State and National Polls in Forecasting Presidential Elections

Jay A. DeSart
Utah Valley State

Thomas M. Holbrook
University of Wisconsin-Milwaukee

ABSTRACT

This paper is an analysis and update of the DeSart and Holbrook (2003) presidential election forecast model. It compares the performance of our original state-level forecast model (with national-level extrapolations) to an updated model that takes into account national-level forces. We also assess the performance of various models that take into account varying lengths of time in our state electoral history variable. While our original model performed remarkably well in predicting the outcome in 2004, the addition of a national-level poll variable and extending our history variable out over four elections (up from two in our original model) generally makes the model perform better across elections. Given these modifications, we update the model with the data from 2004 so an *a priori* forecast can be generated for the 2008 presidential election when the data become available.

Paper prepared for presentation at the Annual Conference
of the Western Political Science Association, San Diego, CA March 20-22, 2008.

State and National Polls in Forecasting Presidential Elections

For the past two presidential elections we have generated forecasts of the outcome a month in advance based upon the trial-heat polls taken in each state during the month of September along with the states' historical presidential outcomes. In our first attempt, the 2000 election, our model went the way of every other model and overpredicted Gore's share of the vote. However, we redeemed ourselves with our prediction of the 2004 election. This paper presents the results and an analysis of the performance of our model in 2004, and examines two possible modifications to our model: expanding the scope of our "history" variable, and including a national election poll variable. Before dealing with the question of how the model preformed, and how it might be improved, it would be appropriate for us to discuss its basic components and our justification for including them in the model.

The Model

The forecast model that we initially developed after the 1996 election (Holbrook and DeSart, 1999) is a simple and quite parsimonious model comprised of two variables: the average Democratic share of support among the major party candidates in all trial-heat polls taken in each state during the month of September (POLLS), and the average Democratic share of the two party popular vote across the two previous elections (PRIOR VOTE). These two variables are meant to capture the short-term and long-term influences that determine the election outcomes in each of the states. Our forecast model, therefore, is represented by the following equation:

$$VOTE_i = \alpha + \beta_1(POLL)_i + \beta_2(PRIOR\ VOTE)_i$$

This model generates predictions of the Democratic share of the two-party popular vote (VOTE) in each of the states, *i*. The poll variable is derived from data obtained from NationalJournal.com's PollTrack. In our previous work, we have shown that these September polls generally are a fairly good indicator of how the election will turn out in each of the states (Holbrook and DeSart, 1999; DeSart and Holbrook, 2003).

In addition, we have shown that the state-level results can be extrapolated to the national-level, both in terms of the popular vote and Electoral College vote (DeSart and Holbrook, 2003). The process for performing these extrapolations is relatively straightforward. For the Electoral College forecast it is a simple matter of awarding a state's Electoral College votes based on the model's point estimate for the outcome in that state.[1]

The national popular vote extrapolation is a little more complex. Since each state's contribution to the total number of votes cast in the election varies considerably according to its population and level of turnout, simply averaging the state-level predictions across the 50 states would produce a biased estimate favoring the less populated/lower turnout states. To account for this, each state needs to be weighted according to its overall contribution to the total number of votes cast nationwide. To calculate this weight for the purposes of a prediction, a state's total number of votes in each of the two previous elections is taken as a proportion of the total number of votes cast in those elections. Each state's weight is then derived by taking the average of its

[1] This, of course, does have the potential drawback of generating possible errors in awarding electoral votes for Maine and Nebraska, which award electoral votes based on both the statewide vote as well as by congressional districts. However, since this has still not yet happened in any of the previous elections, we remain comfortable with this approach.

proportions across those elections. The national popular vote forecast is thus a simple matter of taking the weighted mean of the state-level forecasts.

The 2004 Forecast

The results of our 2004 *a priori* forecast are presented in the first column in the top half of Table 1. The model we used to generate this forecast was based on the coefficients we presented in our 2003 update of the model (DeSart and Holbrook, 2003):

$$\text{VOTE}_i = -3.63 + 0.84(\text{POLL})_i + 0.20(\text{PRIOR VOTE})_i$$

At the end of September, our forecast model predicted that George W. Bush would win reelection with 51.63% (an error of just .39) of the national two-party popular vote and 321 electoral votes (an error of +35). It correctly predicted the outcomes in 47 states. It incorrectly predicted that Bush would win New Hampshire, Pennsylvania, and Wisconsin. Bush held a small lead in the September polls in New Hampshire and Wisconsin, and Kerry's September lead in Pennsylvania was a miniscule .3%. In the end, Kerry went on to win each of these states by an average margin of just 1.43%

Figure 1 presents each of our state-level forecasts in comparison to the actual results.[2] As one can see, there is a very strong correlation between the predicted and actual results (r = 0.963). The mean absolute error of our state forecasts was just 1.88%, which is the lowest of the four elections in our analysis. To be sure, there were a few outlying cases. Five states had prediction errors of greater than 4% of the statewide two-party popular vote: Massachusetts (6.75), Rhode Island (5.28), Vermont (4.84), Alabama (4.30), and West Virginia (4.05). However, these errors did not result in incorrect

[2] For clarity purposes, Figure 1 presents the predicted and actual vote *margins*, rather than values actually generated by our model, the Democratic candidate's share of the statewide two-party popular vote.

predictions of a winner because in each of these states the winning candidate received no less than 56% of the two party popular vote, and in four of the five states the winner garnered more than 60%. Overall, half of the states had prediction errors of less than 1.5%, and nearly three-fourths (72%) of the states had prediction errors of less than 2.5%. In the three states in which the model predicted the incorrect winner, the prediction errors ranged from 1.53% in Pennsylvania to 3.12% in Wisconsin.

In comparison with other models, our model outperformed each of those examined by Campbell (2005). To be fair, however, it should be noted that these other forecast models have a longer lead time, with each generating predictions by late August at the latest. Even so, we were quite pleased with the performance of our model.

Alternative Models

As pleased as we were with these fairly accurate results, we felt that there was still room for improvement. Based on our earlier analyses, there were two lingering questions that caused us to wonder if maybe there was something more that could be done with the model to improve it.

The first came from our observation in the original presentation of our model (1999) that a "year-specific" model (i.e. including year dummy variables) generally yielded more accurate estimates. This was perhaps not surprising, given the fact that the dummy variables in the model were able to capture some of the unique characteristics of the context of each year's campaign. For forecast purposes, however, this simply was not a reasonable approach since it was impossible for us to know ahead of time what the ultimate effect of the forecast year's campaign context would be.

Our second question, somewhat related to the first, was the fact that the coefficients for the variables exhibited a fair amount of variability across years in the out-of-sample forecasts.[3] When we had run the model successively omitting each year's data from the analysis, we found that the coefficient for the September Polls variable ranged from a low of 0.61 in the 2000 model to 0.95 in the 1992 model, and the coefficient for the Prior Vote variable ranged from 0.13 (1992) to 0.43 (2000). Clearly, we felt, there was something about the shifting contexts of the campaigns that produced such variability in the model.

A Broader Sweep of History

One possible solution is in examining the time frame of the Prior Vote variable. It certainly seemed plausible that the history variable could be influenced by the well-documented shifting home-state patterns that can be brought on from one election to the next (Holbrook, 1991, Campbell, 1992). A particularly popular president or favorite son in a prior election could skew the prior vote variable and yield an errant prediction for a particular state. Take, for example, the case of Arkansas. When one considers Bill Clinton's strong showing in his home state in 1992 and 1996, it is perhaps not surprising to find out that the out-of-sample models over-predicted both Gore's support and Kerry's vote share in that state in both 2000 and 2004. This is especially relevant, given the fact that Arkansas had typically gone rather solidly Republican in 1984 and 1988. Similarly,

[3] Out-of-sample models attempt to replicate the performance of a priori "forecasts" for a prior time period. For example, in order to examine how well the model may have performed in predicting the outcome of the 1992 election, new coefficients were generated using data from all but 1992, and then "predictions" were generated for the 1992 outcome with these new coefficients. This process is repeated for each year in the analysis. In this way, the scenario is similar to that which we had in generating a forecast for 2004 and allows for a reasonable basis for comparison of model performance across elections.

as can be seen in Figure 1, the model under-predicted by a fairly wide margin Kerry's support in both Massachusetts and Rhode Island. This would likely end up producing *over*-estimated support for the Democratic candidate in those states in the 2008 election. To be sure, some of that home state/region advantage should be captured by the September Poll variable, but it may well be that some of the cross-election idiosyncrasies are still creeping into the predictions.

One way we can attempt to mute those idiosyncrasies is to extend the time-frame of the Prior Vote variable to encompass more of the true political leaning of the states. In other words, instead of averaging the results across the two previous elections we could try a span of four elections[4].

The National Campaign Context

The other modification involves an attempt to try and capture the shift in the national-level context from one election to the next. Once again, it is reasonable to expect that some of that would be captured in movements in the state-level polls. However, our prior analyses suggested that there was likely a fair amount of national-level effect that remained unaccounted for. In attempt to deal with this, we ran analyses with a third variable: the average Democratic share of two-party support in national polls taken in the second half of September.

The results of these analyses are presented in the remainder of Table 1 and in Table 2. The conclusions based on these results are somewhat mixed. In terms of the Prior Vote variable, extending the time-frame of history generally produced slightly better results in both our original and national poll models. In comparison to our original

[4] We did run analyses including a prior vote variable spanning three elections. However, the four-election variable outperformed it in nearly every year. Therefore, due to space considerations we do not present those results here.

model, using the four prior election variable shrinks the mean absolute error modestly and generates an Electoral College result that is a bit more accurate in a couple of elections. On the other hand, all the other diagnostics show a decrease in predictive power.

The introduction of the national poll variable into the model overall does not seem to produce more accurate results either. In comparing our original model (top half of Table 1) with the national-poll model (top half of Table 2), the mean absolute error in prediction is roughly the same, if not somewhat higher when the national polls are included. However, in examining the models within Table 2, there is a clear pattern wherein the predictions become more accurate the longer the time frame of the prior vote variable. Even so, this cursory analysis does not suggest that our two proposed modifications to the model are bearing much fruit. Clearly, a more thorough investigation is warranted before we can truly judge the efficacy of these modifications to the model.

Out-of-Sample Models

A more in-depth analysis of the Tables 1 and 2 gives a more detailed understanding of the performance of the models across years. Given our rationale for making our two modifications to the model, there are to pieces of information in these Tables on which we should focus: 1) The stability of the regression coefficients, and 2) The various measurements of prediction error.

Recall that one of our reasons for making these modifications was that the out-of-sample models in our earlier analyses had yielded slightly unstable coefficients across the

years. A simple examination of the coefficients in the Original Model across the years shows that this instability has improved somewhat with the inclusion of the 2004 data. The coefficients for the September Polls variable hover around 0.80 with a swing from a low of 0.73 in the 2000 model to a high of 0.85 in the 2004 model. The coefficients for the prior vote variable similarly settle into a pattern in which the coefficients vary in a .12 point range centered around 0.24. This variability in the coefficients is down significantly from that which we found in our 2003 study.

However, extending the time frame of the prior vote variable or including the national poll variable does not really do anything to improve upon the stability of the coefficients. The cross-election variability in the coefficients does not seem fluctuate much at all in each of the iterations of the model. The only real difference shows up in the two national models (Table 2). The swings in the coefficient for the national poll variable and in the constant start to normalize when the Prior Vote variable is changed from a two-year to a four-year lag. This offers some suggestive evidence that each modification alone might not improve the model much, but taken together there might be some overall improvement.

This becomes even more apparent when one considers the various measurements of prediction error shown in the two tables. Simply trying to compare individual measures across individual years and models is somewhat complicated and presents a mixed bag of results. In some instances the modified models produce better predictions in individual years and less accurate predictions in others. As a result, it is difficult just from these tables to determine which model(s) is/are “better.”

Table 3 attempts to simplify the matter somewhat by simply showing the averages in the various measures across each year for each model. From this simple and straightforward analysis it seems fairly clear that, while our model modifications may have produced slightly less accurate predictions for 2004, they generally tend to produce better estimates in the long run. It is interesting to note that, taken individually, each of the modifications by themselves do not do much to improve the performance of the model. Indeed, they produce less accurate results across the board by themselves. However, when they are combined into a single model they help generate more accurate forecasts. Only the States Correctly Predicted measure shows a decrease in predictive power. Even so the decrease amounts to just one more mispredicted state outcome out of a total of 200. Based on this evidence, we feel safe in concluding that our proposed modifications to our forecast model generally make it better, and should be used for future forecasts.

Updating the Model for 2008

Tables 4 and 5 present the in-sample results for our original model and our "new and improved" model. Table 4 presents the new coefficients that can be used to generate the forecasts for 2008, and Table 5 shows a side-by-side comparison of the performance of the model both within and across elections. This evidence further confirms our conclusion that the modified model generally performs better than our original model. According to nearly every diagnostic presented in these Tables, the modified model provides a better fit with the data and generates more accurate predictions.

Therefore, the model we shall use to generate our forecast for the 2008 election is:

$$VOTE_i = -28.85 + 0.59(POLL)_i + 0.55(PRIOR\ VOTE)_i + .43(NATIONAL\ POLL),$$

Where:

VOTE = Predicted Democratic share of the two-party popular vote in the state
POLL = Average Democratic share of the two-party support in polls taken in the state during the month of September prior to the election.
PRIOR VOTE = Average Democratic share of the two-party popular vote in the state across the four previous elections.
NATIONAL POLL = Average Democratic share of the two-party support in national polls taken in the second half of September prior to the election

Conclusion

This analysis suggests that even though the forecast model that we have used for the two previous presidential elections performed remarkably well in generating fairly accurate predictions for the 2004 election, a few minor modifications to the model could very well make the model even more reliable. The 2008 election shall provide the critical test of the predictive power of the "new and improved" model. Whether or not the uncertainty still looming over the Democratic nomination will result in an election that will be difficult to predict still remains to be seen. Regardless of what the future holds for Senators Obama, Clinton and McCain, we will attempt to peer into the future when we generate our forecast at the end of September, 2008.

References

Campbell, James. 1992. "Forecasting the Presidential Vote in the States." *American Journal of Political Science.* 36: 386-407.

_____. 2005. "Introduction – Assessments of the 2004 Presidential Vote Forecasts." *PS: Political Science and Politics.* 38: 23.

DeSart, Jay A. and Thomas M. Holbrook. 2003. "Statewide Trial Heat Polls and the 2000 Presidential Election: A Forecast Model." *Social Science Quarterly*. 84: 561-573.

Holbrook, Thomas M. 1991. "Presidential Elections in Space and Time." *American Journal of Political Science.* 35: 91-109.

Holbrook, Thomas M. and Jay A. DeSart. 1999. "Using State Polls to Forecast Presidential Election Outcomes in the American States." *International Journal of Forecasting*. 15: 137-142.

Table 1: Out-of-Sample Model Performance
Original Model

	2004	2000	1996	1992
Original Model				
September Polls	0.85**	0.73**	0.77**	0.86**
Prior Vote	0.20**	0.30**	0.21**	0.18**
Constant	-3.63	-1.38	-0.77	-3.47*
Adjusted R^2	0.84	0.85	0.86	0.91
$SE_{y/x}$	3.10	3.10	3.14	2.74
Mean Absolute Error	1.88	2.92	2.52	2.99
States Correctly Predicted	94%	86%	94%	88%
Predictions				
Popular Vote	48.37	52.21	53.50	53.23
Error	-0.39	+2.02	-1.16	-0.13
Electoral College	214	335	355	372
Error	-35	+71	-21	+5
w/ Prior 4 Elections				
September Polls	0.85**	0.73**	0.77**	0.86**
Prior Vote	0.20**	0.30**	0.21**	0.18**
Constant	-3.63	-1.38	-0.77	-3.47*
Adjusted R^2	0.84	0.85	0.86	0.91
$SE_{y/x}$	3.10	3.10	3.14	2.74
Mean Absolute Error	1.86	2.31	2.73	2.96
States Correctly Predicted	92%	94%	92%	86%
Predictions				
Popular Vote	49.68	50.95	52.52	54.23
Error	+0.92	+0.76	-2.14	+0.87
Electoral College	240	291	330	402
Error	-9	+27	-46	+35

Table 2: Out-of-Sample Model Performance
National Model

	2004	2000	1996	1992
National Model				
September Polls	0.72**	0.62**	0.68**	0.77**
Prior Vote	0.33**	0.39**	0.32**	0.28**
National Polls	0.46**	0.27**	0.32*	0.15*
Constant	-28.45**	-14.70**	-17.36*	-11.68*
Adjusted R^2	0.85	0.87	0.86	0.91
$SE_{y/x}$	3.13	2.96	3.09	2.71
Mean Absolute Error	3.31	2.90	2.41	2.89
States Correctly Predicted	82%	86%	90%	84%
Predictions				
Popular Vote	45.74	52.12	55.38	52.39
Error	-3.02	+1.93	+0.72	-0.97
Electoral College	150	335	408	360
Error	-99	+71	+32	-7
w/ Prior 4 Elections				
September Polls	0.61**	0.53**	0.57**	0.71**
Prior Vote	0.54**	0.59**	0.55**	0.43**
National Polls	0.51**	0.47**	0.42**	0.30**
Constant	-33.93**	-29.52**	-27.79**	-22.20**
Adjusted R^2	0.84	0.88	0.89	0.92
$SE_{y/x}$	3.10	2.74	2.78	2.55
Mean Absolute Error	2.13	2.35	2.23	2.68
States Correctly Predicted	94%	92%	90%	84%
Predictions				
Popular Vote	47.35	50.21	54.61	53.87
Error	-1.41	+0.02	-0.05	0.51
Electoral College	214	286	371	397
Error	-35	+22	-5	+30

Table 3: Comparison of Overall Cross-Election Model Performance: Out-of-Sample Models

Measure	Original	Original w/ Prior 4	National	National w/Prior 4
Mean Absolute Error	2.58	2.47	2.88	2.35
States Correctly Predicted	90.5%	91%	85.5%	90%
National Popular Vote Error	0.93	1.17	1.66	.50
Electoral College Error	33	29.25	85.5	23

Figures represent the average of each performance statistic across the four elections in the analysis.
The Popular Vote and Electoral College figures represent the average absolute errors in the predictions.

Table 4: General Forecasting Models for State- and National-Level Presidential Election Outcomes

	MODEL	
	Original	**National w/ 4 Prior Elections**
September Polls	0.798*	0.590*
2 Prior Elections	0.225*	—
4 Prior Elections	—	0.546*
National Polls	—	0.430*
Constant	-2.264	-28.845*
Adjusted R^2	0.862	0.892
S.E. $_{y/x}$	3.12	2.76

N = 200
* = p < .01

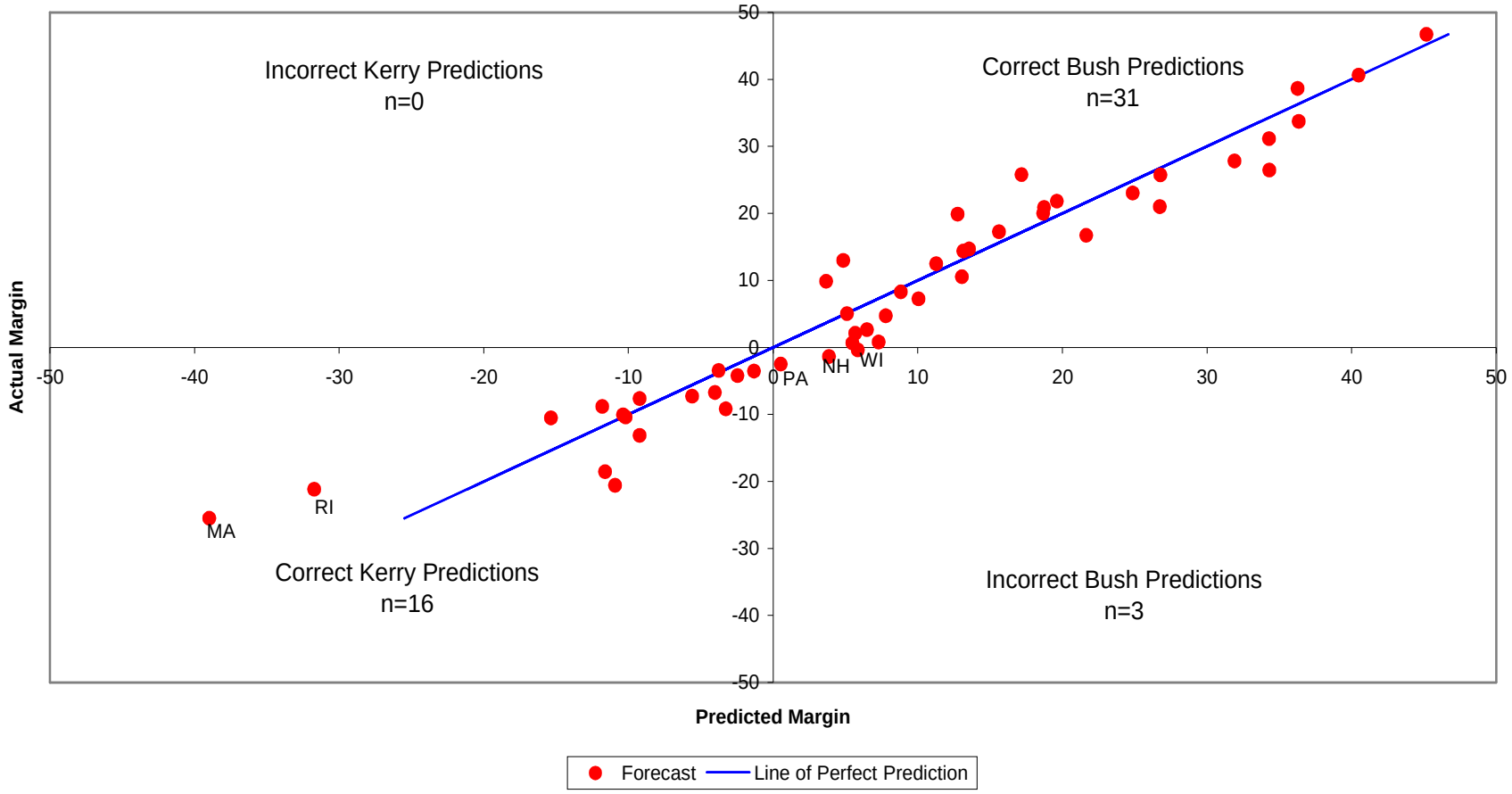
Figures represent unstandardized regression coefficients for each variable in the model.

Table 5: In-Sample Model Performance

	MODEL	
	Original	**Original w/ 4 Prior Elections**
States Correctly Predicted		
1992	86%	88%
1996	92%	90%
2000	90%	92%
2004	94%	96%
OVERALL	90.5%	91.5%
Mean Absolute Error		
1992	2.89	2.43
1996	2.40	2.22
2000	2.24	2.25
2004	1.82	1.74
OVERALL	2.33	2.16
National Prediction Errors		
Popular Vote Error		
1992	-0.67	0.03
1996	-0.72	0.14
2000	0.95	-0.15
2004	-0.18	-0.48
AVG. ABSOLUTE OVERALL ERROR	0.63	0.20
Electoral College Vote		
1992	-5	+18
1996	-18	-5
2000	+42	+23
2004	-35	-14
AVG. ABSOLUTE OVERALL ERROR	25	15

Figure 1

DeSart and Holbrook Forecast Accuracy
2004 State-Level Presidential Election Predictions



Pearson's r = 0.96. Data points represent the model's September and October state-level predictions of Bush's margin of victory compared to his actual margin of victory. Cases in the upper-left quadrant (Ohio) represent incorrectly predicted Bush victories. Those in the lower-right quadrant (New Hamphsire, Pennsylvania, and Wisconsin) represent incorrectly predicted Kerry victories.