State Electoral Histories, Regime Age, and Long-Range Presidential Election Forecasts: Predicting the 2016 Presidential Election

Jay A. DeSart
Utah Valley University

Abstract

This paper presents an adaptation of the DeSart and Holbrook presidential election forecast model for the purpose of making longer-range forecasts of presidential elections up to a year in advance of the election. Relying upon state electoral histories, home state advantage, and "time for change" variables, the model produces in-sample forecasts similar to that of the DeSart and Holbrook September forecast. On the basis of this model, a series of forecasts are generated for the 2016 election.

Paper prepared for presentation at the 2015 Iowa Conference on Presidential Politics,

Dordt College, Sioux Center, IA

October 29-31, 2015

State Electoral Histories, Regime Age,
and Long-Range Presidential Election Forecasts:
Predicting the 2016 Presidential Election

Election forecasting has grown into a significant subfield of political science over the past decade or so. When Tom Holbrook and I first published our own election forecast model (1999), only a handful of models existed and ours was one of the first that was specifically created to generate state-level forecasts for the purpose of generating a forecast for the Electoral College outcome.

Now, with the much greater availability of polling data at both the national and state levels, election forecasting has now grown into a cottage industry that has attracted a great deal of attention, and even a non-academic statistician like Nate Silver has practically become a household name during election seasons. In the past couple of elections we've seen a significant expansion in the development of election forecast models to predict the outcome of American presidential elections. Not only has the number of models grown, but they've also become increasingly complex and sophisticated in their nature.

In 2004, the first APSA Symposium on presidential election forecasts reported predictions from 7 different academic models (Campbell, 2004). In 2008, it reported on 9 models (Campbell, 2008). In 2012, that number had grown to 13 (Campbell, 2012), and the picture presented from this last set of forecasts suggested that the race between Barack Obama and Mitt Romney would likely be a nail-biter. A number of models predicted an Obama victory by margins ranging from about a half of a percentage point to a rather comfortable margin of over 7½ percentage points. But making the picture a bit fuzzier was the fact that a number of models also generated predictions that Romney would win, some by a margin as big as 6 percentage points.

In the midst of all of the mixed signals being sent by this explosion in the number of models, there is one significant shortcoming that we must face: The lead time for such models has appeared to have an upper limit. Most models rely upon polling data and/or economic indicators at both the national and state levels to generate those forecasts and, unfortunately, there has been a paucity of polling data very far in advance of the elections, and the important economic variables featured in many models do not get measured and available until a few months ahead of the elections.

The model developed by Thomas Holbrook and myself is no exception. The critical variable in our model is state-level trial heat polls. One significant limitation of our model is that such polling data is typically not widely collected and available until September of the election year, which effectively narrows the lead time for our prediction to roughly one month before the election, at least a full month behind most other prominent academic forecasts models.

Currently, the academic forecast model of note that has one of the longest lead times is that of Helmut Norpoth. His Primary Model generates a prediction of the outcome based on the candidates' performance in the New Hampshire Primary (Norpoth, 2004). The logic of his model is quite simple: The more divisive a party's nomination contests are, the more difficulty its eventual nominee will have in leading a united party and securing enough votes to win the general election. With this model Norpoth has been able to generate reasonably accurate estimates of the national popular vote 10 months in advance of the election. Using this model, Norpoth has been able to accurately project the national popular vote winner of each Presidential Election since 1996. Furthermore, Norpoth has been able to extend the lead time of his forecast by examining early polls of the New Hampshire primary electorates to attempt to project a winner even further in advance of the election, depending when such polls come out.

Recently, Norpoth outdid himself by developing an even longer range forecast model that would purportedly generate a prediction as long as 4 years in advance of the election (Norpoth, 2013). In other words, relying upon a "Time For Change" sentiment advanced by Alan Abramowitz (1988) in his own forecast model, Norpoth argues that Presidential Elections go through cycles, wherein the longer a party occupies the White House, the more difficult it is for them to hold on to it as regime fatigue sets in with the voters. Based on this model, which simply looks at the national-level outcome of the past two Presidential Elections, he predicts that the Republican candidate, whoever it turns out to be, should be expected to receive 51.4% of the national popular two-party vote.

This paper represents my own attempt at pushing the lead-time envelope by adapting the state-level forecast model developed by Tom Holbrook and myself to generate a prediction a year or more in advance of the election. Before I can present that model, however, it would be most helpful to examine the short-range model upon which it is based.

The DeSart and Holbrook Model

The forecast model that we initially developed after the 1996 election (Holbrook and DeSart, 1999) was a simple and quite parsimonious model comprised of just 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 both the short-term and long-term influences that determine the election outcomes in each of the states. The model, therefore, is represented by the following equation:

$$\text{VOTE}_i = \alpha + \beta_1(\text{POLL})_i + \beta_2(\text{PRIOR VOTE})_i$$

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 various internet sources like NationalJournal.com's PollTrack, PollingReport.com, and Pollster.com. 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). Indeed, it is the case that candidates win the states in which they hold a significant lead in the polls in the month of September. The PRIOR VOTE variable is intended simply as a measure indicating the overall partisan tendency of a state. It serves as a stabilizing influence in the model. While the polls might respond to the short term stimuli of the specific campaign, the prior vote variable reflects the tendency for states to regress back to their typical behavior on Election Day.

Given a less than stellar performance of this model in its first attempt at generating an a priori forecast of the 2000 election, we went back to the drawing board and made some adjustments to it (DeSart and Holbrook, 2003). In spite of the fact that it did quite well in 2004, we did modify the model in advance of the 2008 election by adding a third term, a national poll variable and extending the time frame of the PRIOR VOTE variable from the average of the two previous elections to four (DeSart and Holbrook, 2010).

The reasoning behind making these changes was fairly straightforward. First, the change to the PRIOR VOTE variable would help us mitigate the effects of any home-state advantage that might have been enjoyed by a candidate in a previous election and "normalizing" the average somewhat. This problem was evidenced by the fact that the model over-estimated the Democratic share of the vote in Arkansas in both 2000 and 2004, no doubt due to Clinton's strong "favorite son" showing in that state in 1992 and 1996.

The purpose of the national poll variable was to address an issue that we first identified in our original presentation of the model (Holbrook and DeSart, 1999). In that analysis we determined that there were significant year-specific effects in the forecasts that weren't being picked up by the state-level polls. We added the national poll variable in an attempt to capture those year-specific contextual effects.

Thus, the model we used to generate state-level forecasts for both the 2008 and 2012 elections was:

$$\text{VOTE}_i = \alpha + \beta_1(\text{POLL})_i + \beta_2(\text{PRIOR VOTE})_i + \beta_2(\text{NATIONAL POLLS})_i$$

From those state-level forecasts, we can then generate predictions of both the Electoral College vote and national popular vote. For the Electoral College forecast it is a simple matter of awarding a state's Electoral College votes to the candidate that the model's point estimate suggests will win in that state.

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 the previous election is taken as a proportion of the total number of votes cast in those elections.[1] The national popular vote forecast is thus a simple matter of taking the weighted total of the state-level forecasts.

With this model, then, we can generate a fairly accurate estimate of the outcome a month in advance of the election, once all of the polls in September have been administered and released.

Data Availability and Longer Range Forecasts

As well as this model has performed, its biggest drawback has been that it is limited by the fact that its most important variable, the state-level polls, are simply not available prior to September. As such, it comes out too late to be included in the APSA's quadrennial forecast symposium. Finding a way to break past the time constraint imposed by the unavailability of data, would require us to find alternative measures for our variables; measures that are available much longer in advance of the election.

As it turns out, the solution to this problem presented itself to us in the midst of the 2012 campaign. For the first time since the beginning of our efforts, September came and went and there were a number of states in which no polls had been conducted. Left without a complete set of data with which to generate predictions, we imputed the POLL variable by

---

[1] Previously, we have used a state's average vote total across the two previous elections in an effort to mitigate any potential issues where a particularly contentious down-ticket race in the previous election would spike turnout, however, further analysis showed that any such effect was negligible. Indeed, averaging turnout across two previous elections actually tended to produce *less* accurate estimates of the national popular vote.

averaging poll figures from the 3 or 4 states that were "most similar" to them based on their outcomes in 2008. This proved to be a fairly effective strategy because it allowed us to generate a forecast that missed the national popular vote result by only 1.1% and scored a perfect 51 out of 51 in predicting the state-level winners (and the District of Columbia).

Of course, that still doesn't solve the current problem because it still requires us to have state-level polling data which simply isn't available. However, it does suggest that another way to do a long-range imputation of a state's short-term partisan proclivity is to simply rely upon its result from the previous election. This does, of course, change the nature of the variable somewhat. The purpose of the POLL variable is to capture the short-term influences in each of the states. Can that really be picked up by simply looking at how the state voted in the previous election? Probably not, and it is obviously one of the risks of producing a long-range forecast.

The real question, however, is how well does a state's previous result help us know how it will vote in the next? Simply looking at the bivariate correlation between the states' results with that of their result in the previous election suggests that it doesn't work very well at all. For the elections of 1952-2012, the Pearson's correlation is only .48. However, if we drill down a bit further into the data we can see that not all elections are created equally in terms of their ability to predict the outcome of the next.

Figure 1 presents the simple bivariate scatterplots for each pair of elections from 1952 to 2012. One can see that for a number of elections, the previous result does not do very well at all in indicating what will happen the next time around. However, it is also very clear to see that there is a definite pattern as to when the previous result works and when it doesn't.

The plots from 1960 to 1976 show a tremendous amount of flux from one election to the next, but from that point on the plots have settled into fairly tight and predictable patterns. It is probably no coincidence that the 1960s and 70s have been identified as the time during which the US went through a significant partisan realignment (Carmines and Stimson, 1980, Petrocik, 1981). Indeed, this is pretty clear evidence of the partisan upheaval during that time. More to the point, this shows that during periods of partisan realignment, a state's prior electoral history is understandably not a very good indicator of how hit will go in the next election.

However, we can also see that once the new partisan alignment settled in, state outcomes have become quite predictable simply by looking at the previous election. Since 1980, state outcomes have been correlated with the result from the previous election at an average of .89, and since 1996 the average correlation is .94.

What all of this suggests is that during periods of "normal" partisan alignments state outcomes are fairly consistent from one election to the next, and can be quite useful to project the outcome in the next election. Even so, they are likely not going to be very good to pick up the short-term influences that will be unique from one election to the next. We will still need to find some way of capturing the unique effects of each election context if we are going to be able to generate accurate predictions of outcomes.

Recall that one way in which we capture the unique contextual effects of each election year in our model is a national poll variable. Fortunately, long range national match-up polls are far more common many months in advance of the election than are state-level polls.

It is difficult to find long-range national poll data prior to 2000. The PollingReport.com archives only have national polling data back to 1998. Even so, we can get a sense of how well long-range national polls have projected the actual result for the past 4 elections. Figure 2 presents the scatterplot showing the relationship between the national popular vote result and the average national match-up polls between the two eventual nominees taken in October of year prior to the election. Granted there's only four data points in the plot, but they clearly demonstrate an obvious pattern of a strong positive association. Indeed, Pearson's r is .76. However, it should be pointed out that, for comparative purposes, the NATIONAL POLL variable that we use in our September forecast model for the same time frame is correlated with the eventual result with a Pearson's r of .93. What this suggests is that while certainly not perfect or as strong as those taken much later, national polls taken up to a year ahead of the election appear to be somewhat indicative of what will happen at the polls the following year and may indeed help generate reasonably accurate estimates.

Building a model

With our two long-range surrogates for the poll variables in our model we can now assess how well it performs in generating forecasts a year in advance of the election. This long-range model is specified as

$$\text{VOTE}_i = \alpha + \beta_1(\text{PREVIOUS RESULT})_i + \beta_2(\text{PRIOR VOTE})_i + \beta_2(\text{PRIOR OCTOBER NATIONAL POLLS})_i$$

The performance of this model can be found in the first column, labelled "Simple Model," in Table 1. All three variables in the model perform well, have correlations in the expected direction, and have no trouble achieving statistical significance at the .01 level. The performance diagnostics also show that this model does reasonably well in producing accurate state-level forecasts, correctly predicting the winner in 93% of all cases and generating a mean absolute error in prediction of 2.53. For comparative purposes, it is helpful to note that the model performs almost as well as our September Forecast model.

It, too, correctly predicts the outcomes in 93% of all states, with a slightly lower mean absolute error of 2.30. Perhaps not surprisingly, the long range model has a slightly lower R^2 (.88 compared to .89) and a higher standard error of the estimate (3.30 compared to 2.95) (DeSart and Holbrook, 2013). What all of this suggests is that reasonably accurate forecasts of the final outcome are indeed possible up to a year in advance of the election, even with a lack of polling data.

Additional Model Specification Issues

As well as this model seems to perform, there are still issues that it needs to address given its much longer lead time. Because of the lack of any polling data at the state-level, some short-term factors that we assume are picked up by polls in our September Forecast Model remain unaccounted for in this long-range model. It is these concerns to which we now turn.

Home State Advantage

Perhaps one of the most important, and most easily addressed, factors is the home state advantage. Previous authors have noted a significant home state advantage for presidential candidates in explaining state-level outcomes (Campbell, 1992; Holbrook, 1991). Accounting for this in our September Forecast Model is typically not necessary, because it most likely already comes through in the September polls. When we were developing our model, we consistently found that a Home State dummy variable failed to achieve statistical significance. Simply put, voters in a state most likely already know that a candidate is from their state in September, and any such advantage that brings will already be reflected in how they respond to a trial heat poll.

For a long-range forecast, however, the idea of a home state advantage is problematic for two reasons: First, the lack of state-level polling data means that such an effect is not likely to be picked up in the model unless we specifically specify for it. The second is that a year in advance of the election we often do not know for sure who the eventual nominees will be, let alone the state from which they hail.

These problems become exacerbated by the fact that our long-range surrogate for the state poll variable, PREVIOUS RESULT, is doubly sensitive to the influence of this effect because not only does it not pick up the effect it will have in the upcoming election, but it is retroactively affected by it given who the candidates were in the previous election and where they were from. Therefore, any attempt to specify it in this model must account for both affects.

To do this, I created a simple dummy variable that would attempt to capture these fluctuations in the home state advantage from one election to the next. Clearly, we would expect the support for a candidate to go up in the state wherein they reside, so a simple dummy variable that indicates that the state is the candidate's home state would work. However, since the dependent variable in the model is the Democratic share of the two-party popular vote (2PPV), it needs to be adjusted to whether or not the state is the home state of the Democratic candidate or the Republican candidate to correctly capture the estimated effect. We would expect the Democratic share of the 2PPV to increase in the home state of the Democratic candidate and decrease in the home state of the Republican candidate. So, to capture this effect the home state of the Democratic candidate in any given election is coded 1, and the home state of the Republican candidate is coded -1, to keep it consistent.

In addition, the retroactive home state effect needs to be accounted for as well given the PREVIOUS RESULT variable. Therefore, the opposite coding scheme is used to account for the expected decline the Democratic share of the 2PPV in the home state of the Democratic candidate in the previous election and the expected increase of the dependent variable in the home state of the Republican candidate in the previous election. If we have a president running for reelection, like we did in 2004 and 2012, those increases and decreases would not be present. Therefore, Texas in 2004 and Illinois in 2012 are coded 0, along with the rest of the non-home state states.

The effect of including this variable in the model can be found in the second column, labeled Home State Model, of Table 1. As you can see, the variable does add explanatory power to the model and improves the accuracy of the predictions slightly. The Adjusted R^2 increases slightly and the standard error of the estimate decreases slightly. The state-level predictions improve slightly as well, generating a somewhat smaller Mean Absolute Error in prediction as well as correctly predicting one additional state. Not surprisingly, that one extra state is Arkansas in 2000, which is exactly the state we would expect to flip in the forecast as a result of this specification of the model.

Regime Age and "Time for Change"

One additional factor that needs to be addressed given the lack of state-level polling data is the impact of a "Time for Change" influence. Both Abramowitz (1988) and Norpoth (2013) have demonstrated that the longer a party holds on to the White House, the more difficult it becomes for them to hold on to it. Again, in our September Model, we assume that we are capturing the effects of this factor through the use of the poll variables, since if there is indeed such an effect, it would most likely exhibit itself in the responses of voters a month or two in advance of the election. It *may* be picked up somewhat by the PREVIOUS OCTOBER NATIONAL POLL variable, but to be certain it would make sense to add a term to the model to account for this.

A simple variable which counts the number of consecutive terms one party has held the White House going into the election can be used to operationalize the age of a party's occupancy of the White House. However, like the Home State variable, this will need to be adjusted depending upon whether or not the current occupant of the White House is a Democrat or a Republican. For example, the more consecutive terms the Democrats occupy the White House, the lower we would expect the Democratic candidate to receive in the current election. On the other hand, the more consecutive terms a Republican occupies the White House, we would expect the Democratic candidate to receive a higher percentage of the vote. Therefore, to keep the coding consistent with our expectations the REGIME AGE variable added to the model is the count variable mentioned above, but given a negative sign if we are counting consecutive Republican terms and positive if we are counting consecutive Democratic terms.

The results of including this variable in the model can be found in the third column of Table 1, labeled “Regime Age Model.” Once again we see some slight improvement in the model as a result of adding this new variable. The R^2, standard error of the estimate, and mean absolute error all improve slightly as a result of the respecification of the model. We do however, see one state fewer predicted correctly by this model. Even so, this suggests that including a Regime Age variable would make sense in a long-range forecast model such as this.

It is worth noting as well that one consequence of including the Regime Age variable is that the PRIOR 4 variable becomes insignificant. This is perhaps not surprising given that this variable is essentially trying to account for the fluctuations in vote percentages across the previous couple of elections, depending on who won when. The result of this is that leaving the insignificant PRIOR 4 variable in the model may unnecessarily add error to the prediction. Therefore, the final column of Table 1, labelled “Reduced Regime Age Model” shows the results of removing that variable from the analysis.

As you can see, eliminating the PRIOR 4 elections variable does indeed improve the model slightly on one performance diagnostic. The R^2 is virtually unchanged, but the standard error of the estimate dips slightly. One drawback to this model is that it incorrectly projects the winner in one additional state than the “unreduced” Regime Age model. Even so, it seems to be the case that this would be the most preferred of the four models presented here, and the one we should use in generating a prediction for the 2016 Election.

The 2016 Prediction

Generating a prediction of how the election will turn out next November is a simple matter of plugging the values of the variables into the regression equation presented in Table 1 and seeing who will come out on top. However, one thing that we do not know at this point is who the nominees will be. This makes the selection of the national matchup polls to include in the model difficult, as well as the assignment of the home state advantage.

Dealing with this issue is fairly straightforward, however, because there have been multiple polls taken this month that show various matchups for various candidates. By examining a number of those matchups, we can examine the projections generated by the model for several different pairings. Table 2 presents the matrix of possible outcomes for various pairings based on the model.

The results do not bring good news for the Democrats. In almost all of the matchups, the model projects that the Republican candidate will win. The only Republican candidate that the model suggests will NOT win is Ted Cruz, regardless of whether he faces Hillary Clinton or Bernie Sanders in the general election.

Interestingly, there are two instances where the model projects an Electoral College misfire where Hillary Clinton wins an Electoral College majority but fails to win the national popular vote. According to the model, this happens when Clinton is matched up against Fiorina or Rubio. In those races, she is projected to squeak out a narrow win in the Electoral College against these two candidates despite losing the national popular vote by less than one half of 1%.

Clearly, the "Time For A Change" factor seems to be looming fairly large here and undoubtedly explains these results. Whether or not its influence will be felt as significantly next year as this model suggests remains to be seen, as 2016 will be the first a priori test of its predictive power.

Discussion

The goal of this work was to try extend the traditional lead-time of most election forecast models from a couple months to at least a year. While it succeeds in developing a model that generates such forecasts, it remains to be seen whether it generates *accurate* forecasts. There are a few caveats to consider as we move forward to November 2016.

The first is that this model is only based on four elections worth of data. While that results in 200 data points, it is still uncertain that those four election contexts constitute a significant enough representation of the variation that is likely to occur across time. Furthermore, the analysis of Figure 1 suggests that this model would only work reasonably well during periods of relatively stable partisan alignments. During periods of partisan upheaval, it is quite likely that this model would generate wildly incorrect predictions.

In addition, the model does not have the ability to take into account any idiosyncratic factors that may make any given election unique. For example, it seems fairly clear at this point that one of the things that clearly benefited Barack Obama in 2008 was the undoubtedly historic nature of his candidacy. As the first African-American candidate to win the nomination from a major party, he most likely was able to ride a wave of enthusiasm as people flocked to the polls to cast a historic vote. So, too, it might be the case that Hillary Clinton (or even Carly Fiorina, for that matter) would be a historic candidacy and might similarly generate a groundswell of support and could lead the forecast to be wrong.

Finally, and certainly not least, is the obvious observation that there is clearly a lot of time between now and Election Day. There is certainly enough time for events, campaign or otherwise, to dramatically change the context and completely render the forecast meaningless. Simply put, only time will tell whether this new model is reliable, or whether the attempt to extend the lead-time in this manner was foolhardy.

References

Abramowitz, Alan I. 1988. “An Improved Model for Predicting Presidential Election Outcomes.” *PS: Political Science and Politics*. 21: 843-847.

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

Campbell, James E. 2004. "Introduction - The 2004 Presidential Election Forecasts." *PS: Political Science and Politics*. 37: 733-735.

Campbell, James E. 2008. "Editor's Introduction: Forecasting the 2008 National Elections." *PS: Political Science and Politics*. 41: 679-82.

Campbell, James E. 2012. "Forecasting the 2012 American National Elections: Editor's Introduction." *PS: Political Science and Politics*. 45: 610-613.

Carmines, and Stimson. 1980. *Issue Evolution: Race and the Transformation of American Politics.* Princeton, NJ: Princeton University Press.

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.

DeSart, Jay A. and Thomas M. Holbrook. 2010. "A Tale of Two Models: The DeSart and Holbrook State-Level Forecast Model in 2008." In Chandra S. Putcha, ed. 2010. *Methods of Forecasting American Election Outcomes: Studies in Strategies for Prediction*. Lewiston, NY: Edwin Mellen Press.

DeSart, Jay A. and Thomas M. Holbrook. 2013. "Forecasting Hurricanes and Other Things." Presented at the Annual Meeting of the Western Political Science Association. Hollywood, CA. March 28-30, 2013

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

Norpoth, Helmut. 2004. “From Primary to General Election: A Forecast of the Presidential Vote,” *PS: Political Science and Politics.* 37: 737-740.

Norpoth, Helmut. 2013. “Time For Change: A Forecast of the 2016 Election.” Presented at the 2013 Annual Meeting of the American Political Science Association. Chicago, IL. August 29 – September 1, 2013.

Petrocik, John. 1981. *Party Coalitions: Realignment and the Decline of the New Deal Party System.* Chicago, IL: University of Chicago Press.

Table 1 – Long-Range Presidential Election Forecast Models

Independent Variable	Simple Model	Home State Model	Regime Age Model	Reduced Regime Model
Previous Result	.789*	.794*	1.018*	1.030*
Prior 4	.273*	.273*	.015	—
Prior October National Polls	.687*	.689*	.469*	.471*
Home State	—	2.683*	2.767*	2.768*
Regime Age	—	—	-1.329*	-1.353*
Constant	-19.996*	-37.826*	-25.520*	-25.454*
Adjusted R^2	.882	.886	.899	.899
S.E. $_{y\|x}$	3.301	3.243	3.054	3.046
N = 200				

Model Performance

Percent of States Correctly Predicted	93%	93.5%	93%	92.5%
Mean Absolute Error	2.53	2.44	2.30	2.30

Table 2 – 2016 Long-Range Forecast Outcome Matrix

		Trump	Carson	Fiorina	Bush	Rubio	Cruz
Clinton	Popular Vote	48.89	48.15	49.86	48.85	49.88	51.92
	Electoral Vote	233	201	272	233	272	272
Sanders	Popular Vote	49.00	47.73	46.88	–	48.65	51.92
	Electoral Vote	239	207	207		233	332

Cell entries represent the projected Democratic vote totals for each pairing. Cell shading represents whether the model projects that the Democrat (blue) or Republican (pink) will win. No national polls presenting a matchup between Sanders and Bush were conducted and/or reported during October of 2015.

Figure 1 – Continuity and Discontinuity in Statewide Presidential Election Outcomes, 1952-2012

Plots represent the correlations between the statewide Democratic share of the two-party popular vote of two successive elections.

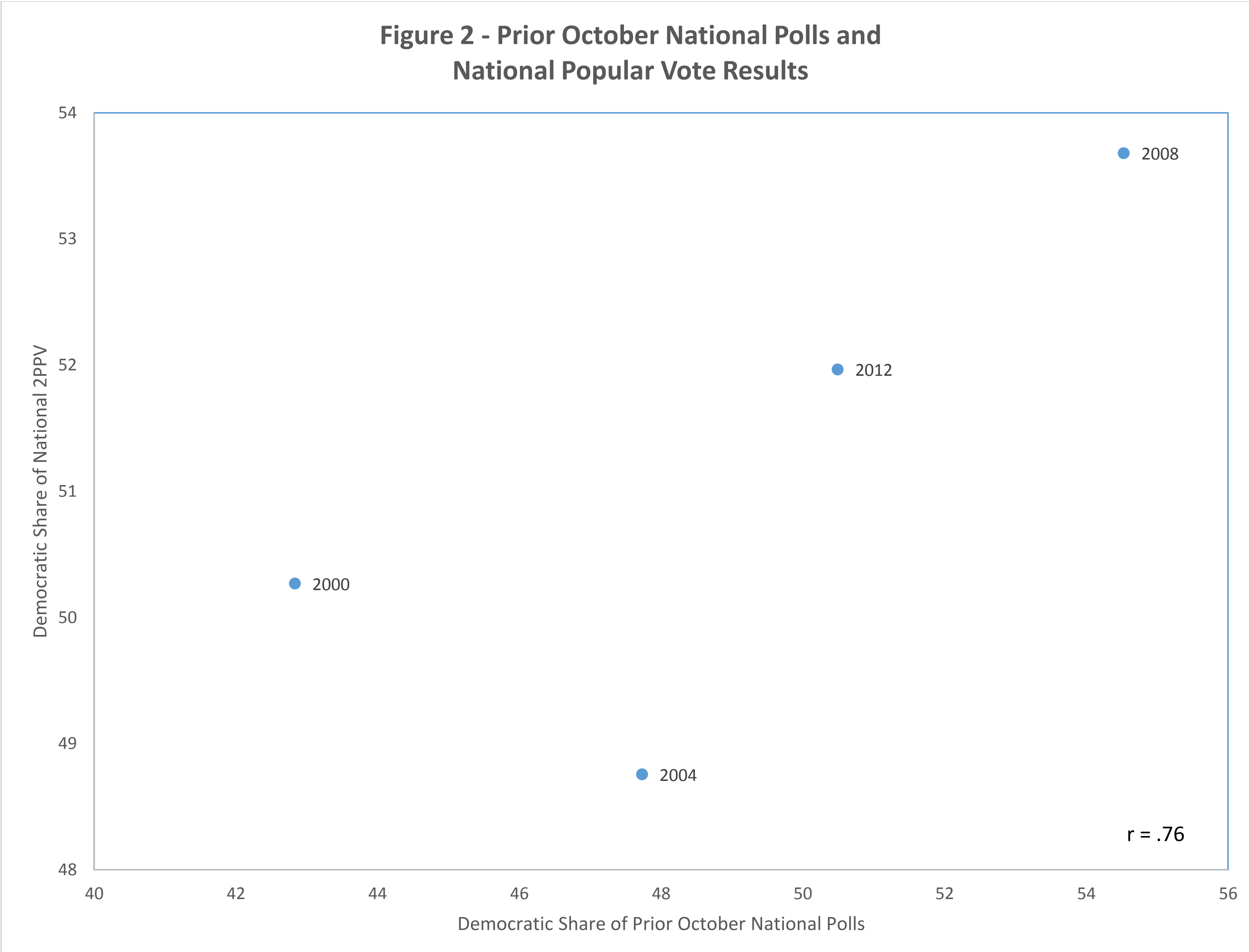
Figure 2 - Prior October National Polls and National Popular Vote Results
54
53
52
51
50
49
48
Democratic Share of National 2PPV
2008
2012
2000
2004
r = .76
40
42
44
46
48
50
52
54
56
Democratic Share of Prior October National Polls