SYMPOSIUM

State-Level Forecasts of the 2012 Federal and Gubernatorial Elections

Carl E. Klarner, *Indiana State University*

The election forecasts presented in this article indicate that control of the White House after the 2012 election is a tossup, that control of the US House will likely remain in Republican hands, and that although closely fought, the Republicans have the edge for control of the US Senate. These forecasts were made on July 15, 2012. Obama was predicted to receive 51.3% of the two-party popular vote, 301 electoral votes, and to have a 57.1% chance of winning the Electoral College. The year 2012 was forecast to be one of stasis for the US House, with almost no change in the number of seats controlled by the Republicans: they were forecast to pick up two seats, and to have a 75.6% chance of maintaining their majority. Lastly, the Republicans were predicted to pick up five seats in the US Senate and have about a 61.6% chance of attaining majority control.

PREDICTION MODELS

These forecasts were made by analyzing the relationship between prior election outcomes and causal factors that can be measured before elections. Then, the values of these causal factors in 2012 were used in conjunction with the strength of the relationships in the past to make forecasts for the 2012 election. Accordingly, the dependent variable in the following four models was the percentage of the two-party vote going to the Democrats in the presidential, US Senate, US House, or gubernatorial elections, using the years from 1948 to 2010.[1] The predictor variables used for these models followed the forecasts of Klarner (2008) with some alterations mentioned later in the text. Space constraints preclude theoretical justification for the independent variables, but this justification has been described elsewhere (Klarner 2008), and all have been used by multiple scholars (see, for example, Abramowitz 2008).

The strength of statistical models as a method of forecasting elections is that they give more insight into why an election turned out the way it did than do forecasting methods that are based on expert judgments or the Iowa Electronic Market. In statistical models, it is tempting to use predictors that ask people how they will vote in the upcoming election. However, we gain little insight into why an election turned out the way it did because that was how most people said they were going to vote three months before Election Day.[2] Accordingly, vote intention variables are not used here, unlike in Klarner (2008).

The difference between the forecasting models presented here and most others is that the *state-year* is the unit of analysis, allowing state-level forecasts (but see Campbell, Ali, and Jalalzai 2006). Klarner (2008) used the *district-year* as the unit of analysis for the US House of Representatives prediction models, but redistricting complicated making district level forecasts for 2012. This unit of analysis was changed partly because lagged vote share mapped to 2012 US House districts were not available, but, more importantly, because past vote share may have a fundamentally different impact in elections after redistricting. Hierarchical linear modeling was used, with a level-2 error term at the year level, which took into account aspects of the partisan wave that were not captured by the national-level variables described in this article. After the models were run, Monte Carlo simulation was used to compute the probability of various national-level outcomes, such as the probability that Obama or Romney would win a majority in the Electoral College.

Cuzán and Burdick (2009) identified major problems with basing election forecasts on limited numbers of cases. Because using pooled cross-section time-series data considerably increases the number of cases for analysis, these datasets may side-step this problem. The availability of so many cases for analysis also allows the use of lagged independent variables, which more fully capture the dynamics of the phenomena being forecast (De Boef and Keele 2008). For models with the US Senate, variables were lagged in terms of the last election for a particular class of senator in a state, unlike in Klarner (2008). Another alteration from Klarner (2008) was that lagged independent variables were included whether or not these attained statistical significance. Finally, missing data were not imputed with a missing data algorithm, unlike in Klarner (2008).

State-years with uncontested elections were not used in the relevant model.[3] To take into account the fundamentally different nature of states in the "solid South," state-years were excluded from analysis when more than 84% of its statehouse members were of a particular major party in the prior year. State-years were also excluded for a particular office if a nonmajor party candidate received more votes than either major party candidate, received more than 25% of the vote, or was an incumbent.

The variables used to predict election outcomes can be put into four general categories: candidate attribute, partisan disposition, national level, and state level.

Candidate Attribute

For the presidential model, one candidate attribute variable measured whether a state was the home state of the Democratic (coded "1") or Republican (coded "−1") presidential candidate (coded "0" otherwise), along with an analogous variable for vice presidential candidates.

doi:10.1017/S1049096512000960

Candidate attribute variables for the US House measured the percent of Democratic candidates in a state and year who possessed a particular characteristic, subtracted by the percent of Republican candidates who possessed that characteristic. Such variables measured the percent of seats that (1) were uncontested by a party, (2) had an incumbent, (3) had an unelected incumbent, (4) had a nonincumbent who held an elective position currently or in the past, and (5) had a nonincumbent candidate who was a former US House member.[4] Incumbents who are more moderate in their roll-call voting behavior have been found to garner more votes than those who are less moderate. Accordingly, the average of Keith Poole's first-dimension DW-Nominate scores for the incumbents in a state who were running for reelection was included, where higher values indicated more conservative voting. An interaction between this variable and the percent of seats in a state in which an incumbent was running was also included, as average DW-Nominate scores would be more influential the greater the proportion of elections in which incumbents were running.[5]

For the US Senate and gubernatorial elections, candidate attribute variables were coded "1" when the Democratic candidate had the attribute in question and "-1" when the Republican did. These measured (1) whether the prior election was uncontested, (2) incumbency, and (3) unelected incumbency. A five-point elective office holder scale was also used for nonincumbents, with negative values for Republicans.[6] The DW-Nominate scores of incumbent senators were also included.

Partisan Disposition

Numerous variables were used to measure the partisan disposition of a state. The first variable was the lagged dependent variable for the relevant type of election. The lagged election returns of three other offices—centered around the average for the year in which they were measured—were also used to measure the partisan disposition of a state. These were the Democratic percent of the two-party vote for president and US House, as well as the percent of major-party state house legislators who were Democrats. Centered presidential returns were not used for the presidential model, and centered US House returns were not used for the US House model.

National Level

The national level variables were as follows. A two-term penalty variable was included in the presidential model, coded "1" when a Democrat had been in the White House for two terms or more, "−1" for such Republican administrations, and "0" in other years (Abramowitz 2008). For the other three models, a midterm penalty variable was used, coded "1" in nonpresidential years when a Democrat was in the White House, "−1" in such years with a Republican administration, and "0" in presidential election years.

Two national level variables were utilized in all models: First, the percent approving of the job of the president in the Gallup Poll was measured as close as possible to 129.5 days before the election.[7] This variable was multiplied by "−1" when a Republican was in the White House. Because of this multiplication, a variable coded "1" when the president was a Democrat and "−1" when the president was a Republican was also included. A second national level variable used in all models measured the health of the economy. Many forecasting models have used gross domestic product to measure the health of the economy, but real personal income is a more theoretically sound way to measure the electorate's economic well-being (Bartels and Zaller 2001). Accordingly, the health of the national economy was measured by the percent change in real disposable income from July of the year prior to the election to May of the election year.[8] This variable was also multiplied by "-1" when a Republican was in the White House.

State Level

Only one aspect of state conditions was used: state economic growth. This was measured by the percent change in real disposable income from the third quarter of the calendar year prior to the election to the first quarter of the election year. In all models, this was multiplied by "−1" when a Republican was in the White House. In the gubernatorial models, an additional variable which represented an interaction between state economic growth and the party of the governor (with a Republican governor coded "−1" and a Democratic governor coded "1") was included to take into account the possibility that the electorate holds governors responsible for state economic growth.

As explained earlier, all independent variables were lagged. With only 16 election years available for analysis in the presidential model, multicollinearity was a problem for the national level variables, so no lagged national-level variables were used. The lagged midterm loss variable had to be excluded for the US House because it was perfectly collinear with other variables. Lagged independent variables were expected to have the coefficient opposite in sign of their nonlagged counterparts.

Table 1 displays the findings for the presidential model, table 2, for the US House model, and table 3, for the gubernatorial and US Senate models. The majority of variables behaved as expected and were statistically significant. Due to space constraints, no comments about the findings of the models are published here.

The difference between the forecasting models presented here and most others is that the state-year is the unit of analysis allowing state-level forecasts.

TRANSLATION OF VOTES INTO SEATS FOR THE US HOUSE

The translation of votes into seats at the state level for the US House was also necessary to model. When a state had only one congressional district, the translation of votes into seats was simply determined by which party had a majority of the

Table 1

Determinants of State Level Democratic Vote: Presidential Elections, 1948–2008

INDEPENDENT VARIABLE	MODEL ONE	
Lagged Vote$_{t-1}$	.849*	.027
Home State Advantage	2.900*	.812
Home State Advantage$_{t-1}$	−1.909*	.793
VP Home State Advantage	1.027	.755
VP Home State Advantage$_{t-1}$	−.511	.811
State House Vote Centered	.054*	.020
State House Vote Centered$_{t-1}$	−.060*	.018
US House Vote Centered	.009	.020
US House Vote Centered$_{t-1}$	−.010	.019
President's Party	−12.767*	5.661
Two-Term Penalty	−7.727*	2.112
Presidential Approval	.202*	.073
National Per Capita Personal Income Growth (%)	.540	.355
State Per Capita Personal Income Growth (%)	.263*	.086
Constant	7.224*	1.649
Σ_u: Year Level Error	3.770*	.687
Σ_e: State-Year Level Error	3.818*	.107
Log Likelihood	−1836.469	
N	655	
Number of Groups (Years)	16	

Note: The dependent variable is the percentage of the major-party vote obtained by Democrats in a state-year. Cell entries are the unstandardized random-effects ML regression coefficient, followed by its standard error in the last column. Random effects are grouped by year. *$p < .05$, one-tailed test.

two-party vote in one of the 2,000 simulations. For all other states, prediction models with the following independent variables were used. The first was the Democratic percent of the two-party vote in a state. The second and third were variables representing the percent of seats in a state that had a Republican (Democratic) incumbent. Given the same percentage of votes in a state for a party, we would expect a party to obtain more seats the more incumbents they have running. The fourth and fifth variables represented the percent of seats for which the Republicans (Democrats) did not run a candidate. Given two states with the same percent of Republican votes, we would expect the Republicans (Democrats) to win more seats in a state where they ran fewer candidates.

Ordered logistic regression was used for predicting the number of seats won by Democrats in states with two congressional districts. For states with three or more congressional districts, ordinary least squares was used as it was found to be as effective at predicting seats as ordered logistic regression. The dependent variable in these models was the percent of seats in the state won by the Democrats. States were divided into four categories for analysis based on their number of districts (3–5, 6–8, 9–13, and 14–53). For brevity, the results of these five models are not displayed here.[9]

Variables that measured partisan control of the redistricting process were not included. However, prior scholarship finds little partisan bias in redistricting (Friedman and Holden 2009). Furthermore, because incumbency was controlled for, the impact of redistricting through retirements and forcing two or more incumbents to face off in the same congressional district was taken into account. Furthermore, the impact of reapportionment is considered.

SIMULATIONS

After the models described here were run, the values of all independent variables were measured in 2012, and the coefficients and standard errors associated with those variables displayed in tables 1, 2, and 3 were used to generate estimated values at the state level for 2012 in the manner described in the following text.[10] A second stage of simulations was carried out for the translation of votes into seats for the US House.

For the first stage of simulations, election outcomes were generated 2,000 times. It is not justified to take the estimated values generated from independent variables' coefficients and then add a random quantity to this estimated value with the standard deviation matching the standard error of the estimate. For out of sample predictions, the uncertainty associated with each independent variables' coefficient must also be considered (Beck 2000). As a result, for each of the 2,000 simulations, a normally distributed random variable with a mean of zero and a standard deviation reflecting the standard error of each coefficient was added to each coefficient. Then, these were multiplied by the values of the independent variables in 2012 and added together to yield predicted values. Next, these quantities were added to two random variables: one that varied across all states in one simulation which also had a standard deviation that matched the level-1 error term, and another that was constant across all states in one simulation, which also had a standard deviation that matched the level-2 error term.[11] The next step in the presidential and US Senate models was to recode each state as being won by the Democrats or Republicans, and then summing up the number of electoral votes or Senate seats for the Democrats for each of the 2,000 simulations. The popular vote for the presidency was computed by weighting Democratic vote share by the number of major party votes in a state in the prior election.[12]

For the translation of votes into seats for the US House, the percent of the vote for the Democrats in a state from each of the 2,000 simulations was put into the second-stage simulation model, and then analogous simulations to those described earlier were done. If the number of seats predicted from one of the 2,000 simulations was out of bounds on the basis of the number of a state's congressional districts and uncontested elections, these were brought to the minimum or maximum bound of the possible range as appropriate.

Table 2

Determinants of State Democratic Vote: US House, 1948–2010

INDEPENDENT VARIABLE	MODEL ONE: HOUSE	
Lagged Vote$_{t-1}$	.570*	.025
Unopposed Difference	–.302*	.014
Unopposed Difference$_{t-1}$	.130*	.015
Incumbency Advantage	.116*	.008
Incumbency Advantage$_{t-1}$	–.054*	.008
Unelected Incumbency	.002	.019
Unelected Incumbency$_{t-1}$	.046*	.018
Previous Office Holder Advantage	.027*	.005
Previous Office Holder Advantage$_{t-1}$	–.007	.005
Former US House Member	.063*	.015
Former US House Member$_{t-1}$	–.017	.014
DW-Nominate	–3.952	2.428
DW-Nominate$_{t-1}$	5.506*	2.424
DW-Nominate ∗ Percent Incumbents	.127*	.032
DW-Nominate ∗ Percent Incumbents$_{t-1}$	–.109*	.032
State House Vote Centered	–.067*	.019
State House Vote Centered$_{t-1}$	.069*	.018
Presidential Vote Centered	.155*	.049
Presidential Vote Centered$_{t-1}$	.003	.051
President's Party	–6.402*	1.804
President's Party$_{t-1}$	6.354*	2.022
Midterm Penalty	–4.398*	.775
Presidential Approval	.122*	.022
Presidential Approval$_{t-1}$	–.096*	.027
National Per Capita Personal Income Growth (%)	.188	.125
National Per Capita Personal Income Growth (%)$_{t-1}$	–.210*	.119
State Per Capita Personal Income Growth (%)	.131*	.066
State Per Capita Personal Income Growth (%)$_{t-1}$	.052	.061
Constant	21.494*	1.311
Σ_u: Year Level Error	1.493*	.251
Σ_e: State-Year Level Error	4.635*	.094
Log Likelihood	–3692.802	
N	1242	
Number of Groups (Years)	32	

Note: The dependent variable is the percentage of the major-party vote obtained by Democrats in a state-year. Cell entries are the unstandardized random-effects ML regression coefficient, followed by its standard error in the last column. Random effects are grouped by year. *p < .05, one-tailed test.

A "drop one year" analysis was conducted for the presidency. In this analysis, one election year from the presidential dataset was dropped in turn, and the model described earlier for all election years was run on the remaining years. Then, estimated values for each state were computed for the omitted year, and the national popular vote was computed by a weighted average, with weights equal to the number of major party votes from the last election. Next, these were compared with the actual popular vote, but only for states that were not excluded from the analysis (i.e., not the solid South). Subtracting the actual from the forecast percent of the Democratic two-party popular vote yielded the following for each year (positive numbers indicate a pro-Republican bias in the forecast): 1948 (7.1), 1952 (7.5), 1956 (5.5), 1960 (3.3), 1964 (8.7), 1968 (−6.4), 1972 (−7.2), 1976 (5.8), 1980 (−9.1), 1984 (2.8), 1988 (.6), 1992 (−2.4), 1996 (1.0), 2000 (.0), 2004 (−1.4), 2008 (−5.3). These numbers imply uncertainty bounds that are smaller than those created by the Monte Carlo simulation presented later in the text, which imply that the simulation method described earlier may be too strict. Second, the presidential model presented in Klarner (2008) was altered in ways that many scholars might consider to be insubstantial. Yet for the 2008 election, Klarner (2008) forecast that Obama would get 53.0%, while the "drop one year" analysis resulted in a prediction of 58.6%. This illustrates that political scientists should put renewed emphasis on justifying modeling decisions that many might believe are small ones.

FORECASTS

The following forecasts were finalized on July 15, 2012. The simulations indicated that there was a 57.1% chance of a Democratic victory in the Electoral College, with a .1% chance of a tie. Obama was forecast to receive 301 electoral votes, the median number of electoral votes across the 2,000 iterations. The simulations indicated a great deal of uncertainty about the outcome, with a 95% confidence interval of 7 to 526 electoral votes for the Democrats, and a 66.7% confidence interval of 119 to 450. The median estimate from the simulations of the Democratic percentage of the two-party popular vote was 51.3%, with a 95% confidence interval of 34.9 to 67.0%, and a 66.7% confidence interval of 43.4% to 59.3%. The probability of Obama being the Electoral College winner but popular vote loser was 2.5%, while the probability of that happening to Romney was 1.9%. Table 4 displays forecasts at the state level for all four offices. Florida and Ohio were forecast to be tossups between Obama and Romney, with Obama predicted to receive 49.7% and 50.3% of the vote, respectively. This forecast provides insight into why there was so much uncertainty about the overall outcome. Although not as closely fought, Pennsylvania was predicted to yield 52.8% for Obama. Numerous medium-sized states were also predicted to be closely fought, including Colorado, Indiana, Iowa, Minnesota, Missouri, North Carolina, and Virginia.[13]

Table 3

Determinants of State Democratic Vote: Gubernatorial (1946–2010) and US Senate (1950–2010) Elections

INDEPENDENT VARIABLE	MODEL ONE: GOVERNORS		MODEL TWO: US SENATE	
Lagged Vote$_{t-1}$	.280*	.050	.227*	.039
Uncontested$_{t-1}$	11.229*	4.699	9.515*	2.675
Incumbency	7.197*	.681	9.535*	.747
Incumbency$_{t-1}$	−3.034*	.706	−1.104	.709
Lagged Vote$_{t-1}$ *Incumbency Dummy	.010	.011	.029*	.012
Unelected Incumbent	−1.571	1.360	−3.916*	1.602
Unelected Incumbent$_{t-1}$	2.542*	1.361	2.955*	1.362
Previous Office Holder	.830*	.210	1.362*	.148
Previous Office Holder$_{t-1}$	−.074	.208	—	—
DW-Nominate	—	—	5.625*	1.901
DW-Nominate$_{t-1}$	—	—	−1.871	1.858
State House Vote Centered	.072*	.036	−.010	.031
State House Vote Centered$_{t-1}$	−.023	.033	.006	.028
US House Vote Centered	.004	.037	.066*	.030
US House Vote Centered$_{t-1}$	.007	.036	−.015	.029
Presidential Vote Centered	.013	.074	.363*	.063
Presidential Vote Centered$_{t-1}$	.049	.077	−.026	.065
President's Party	−6.561*	1.723	−6.976*	2.117
President's Party$_{t-1}$	.483	1.706	3.318	2.098
Midterm Penalty	−2.881*	.643	−2.543*	.830
Midterm Penalty$_{t-1}$	−.100	.612	.646	.787
Presidential Approval	.088*	.022	.093*	.027
Presidential Approval$_{t-1}$	−.014	.023	−.043	.026
National Per Capita Personal Income Growth (%)	.300*	.117	.220	.144
National Per Capita Personal Income Growth (%)$_{t-1}$	.095	.112	−.168	.140
State Per Capita Personal Income Growth*Prez Party (%)	−.032	.109	.148	.117
State Per Capita Personal Income Growth*Prez Party (%)$_{t-1}$	−.120	.100	.099	.113
Governor's Party	−1.702*	.596	—	—
Governor's Party$_{t-1}$	.916	.486	—	—
State Per Capita Personal Income Growth*Gub Party (%)	.055	.101	—	—
State Per Capita Personal Income Growth*Gub Party (%)$_{t-1}$	.021	.095	—	—
Constant	36.151*	2.491	38.294*	2.001
Σ_u: Year Level Error	.469*	.635	1.338*	0.395
Σ_e: State-Year Level Error	7.095*	.193	7.075*	0.176
Log Likelihood	−2379.792		−2869.545	
N	704		847	
Number of Groups (Years)	33		31	

Note: The dependent variable is the percentage of the major-party vote obtained by Democrats in a state-year. Cell entries are the unstandardized random-effects ML regression coefficient, followed by its standard error in the third and fifth columns. Random effects are grouped by year. *$p < .05$, one-tailed test.

Simulations for the US House indicated there was a 75.6% chance the Republicans will keep majority control. The median estimate from the simulations was that the Democrats would be left with 191 seats after the election, a decrease of two seats from their current 193. Again, the margin of error from the simulations was very large, with a 95% confidence interval of 115 to 270 Democratic seats, and a two-thirds confidence interval going from 153 to 228 Democrats.

Table 4

State Level Forecasts

	PRESIDENT: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY	FORECAST NUMBER OF DEMOCRATIC US HOUSE SEATS FOLLOWED BY NUMBER OF SEATS IN STATE	SENATE: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY	GOVERNOR: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY
Alabama	38.2/9.4	1/7	No Election	No Election
Alaska	39.7/12.9	0/1	No Election	No Election
Arizona	46.2/33.9	4/9	38.7/8.9	No Election
Arkansas	38.4/10.4	1/4	No Election	No Election
California	59.0/85.1	36/53	61.8/91.2	No Election
Colorado	51.6/57.3	3/7	No Election	No Election
Connecticut	57.8/80.4	4/5	58.7/83.1	No Election
Delaware	60.9/87.9	1/1	65.0/95.3	61.5/89.0
District of Columbia	84.9/100.0		No Election	No Election
Florida	49.7/48.1	8/27	53.4/64.9	No Election
Georgia	46.0/33.2	5/14	No Election	No Election
Hawaii	68.0/97.2	2/2	57.5/79.7	No Election
Idaho	37.3/8.2	0/2	No Election	No Election
Illinois	60.2/87.0	9/18	No Election	No Election
Indiana	48.3/43.0	3/9	46.6/33.8	45.6/31.8
Iowa	52.1/59.4	2/4	No Election	No Election
Kansas	41.4/16.2	0/4	No Election	No Election
Kentucky	41.5/16.3	2/6	No Election	No Election
Louisiana	40.7/14.8	1/6	No Election	No Election
Maine	55.6/73.8	2/2	Not Available	No Election
Maryland	59.3/84.2	6/8	60.5/88.8	No Election
Massachusetts	56.1/73.5	8/9	48.0/41.1	No Election
Michigan	55.4/72.4	6/14	56.4/78.8	No Election
Minnesota	51.9/58.0	4/8	58.4/83.0	No Election
Mississippi	42.6/20.3	1/4	35.1/4.7	No Election
Missouri	48.2/40.7	2/8	51.8/58.1	55.0/72.3
Montana	47.3/37.5	0/1	49.7/48.7	46.9/36.5
Nebraska	41.7/18.7	0/3	48.0/41.7	No Election
Nevada	53.3/64.7	1/4	44.9/28.2	No Election
New Hampshire	51.1/55.3	1/2	No Election	45.4/28.7
New Jersey	55.3/72.1	6/12	59.1/86.4	No Election
New Mexico	54.5/69.6	1/3	54.3/69.1	No Election
New York	59.7/86.2	19/27	69.2/97.9	No Election
North Carolina	47.9/40.1	6/13	No Election	48.5/42.7
North Dakota	45.4/29.8	0/1	45.4/29.8	37.1/6.3
Ohio	50.3/51.6	5/16	53.9/67.3	No Election
Oklahoma	34.8/4.3	1/5	No Election	No Election
Oregon	55.5/72.8	4/5	No Election	No Election
Pennsylvania	52.8/61.6	6/18	58.9/84.7	No Election
Rhode Island	60.1/86.4	2/2	60.7/88.9	No Election

(continued)

Table 4 (Continued)

	PRESIDENT: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY	FORECAST NUMBER OF DEMOCRATIC US HOUSE SEATS FOLLOWED BY NUMBER OF SEATS IN STATE	SENATE: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY	GOVERNOR: DEMOCRATIC VOTE SHARE FOLLOWED BY PROBABILITY OF DEMOCRATIC VICTORY
South Carolina	44.7/26.2	2/7	No Election	No Election
South Dakota	45.0/30.3	0/1	No Election	No Election
Tennessee	40.9/15.4	2/9	36.3/5.9	No Election
Texas	42.5/20.4	8/36	41.2/16.1	No Election
Utah	36.2/5.9	1/4	31.6/1.6	No Election
Vermont	64.8/94.4	1/1	64.3/94.3	55.7/73.3
Virginia	51.2/55.8	4/11	42.7/18.8	No Election
Washington	55.6/72.7	5/10	57.9/82.7	46.0/32.4
West Virginia	42.6/20.3	1/3	57.1/79.3	58.9/85.2
Wisconsin	54.1/66.3	3/8	51.9/58.6	No Election
Wyoming	34.3/3.8	0/1	25.5/.3	No Election

The forecasting models here did not consider the impact of large nonmajor party candidacies, although Sanders (VT) was treated as a Democrat for the analysis and forecasts. Because Angus King, former independent governor of Maine, was running for the US Senate, no forecast was made for that state, leaving 32 states with forecasts.[14] The results of the US Senate forecasting model and simulations indicated that the Democrats will win 18 of the 32, implying the Republicans will pick up five seats, as the Democrats hold 23 of those that are up for election.[15] Of the seats not up for election, Democrats hold 30, meaning they were predicted to be left with 48 seats, causing them to lose majority control. Of course, the necessary number of senators for majority control is contingent on the party of the vice president. Accordingly, the simulations indicated there was a 42.0% chance the Democrats will have 50 seats or more, and a 33.6% chance they will have 51 seats or more.[16] Together with the probability of the Democrats winning the presidency cited earlier, this implied that the Democrats had approximately a 38.4% chance of retaining control of the Senate.[17] Another contingency is that Angus King will caucus with the Democrats, and the Democrats win the presidency, meaning they need 49 seats or more for majority control. In that contingency, there is a 49.8% chance the Democrats will have majority control.

Last, table 4 indicates that the simulations predicted easy Democratic victories in the Delaware and West Virginia gubernatorial elections, and a clear Republican victory in North Dakota. Missouri and Vermont are leaning Democratic victories, while Indiana, Montana, New Hampshire, and Washington are leaning Republican. North Carolina is highly competitive with a 42.7% chance of a Democratic victory.

Overall, the 2012 elections promise to be highly competitive ones, with control of both the presidency and US Senate within the grasp of both of the major parties, and only the outcome for the US House predicted to be one of constancy. ■

NOTES

1. Except the US Senate dataset, which begins in 1950, and gubernatorial elections, which begin in 1946.
2. Some argue that the use of presidential approval as a predictor variable falls into this category.
3. For the US House model, a state-year was dropped if all of its elections were uncontested. Initially, for this model, uncontested elections were dropped from a district-level dataset, and the dependent variable was the Democratic vote share in the remainder of the state aggregated to the state-year level. This modeling strategy was highly problematic and was not pursued.
4. These data were shared by Gary Jacobson.
5. The modeling decision for the DW-Nominate scores is nonintuitive but valid. These data are from http://voteview.com, accessed June 30, 2012. Scores for the 112th session of Congress were estimated from the single-session files via regression, and with a high degree of predictive accuracy (R-squared of .98 for both the House and the Senate scores).
6. Democratic nonincumbents with the following attributes were given the following number of points: 5 = US Senator or governor, 3 = US House member, 2 = statewide office holder, 1 = state legislator. US House members have the proportion of the state they represented added to their score, meaning a US House member from a state with one congressional district will receive a score of "4." Republicans are given the same scores, but multiplied by "−1." When a position was not held immediately in the past, the number of points is multiplied by 2/3. These weightings were empirically validated.
7. The poll for 2012 was conducted between July 3 and 6. Polls for other years were conducted between 114.5 and 144.5 days before their elections. These data were shared with the author by Gallup.
8. 2012q1 state income data became available on June 27, 2012, while May 2012 national income data became available June 29, 2012. These data are from http://www.bea.gov/, accessed June 30, 2012.
9. "Percent votes Democratic" was always statistically significant in these prediction models. The incumbency variables failed to attain statistical significance two out of ten times, and variables measuring non-contestation failed to attain statistical significance three out of 10 times.
10. As of June 27, the filing deadline for all states' primaries had passed, and states with 294 out of 435 congressional districts and 18 out of 33 US Senate races and six out of 11 states with gubernatorial elections had had their primaries. Louisiana does not have primaries, but the filing deadline for its elections is not until August 17, introducing a greater chance of misclassifying candidate characteristics in that state. For states that had not had their primaries, it was assumed that the candidate with

the highest scoring candidate characteristics would win that primary. Uncertainty about which candidate would win a primary was seldom a problem, due to uncontested primaries, etc.

11. These error terms also varied from simulation to simulation on the basis of random variables with standard deviations matching their standard errors.

12. When the number of major party votes for president in a state is regressed on the prior number of major party votes, the Pearson's R of such an equation varies between .993 and .999 depending on the year.

13. For the state level forecasts, average Democratic vote share from the simulations, median Democratic vote share from the simulations, and the estimated values generated from the model were nearly identical, so which were reported does not matter.

14. King has not stated whether he will caucus with the Democrats or the Republicans after the election.

15. In other words, the median estimate from the simulations was that the Democrats would win 18 of 32 seats.

16. The 95% confidence interval for the number of elections the Democrats will win in the Senate is 10 to 25 seats. The 66.7% confidence interval is 14 to 23 seats.

17. Computed by $((.571 * 42.0) + (.429 * 33.6))$. However, this estimate did not take into account the correlation between outcomes in the two institutions.

REFERENCES

Abramowitz, Alan. 2008. "Forecasting the 2008 Presidential Election with the Time-for-Change Model." *PS: Political Science and Politics* 41 (4): 691–95.

Bartels, Larry M., and John Zaller. 2001. "Presidential Vote Models: A Recount." *PS: Political Science and Politics* 34 (1): 9–20.

Beck, Nathaniel. 2000. "Evaluating Forecasts and Forecasting Models of the 1996 Presidential Election." In *Before the Vote: Forecasting American National Elections*, eds. James E. Campbell and James C. Garand. Thousand Oaks, CA: Sage.

Campbell, James E., Syed Ali, and Farida Jalalzai. 2006. "Forecasting the Presidential Vote in the States, 1948–2004: An Update, Revision, and Extension of a State-Level Presidential Forecasting Model." *The Journal of Political Marketing* 5 (April): 33–57.

Cuzán, Alfred G., and Charles M. Burdick. 2009. "Predicting Presidential Elections with Equally Weighted Regressors in Fair's Equation and the Fiscal Model." *Political Analysis* 17: 333–40.

De Boef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52 (January): 184–200.

Friedman, John N., and Richard T. Holden. 2009. "The Rising Incumbent Reelection Rate: What's Gerrymandering Got to Do With It?" *Journal of Politics* 71 (2): 593–611.

Klarner, Carl E. 2008. "Forecasting the 2008 U.S. House, Senate, and Presidential Elections at the District and State Level." *PS: Political Science and Politics* 41 (4): 723–28.