

either of these. At the same time it preserves information about the uncertainty of the ensemble average.

It seems easy to predict that ensemble averaging will continue to be a part of election forecasting, given Drew Linzer and Simon Jackman’s success this year, as well as the accuracy and popularity of Nate Silver’s efforts. Doubtless it will be advantageous—and fruitful—to continue to broaden the range of models included in the ensemble in the next symposium.

REFERENCES

Montgomery, J.M., F.M. Hollenbach, and M.D. Ward. 2012a. “Ensemble Predictions of the 2012 US Presidential Election,” *PS: Political Science and Politics* 45 (4): 651–54.

Montgomery, J.M., F.M. Hollenbach, and M.D. Ward. 2012b. “Say Yes to the Guess: Tailoring Elegant Ensembles on a Tight (Data) Budget,” Annual Meeting of the American Political Science Association, New Orleans, LA.

2012 PRESIDENTIAL, US HOUSE, AND US SENATE FORECASTS

Carl E. Klarner, *Indiana State University*

For the second presidential election in a row the prediction of the “Klarner Model” was within half a percent of the actual result. This model forecast that Obama would receive 51.3% of the two-party vote, while he actually received 51.8%.¹ The model also called all states correctly, with the exception of Florida, which was predicted to be narrowly lost by Obama with 49.7% of the vote. These forecasts were made on July 15, 2012. This success does not indicate the model is better than the other models in the symposium as luck is a major determinant of which model gets closest to the mark. Although the Klarner Model correctly called 49 out of 50 states, that means little in an election where merely calling states on the basis of 2008 results would have called all but two correctly.

Two other state-level presidential forecasting models were presented in the October *PS* election forecast symposium: that of Jerome and Jerome-Speziari (J&J), which called two states incorrectly, and that of Berry and Bickers (B&B), which called nine incorrectly. The accuracy of the J&J Model for the national popular vote was essentially tied with the Klarner Model, while B&B were 4.3% off. If we stopped here, we could conclude that the J&J Model did as well as the Klarner Model whereas the B&B Model did worse. However, all three models were equally unimpressive in one way, and the B&B Model brought information to the table that was not in a variety of other sources, as explained next.

Making accurate forecasts at the state level is a function of two things: correctly calling the national tide and correctly ordering Democratic success across states. One standard of how good a model is at ordering the states is determined by its ability to add prediction success to a variable measuring the percent of the vote obtained by the Democrat in the last election. Table 1

reports the results of five regressions, with the dependent variable in all being the 2012 Democratic percent of the two-party vote. All variables are centered around 50% to facilitate assessment of bias.

The first regression reported in table 1, in row two, indicates that the 2008 vote accurately ordered the states from most to least Democratic. When using the predictions from the three state-level forecasts as independent variables (rows three, four, and five), none perform better than lagged vote. The fifth regression in the table uses Nate Silver’s November 5, 2012, state predictions as an independent variable as a useful summary of forecasts from polls immediately before Election Day: it only marginally improves on 2008 vote share in its ability to order states, but this indicates that it is possible to improve over prior vote share. When the four forecasting variables are added to a model with lagged vote share in turn, only those of B&B and Silver attain statistical significance (analyses not shown). Perhaps the care that B&B took in modeling state economic conditions yielded predictive capacity not found in other sources. When pitted against Silver’s Election Eve forecasts, their forecasts still attained statistical significance ($p < .05$, analyses not shown).

Figure 1 reports the unstandardized regression coefficient from bivariate regressions of the state Democratic two-party vote on lagged vote share for each of the 17 elections from 1948 to 2012, but excluding the “one party” states not used in the prediction model (Klarner 2012, 655). The figure indicates that the impact has gone up over time, while prediction error (measured by the standard error of the estimate from those 17 regressions, not shown) has gone down, with lagged vote doing better at predicting current vote in 2012 than any other postwar election. This trend is what one would expect with the increased polarization of the electorate over time. These findings suggest that the longer time period used in the Klarner Model may have decreased its ability to order states correctly

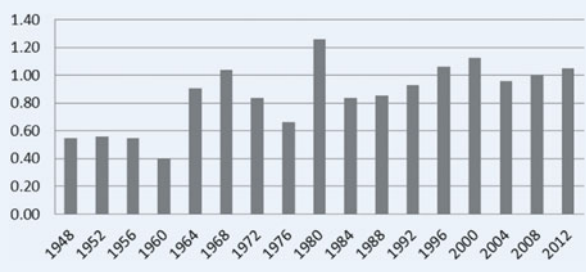
Table 1
Five Regressions Assessing the Relationship between 2012 State Democratic Vote, Forecasts, and Lagged Vote

INDEPENDENT VARIABLE	INTERCEPT	INDEPENDENT VARIABLE COEFFICIENT	STANDARD ERROR OF THE ESTIMATE	R-SQUARED
Lagged Vote	-2.408* (.320)	1.049* (.029)	2.241	.966
Klarner Prediction	-.190 (.346)	1.229* (.037)	2.468	.958
J&J Prediction	-2.163* (.777)	1.239* (.089)	5.449	.797
B&B Prediction	4.711* (.344)	1.059* (.029)	2.267	.965
Nate Silver Prediction	.151 (.264)	1.049* (.024)	1.882	.976

Note: standard error in parentheses. * = $p < .05$. $N = 51$

Figure 1

Impact of Lagged Presidential Vote



from least to most Democratic, as it “overaggregates.” In contrast, the longer time period of the Klarner Model in comparison to the other two state-level models may have helped it call the national tide more accurately. Combining the advantages of both will be a goal for 2016.

The Klarner state-level US House Model predicted that the Republicans would pick up two seats and leave the Democrats with 191. In actuality, the Democrats won 200 seats, the Republicans 234, with the outcome of one race still undecided, making the model nine or ten seats off. Therefore, the US House Model performed reasonably well, as it did in 2006 and 2008.

The Klarner state-level US Senate Model did much worse. It predicted that after the election Democrats would have 48 seats (including Bernie Sanders), and the Republicans would have 51 (a forecast was not made for Maine). Actually, the Democrats were left with 54 seats, and the Republicans with 45, making the forecast six seats off. The Klarner US Senate Model has never performed well, being off by three seats in 2006 and five seats in 2008. US Senate elections appear to be influenced by race-specific factors that are difficult to include in forecasting models.

NOTE

1. 2012 vote percentage based on figures from <http://uselectionatlas.org/RESULTS> accessed on November 28, 2012. The summary table to the symposium incorrectly reported that the Klarner Model predicted 51.2% of the vote for Obama. Campbell's 2012 forecast was tied with the Klarner Model in accuracy. In 2008, the Klarner Model predicted Obama would receive 53.0% of the two-party vote while he received 53.4%.

REFERENCE

Klarner, Carl E. 2012. “State-Level Forecasts of the 2012 Federal and Gubernatorial Elections.” *PS: Political Science and Politics* 45 (4): 655–62.

WHY THE STATE-BY-STATE POLITICAL ECONOMY MODEL DID IT RIGHT

Bruno Jérôme, *University of Paris 2*

Véronique Jérôme-Speziari, *University of Paris Sud 11*

One hundred and forty two days before the 2012 US presidential election our final State-by-State Political-Economy Model gave an advantage to Barack Obama with 51.6% of the popular

vote (error margin ± 4.47) and 324 electoral votes (Jérôme and Jérôme-Speziari 2012). On November 6, 2012, with 51.6%¹ of the vote and 332 electoral votes, the Democratic incumbent wins a second term. Regarding certainty of an Obama plurality, the model gave a probability of victory by 64%. In 2012, it seems that this was enough to ensure a good predictability.

Thus, our model successfully predicted the correct Democratic/Republican balance of power. Moreover, this ratio was correctly forecast in 48 states (+DC) out of 50 (+DC) with the exceptions of Virginia (given to Republicans) and West Virginia (given to Democrats), for which an explanation of the gap between forecast and actual results should be provided. (See figure 1.)

Note that the model correctly predicted the results in the main battleground states such as Colorado, Florida, Iowa, and Ohio. At last, it has been sensitive enough to forecast that Indiana and North Carolina would return to the Republican side.

The vote predictions based mainly on economic determinants (local change in unemployment) and political determinants (president's job approval and parties' local partisan dynamics) helped to correctly predict the election's outcome.

If we decode our results, it seems that the change in unemployment—and not the level—was a decisive factor in the vote, more or less amplified depending on the states. From a 10% peak in October 2009, the unemployment decreased to 7.8% in October 2012 (this rate was the same when Barack Obama took office). It should be added that Obama succeeded in keeping afloat his popularity. Over the long term, from 68% satisfied at the beginning of his term, his popularity decreased to 56% in October 2009 before reaching 51% in November 2012 (Gallup Poll).

At least, since 2008, partisan dynamics across the states seem to have been particularly stable given that only Indiana and North Carolina flipped back compared to the previous election.

As regards exogenous parameters (e.g., not in the model) that could perturb it, it seems that “Obama's failure” in the first debate had no negative effect on the Democratic vote, neither was there a “Sandy” effect that might have had a particular impact within some kind of Rally-around-the-Flag effect. Similarly, the likelihood of an “anti Obama” referendum and the Romney's slogan “are you better off than four years ago” seem not to have been decisive, nor a “racial cost” possibly offset by the growth in the population of visible minorities and their high level of participation once again.

Finally, even if we obtained a very satisfactory performance, how could we improve our model to do even better in 2016?

First, we certainly have to pay more attention to two states, Virginia and West Virginia. First, in Virginia, the model failed because the partisan dynamics variable (establishing this state as a solid historical Republican stronghold) did not capture the silent change in the Virginian electoral sociology (in suburban areas especially). This means that the components of our state-level partisan dynamics index, which performs rather well but is still a black box, need to be more precisely measured. Second, the error on West Virginia comes from our Job Approval Index, measured at a national level, which overestimated Obama's popularity in a state where the incumbent has