exactly 100 percentage points. We handle this symmetrically to the undecided problem. A summary of this “adding up” problem is provided in Web appendix Table 3.

¹²Though the adjusted results present a more “optimistic” assessment of poll accuracy, we present some results using the unadjusted data in the Web appendix.

¹³See Web appendix Table 5 and Table 6 for a complete analysis. Web appendix Table 6 presents results for the three elections separately and the patterns roughly apply. We also conducted several other analyses available in the web appendices. Of particular note is the fact that in the 2000 elections, for example, polls that included any third party candidate provided forecasts with more bias for the Democratic candidate, less bias for the Republican candidate, and much less disperse forecasts for both. However, in 2004 we see precisely the opposite pattern. (See Web appendix Table 7.)

TABLE 1: ELECTION RESULTS FOR POLL SAMPLE

	One Observation Per Poll	
	All Polls	"Likely Voters"
	N = 1,857	N = 1,554
Democratic vote share	49.99	49.98
	{5.93}	{5.66}
Democratic victory	61.60	61.07
Mispredicted victor	24.23	24.26
using prior race		
	One Observation Per Race	
	N = 143	N = 136
Democratic vote share	47.69	48.09
	{8.92}	{8.53}
Democratic victory	46.85	47.79
Mispredicted victor	16.08	16.18
using prior race		

Notes: This table presents election results corresponding to each poll in our sample. In the top panel, each observation represents one poll; in the bottom panel, each observation represents one election. Shares are in units of percentage points. Standard deviations in braces.

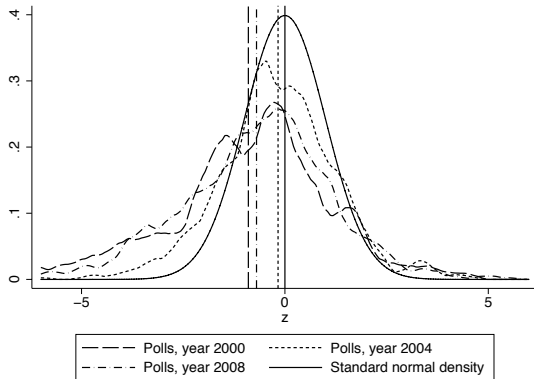


FIGURE 1: DENSITY ESTIMATES OF STANDARDIZED PRE-ELECTION ERRORS OF DEMOCRATIC CANDIDATES

TABLE 2: PRE-ELECTION POLLS

	All Polls	"Likely Voters"
	N = 1,857	N = 1,554
Predicted share	48.95	49.01
	{5.91}	{5.61}
Error	-1.04	-0.96
	{3.45}	{3.29}
Standardized error	-0.55	-0.51
	(0.04)	(0.04)
Variance of	3.20	2.84
standardized error	(0.14)	(0.13)
Predicted victory	55.57	55.15
Mispredicted victor	20.73	20.46

Notes: Results are for the Democratic candidate. Under the null that the poll results are i.i.d. draws from the true distribution, the mean of the standardized prediction error is 0 and the variance is 1. Prediction errors and shares are in units of percentage points. Standard deviations in braces. Standard errors in parentheses. Standard errors on variance estimates are bootstrapped with 1000 repetitions.

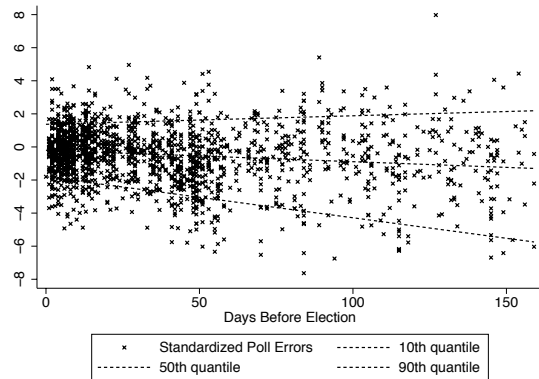


FIGURE 2: SCATTER PLOT OF DEMOCRATIC PREDICTION ERRORS FOR 2000, 2004, 2008 ELECTIONS

estimate of the standardized prediction errors in Figure 1.¹⁴ The estimated densities are too disperse, are not centered at 0, and generally do not share the shape of the standard normal density.

In a subsequent section, we further demonstrate that the difference between the polls and the election outcomes does not appear to be pure noise, but rather is correlated with information available to pollsters (and everyone) at the time the poll is taken.

Table 2 also makes clear that the polls predict the winner more often than not, but the polls guess the

¹⁴See the Web appendix for density estimates of the prediction errors for Republicans; the appendix also includes density estimates for subsamples of the polls we analyze.

winner incorrectly about 20 percent of the time. A very crude “benchmark” model uses the outcome from the previous election as a prediction for the subsequent presidential race. Table 1 documents the fraction of mispredicted races using the prior race outcome. Perhaps surprisingly, by this benchmark pre-election polls do not fare too well. If we compute one prediction per race (as opposed to one prediction per poll) the crude model generally outperforms the polls.¹⁵ Focusing instead on the fraction of mispredicted races using the prior race using one prediction per poll (top panel) demonstrates that the success of the crude benchmark forecast is only partly explained by the fact that more polls are conducted for “hard to predict” races.¹⁶

Web appendix Table 5 analyzes separately the subset of polls conducted within two weeks of the election date. This analysis demonstrates that although there is some *slight* improvement in the poll forecasts closer to the election date, the key features of the errors – bias and over-dispersion – are unchanged. Figure 2 displays the median, and the 10th and 90th quantile regression lines of the prediction errors for all three presidential elections we analyze (Democratic candidates only), demonstrating some decline in the amount of over-dispersion as election day approaches.

The point estimates from the quantile regression of the forecast error for the Democratic candidate on a constant and the number of days confirms the impression from the figure. If a simple linear trend is correct for all three quantiles, the estimates suggest that 100 days closer to the election moves the 90th quantile by 2 standardized units (quite a large amount), and the 10th quantile by about 0.6. Both move in the expected direction – dispersion in the polls diminishes over time. The constant term in the quantile regressions can be interpreted as the hypothetical distribution of poll errors on the day of the election.

Even at this hypothetical best-case there is significant over-dispersion. The 95 percent confidence interval for the constant term for 10th quantile regression is -1.97 to -1.63 and does not cover its theoretical value of -1.28. Likewise the 95 percent confidence interval for the constant term in the median regression is -0.32 to -0.08 and does not cover its theoretical value of zero. For the 90th quantile, the theoretical value suggested by standard normality (1.28) just lies inside the upper part of the estimated 95 percent confidence interval, -1.22 to 1.52.

¹⁵The polls compare more favorably if the sample is limited to those conducted two weeks before the election campaign. See Web appendix Table 5 for these results.

¹⁶See Web appendix Figure 7 for a visual description of where polls are most likely to be conducted.

IV. How “Informative” are the Polls

There are many reasons that polls should be biased. One simple reason is that pollsters may act as “honest Bayesians” and report their posterior distribution instead of the actual poll result.

For instance, imagine a pollster response to a “rogue poll” – a polling result that is wildly inconsistent with other reliable information (such as previous polls). This will happen infrequently of course, but it will happen. Faced with an “unrepresentative” or “unusual” sample, the pollster may “honestly” decide not to report the result of the polling, but massage the answer with his/her prior information to be more consistent with what s/he knows.

The canonical Bayesian approach to this procedure is sometimes referred to as the “Beta–binomial model” which takes the usual binomial distribution likelihood and combines it with a (conjugate) prior of the Beta distribution. Taking the prior and likelihood together generates a posterior distribution for the “honest” Bayesian. It can be shown¹⁷ that the mode of this posterior is merely the weighted average of the prior and the mean from the (unreported) actual polling sample, where the weights reflect the strength of the prior. This suggests an OLS regression,

$$(1) \quad \text{poll}_i = \text{constant} + a * \text{Prior}_i + b * \text{Actual}_i$$

where the parameters a and b are respectively the weights that the typical pollster puts on his prior and the (unreported) actual polling result. If the pollster was merely reporting the results obtained from sampling, then on average the polls would provide the true result, and both a and the constant would be equal to zero.

The “model” as described is easily rejected by the data (although it does remarkably well considering how tightly parameterized the model is) so we instead consider a “just identified” version of equation 1 where we allow an additional parameter that allows the identical priors to vary from the previous election result by a constant γ and assume that the prior can be summarized by a linear combination of previous election results E :

$$\text{poll}_i = a * \left(\sum_{j=1}^J \phi_j E_i^{(t-1)} + \gamma \right) + b * \text{Actual}_i = a * \text{constant} + \sum_{j=1}^J \phi'_j E_i^{(t-j)} + b * \text{Actual}_i$$

¹⁷See the “longer version” of this article.

TABLE 3: THE RELATIONSHIP BETWEEN FORECAST ERRORS AND PRIOR INFORMATION

	Dependent Variable = 2008 Polls	
	(1)	(2)
2008 Outcome	0.821 (0.041)	0.492 (0.099)
2004 Outcome		0.500 (0.154)
2000 Outcome		-0.144 (0.106)
1996 Outcome		0.023 (0.135)
Constant	7.967 (2.098)	7.007 (2.591)
R-squared	0.715	0.736
Observations	677	677

where the constant term (up to scale) identifies a shift from the previous election result, ϕ_j is the weight on the previous election result, and J is as large as two previous election results. These are reported in Table 3 for the 2008 election.¹⁸ Our main result is that the coefficient on the actual outcome is always below one (the prediction from a pure sampling error model.) When we include two previous races in the regression, the coefficient on the actual outcome is about 0.5 for the 2008 election. This suggests that for “honest Bayesians” reported poll results are “one part sample, one equal part prior information.”

This finding helps explain a puzzle: if there are so many reasons for the poll to be biased (non-response, participation model error, the difference between intentions the pollsters questions) why do the polls seem to perform “o.k.”? The simplest answer is that the elections are very easy to predict. Indeed, it is in 2004, when the polls seem to perform the best, that the crude benchmark model most outperforms the pollsters: the 2004 election was, to a large extent, a “replay” of the 2000 election. (See Web appendix Table 6). Indeed, use of the 2000 election result as a prediction would have correctly guessed the winner 94% of the time: the polls we analyzed guessed the victor less than 74% of the time.

V. A Poll that Allows for Uncertain Preferences

While a large literature (see Crespi (1988) for a nice summary) suggests that horse race polls – those that ask respondents about who they intend to vote for in an election – should, if conducted properly and under the right conditions, reflect actual outcomes, an old statistical literature, most recently Manski (1990)

¹⁸See Web appendix Table 8 for the complete analysis.

suggests the opposite. Manski (1990) observes that if a potential voter is uncertain about for whom s/he will vote then a simple “intention” question: “who are you likely to vote for” will be biased in general for the outcome even if agents are perfectly rational, etc. The only hope for generating an unbiased prediction of an outcome from intentions data requires asking the question in such a way that allows the voter to express his or her uncertainty. (See the Web appendix for a further discussion of the intentions problem.) Instead of asking, “If the election were held today, would you vote for X, Y or Z?” one should ask the question in terms of *probabilities* for voting for each of the candidates.

It seems worthwhile to ask whether this “theoretical” source of bias can explain much of the bias we observe in actual polls. In a sense, we would like to see the extent to which this purely statistical problem addresses the following question – are polls biased only because the questions are posed as intentions instead of probabilities?

A. Our Poll

Our poll was conducted by *Time-Sharing Experiments for Social Scientists* (TESS, 2005) and *Knowledge Networks*.¹⁹ Our purpose in designing the questions for the poll was to evaluate the extent to which bias in the polls as forecasts of the outcome are generated by not allowing respondents to characterize their preferences as probabilities. To that end, there are two sets of questions. We call the first set of questions, administered to half of the sample, “the Probabilistic way”; the second set of questions, administered to the (demographically balanced) other half, we call “the Usual way.” Both sets of questions are available in the Web appendix. The Usual style questions are intended to mimic how questions are actually asked in presidential horse race polls, while the Probabilistic style questions allowed respondents to express their voting intentions on a scale of 0 to 100.

B. Results

Web appendix Table 9 presents descriptive statistics of the experimental (Probabilistic) and control (Usual) samples. In both waves we fail to reject differences in mean demographics. Web appendix Table 10 describes the results of the polls. Neither version of the poll does particularly well and use of Probabilistic style questions does not significantly alter the result. The p-value associated with the test of equality of the

¹⁹The data and documentation for our survey is available at <http://www.experimentcentral.org/data/data.php?pid=298>.

predictions for the Democratic candidate in the Probabilistic and Usual style polls is 0.55. The p-value for the joint test of equality for both the Republican and Democratic candidates 0.83. Of course, as is true for any poll results, there are several explanations including non-representative sampling, selection bias and considerable problems with the implementation of the polling by TESS and *Knowledge Networks*. In addition, over 3/4 of the Probabilistic group reported that they were virtually certain of going to the polls, and a similar fraction expressed certainty about their choice of candidate. With such a high degree of certainty among respondents it might have been surprising to see important differences in the preferences of the two groups.²⁰

VI. Conclusion

Voter “uncertainty” and sample selection bias are only two possible problems that might render pre-election polls as unreliable and biased forecasts of the election outcome even when conducted close to the election. There is an enormous literature that proposes other possible reasons which, because of limitations of space, we do not discuss here. Nonetheless, it remains the case that either problem would be *sufficient* to render pre-election polls as unreliable and biased estimates of *trends* – even for the narrowest construct pollsters might care to estimate, i.e. if the election were held *today*

Given the relative ease with which one can arrive a good guess of the outcome of a presidential race at the state level by using the previous election’s result, it is clear that the fact that the polls can often predict the winner is little reason to be sanguine about the “value added” they provide. Our analysis suggests that until a more “severe test” (Deborah G. Mayo, 1996) is proposed there is considerable reason for skepticism.

REFERENCES

- Crespi, Irving.** (1988), *Pre-Election Polling: Sources of Accuracy and Error*, Russell Sage Foundation, New York.
- Manski, Charles F.** (1990). ‘The Use of Intentions Data to Predict Behavior: A Best-Case Analysis’, *Journal of the American Statistical Association* 85(412), 934–940.

²⁰Indeed, the possibility that the the 2004 race was unusual for the high degree of “certainty” most voters had about their intentions, was our primary motivation for attempting to undertake a second poll for the 2008 campaign. We had originally planned and were encouraged to use TESS for a second survey in 2008. Unfortunately, they decided against running the poll at a point too late in the process to find an alternative means to conduct it.

- Mayo, Deborah G.** (1996), *Error and the Growth of Experimental Knowledge*, Science and Its Conceptual Foundations, University of Chicago Press, Chicago.
- TESS.** (2005). ‘Time-sharing Experiments for the Social Sciences’. Data collected by Time-sharing Experiments for the Social Sciences, NSF Grant 0094964, Diana C. Mutz and Arthur Lupia, Principal Investigators.”.
- Traugott, Michael W.** (2001). ‘Assessing Poll Performance in the 2000 Campaign’, *Public Opinion Quarterly* 65, 389–419.
- Zogby, John.** (2004). ‘Interview of John Zogby on the *Daily Show* with Jon Stewart’. <http://www.thedailyshow.com/video/index.jhtml?videoId=127045&title=john-zogby>.