SYMPOSIUM

The Predictive Power of Uniform Swing

Simon Jackman, *Stanford University*

The 2012 election was a fabulous "moment in the sunshine" for those of us using model-based poll-averaging to predict the election outcome, state by state.[1] As is now well known, political scientists Drew Linzer and myself matched the predictive performance of (the much more famous) Nate Silver, accurately forecasting the winner of each of the 50 states plus the District of Columbia.

The focus on state-level forecasts—made possible by an abundance of state-level polling—stands in marked distinction with the bulk of US presidential election forecasting, which concentrates on the national level. For instance, almost all of the models in the October 2012 symposium in *PS: Political Science and Politics* looked at national two-party vote share for the candidate of the party of the incumbent president. Just one of the models presented in the preelection *PS* symposium presented state-level forecasts, that of Berry and Bickers (2012).

The research program centered on forecasting national-level, US presidential election results faces severe practical limitations. There simply is not much data, with a new data point arriving only once every four years. Sample sizes are small. The uncertainty accompanying forecasts is relatively high. Moreover, it might seem that the national outcome (and a forecast of it) has only limited political relevance, because after all, presidential elections are de jure a series of winner-take-all, state-by-state contests. We might well conclude that a state-level approach is the more fruitful way to proceed given the growing abundance of state-level polls, the "triumph of the quants," and the accuracy of their state-by-state calls in 2012.

Maybe. But a critical connection exists between national and state-level election outcomes that is of great utility when making state-level forecasts. Specifically:

1. election-to-election variation in state-level, presidential election results has long been powerfully shaped by national-level factors (e.g., Bartels 1998);
2. the magnitude of the national component of state-level swing in presidential election results has generally been trending up in recent elections;
3. in fact, the election-to-election correlation in state-level outcomes in presidential elections reached .98 in 2012 (the correlation of state-level, two-party vote shares in 2012 with 2008 results), the highest such correlation since World War II (see figure 1).

Accordingly, a model in which state-level swings are assumed to be constant across states—so-called uniform swing—is now a better approximation for state-level election outcomes than at any time in the last 70 years.

This suggests a simple recipe for generating state-level election forecasts:

1. Take the output of a national-level forecasting model, or an average of such models, *a la* Montgomery et al. (2012). Let $\hat{y}_t$ be the national-level forecast of Democratic share of the two-party vote for president from that model.
2. Each $\hat{y}_t$ implies a forecast of the national-level swing, $\hat{\delta}_t = \hat{y}_t - y_{t-1}$ where y_{t-1} is the Democratic share of the two – party vote for president at the previous election.
3. Form predictions for each state s, $\hat{y}_{st}$, by applying uniform swing: that is., $\hat{y}_{st} = y_{s,t-1} + \hat{\delta}_t$.

As we will now see, armed with a reasonable estimate of the national swing, we can do a reasonable job of predicting state-by-state election outcomes. This insight played a key role in Drew Linzer's prediction model (Linzer 2013).

Figure 2 shows the swings in Obama's share of the two-party vote, 2008 to 2012, by state. Swings are not "uniform" in the sense of being constant across states. The national, two-party swing away from Obama was -1.48 percentage points. "Home state" effects account for the biggest departures from the national swing. With Sarah Palin not on the Republican ticket in 2012, Alaska recorded almost a 4-point swing *toward* Obama, while the Mormon "home state" of Utah recorded almost a double-digit swing toward Romney.

It is tempting to look at figure 2 and see considerable variation in swings across states. The standard deviation of the state-level swings is 2.36 percentage points; omitting Alaska and Utah the standard deviation drops to 1.95. But recall the message of figure 1, showing a strengthening of the relationship between state-level presidential election outcomes, election-to-election. By historical standards, the swings displayed in figure 2 are quite tightly distributed around the national swing.

Several implications follow. First, a range of uniform swing estimates exists that would have generated a perfect set of state-by-state predictions. Any uniform swing estimate greater than 0.48 but smaller than 1.27 percentage points would have produced a perfect set of predictions: big enough to get North Carolina and Indiana changing hands, but not so big as to predict that Florida would be won by Romney.

Second, the actual national swing of 1.48 percentage points lies outside this interval. Even with perfect foresight of the national swing, the uniform swing assumption would have led to an erroneous prediction that Florida would fall to Romney.

doi:10.1017/S1049096514000109

So, there are limits to the utility of national, uniform swing when applied to the states, even when the state-level variation around the national swing is small. Still, the predictive performance here would have been 50/51, which is hardly shabby.

Of course, we do not know the *actual* levels of national swing ahead of the election. How did the national, aggregate-level vote forecasts perform when we feed their predictions down to the state level via the uniform swing assumption?

root of the average squared prediction error, and so analogous to $\hat{y}$ in a linear regression setting); (2) the number of states called correctly.

Figure 3 displays levels of RMSE by assumed level of uniform swing, over an interval that spans the levels of uniform swing implied by a set of aggregate forecasting models. Most of the forecasting models—and their point predictions—are those published in the *PS* symposium in fall 2012 (Campbell 2012) I add a model produced by Hill, Sides, and Vavreck that

Of course, we do not know the actual levels of national swing ahead of the election. How did the national, aggregate-level vote forecasts perform when we feed their predictions down to the state level via the uniform swing assumption?

We can compare the predictive performance of the aggregate forecasting models—treated here as generating state-level predictions via the uniform swing assumption—one any one of a number of performance metrics. Space constraints allow me to examine only two metrics: (1) root mean square error (RMSE), perhaps the most commonly used measure of the performance of predictions with respect to a continuous variable, and computed exactly as the name suggests (the square

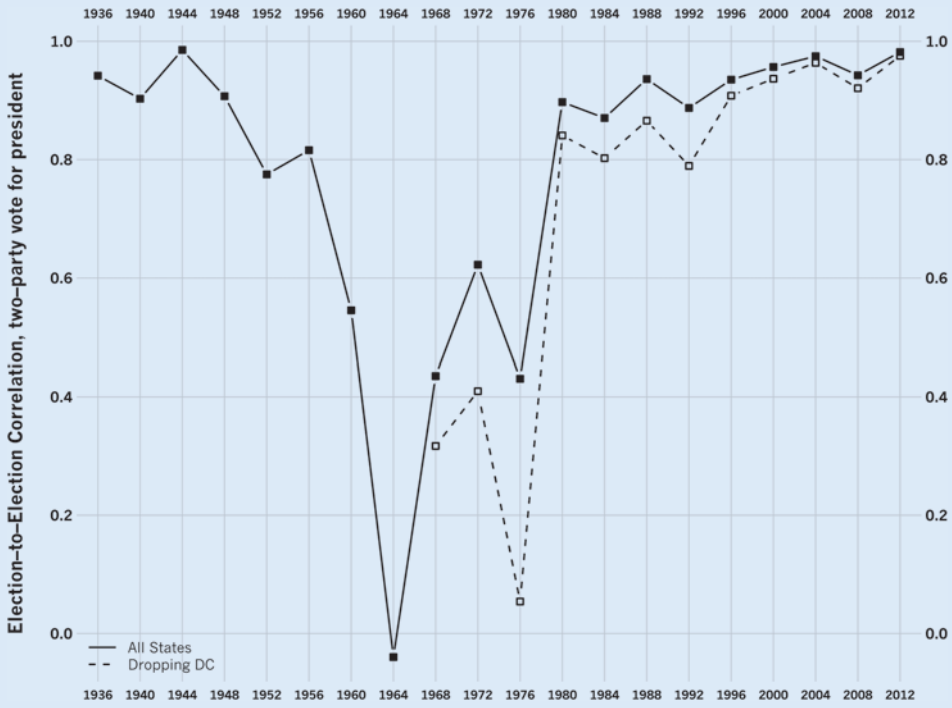
appeared on the *Washington Post* website through the 2012 election campaign[2] and a model of Achen and Bartels (2004).[3] I also include the "ensemble" model-average of Montgomery et al.; the Abramowitz model made the single largest contribution to the ensemble prediction.

State-level swings are never exactly uniform. Thus uniform swing will always predict state-level vote shares with error. The best that uniform swing could have done with respect to the RMSE criterion—even with perfect foresight—would have been to assume a national swing of -2.08 percentage points, generating 2.33 percentage points of RMSE (the minimum of the RMSE curve shown in figure 3). The actual, national swing of -1.48 is not the RMSE-minimizing uniform swing (RMSE = 2.41).

Two of the aggregate forecasting models—published well in advance of the election—fared very well on the RMSE criterion. Campbell's "convention bump" model and the Hill, Sides, and Vavreck model both predicted that Obama would win 51.3% of the two-party vote in 2012. Although slightly underestimating Obama's vote share (forecasting a national swing of -2.14 versus the actual swing of -1.48), the set of state-by-state predictions implied by this

Figure 1

Election-to-Election Correlation in State-Level Presidential Election Outcomes (Two-Party Vote Share), 1936–32 to 2012–08



The correlation between state-level, two-party vote shares in the 2012 and 2008 elections is .98, the highest election-to-election correlation of state-level vote shares observed since the 1944 and 1940 elections. Election results from David Leip's U.S. Atlas of Presidential Elections: http://uselectionatlas.org.

https://doi.org/10.1017/S1049096514000109 Published online by Cambridge University Press

national level forecast (after applying uniform swing) almost attained the RMSE lower bound (2.34, versus the lower bound of 2.33).

We also compare the performance of these models with respect to calling states correctly: that is, are the state-specific predictions of Obama two-party vote share $\hat{y}_{st}$ on the "right side" of 50%? For a given level of uniform swing, I compute the number of states "called correctly" and plot the resulting step function in figure 4.

That is, the models that perform best on RMSE (Campbell's "convention bump" and Hill, Sides, and Vavreck) are not the best performers on the "states correctly called" performance metric. This highlights a common problem in assessing election predictions: are we to assess the accuracy of point predictions or whether the state was "called correctly" (the point prediction was on the right side of 50%), or some mixture of the two? We defer a consideration of this question for another time.

Uniform swing has become a more accurate "first approximation" to state-level election outcomes in recent years. Accordingly, all serious efforts to forecast state-level election outcomes condition on the recent political history of any given state.

The RMSE-minimizing uniform swing of -2.08 incorrectly predicts that Florida and Ohio would be won by Romney. Treating the actual, national swing of -1.48 points as a uniform swing estimate generates 50 correct calls. Two of the aggregate forecasting models go 51 for 51, after converting their estimate of the national swing into a uniform swing estimate: Achen and Bartels and Erikson and Wlezien.

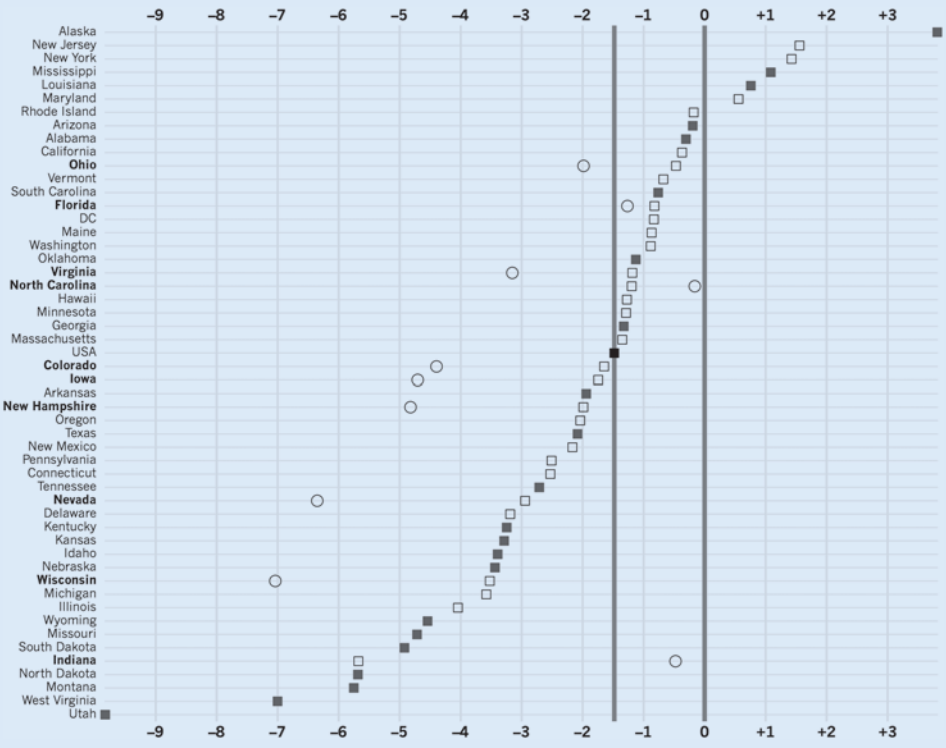
How well do these models fare against more elaborate alternatives, such as models that explicitly try to forecast state-by-state outcomes either well before the election, or by incorporating polls right up until election morning?

One of the models in the fall 2012 *PS* symposium made explicit state-by-state forecasts (Berry and Bickers 2012). The model also implied a national-level swing of more than 6 points, making it the second worst-performing model I consider here using the RMSE criterion (4.84) and suggesting that the model did not perform well state-by-state either. In fact, the Berry and Bickers state-by-state predictions have RMSE of 5.13, which is worse than the predictions one obtains from aggregating their state-by-state predictions and then applying uniform swing.

Models relying on poll averaging—using polls published right up until election morning—ought to perform considerably better than this. My own model, published on the "Pollster" section of the *Huffington Post* website[4], generated point predictions of Obama two-party vote share in 46 states. The RMSE of those predictions is 2.25 points; see table 1. Over those same 46 states, the best RMSE that uniform swing *could* have attained is 2.21. That is, my model generated 102% of

Figure 2

Swing, Obama Share of Two-Party Vote for President, 2008 to 2012, By State

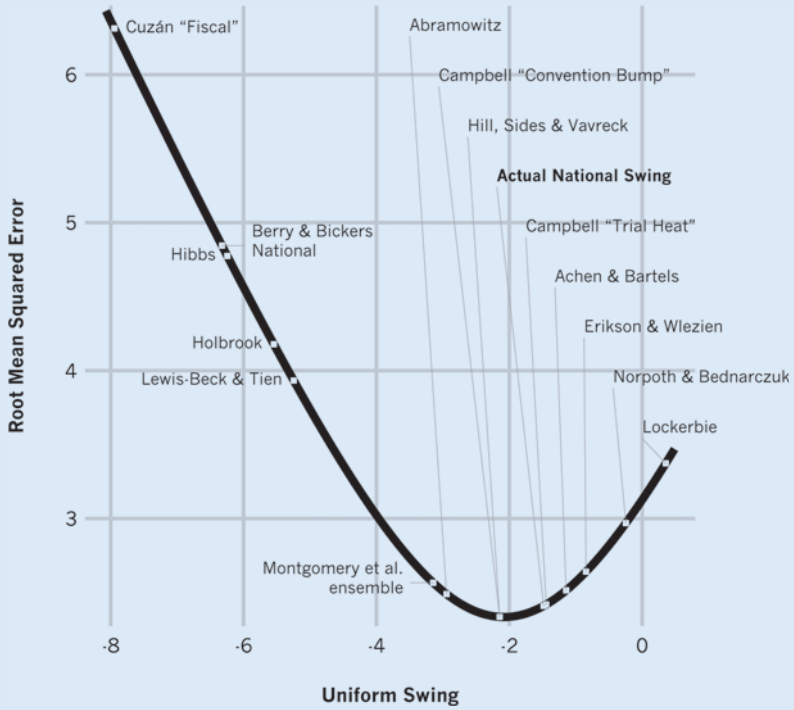


Swing, Obama share of two-party vote for president, 2008 to 2012, by state. Solid squares indicate states won by McCain in 2008. For nine ``battleground'' states plus Indiana (states won by Obama in 2008, labelled in bold type), open circles indicate the swing needed for Romney to win the state. Thicker, vertical lines indicate zero swing and the observed, national swing of -1.48 percentage points.

https://doi.org/10.1017/S1049096514000109 Published online by Cambridge University Press

Figure 3

Predictive Performance of UniformSwing, Root Mean Squared Error, with Respect to ObamaTwo-Party Vote Share, by State, 2012 US Presidential Election



Root mean square error (vertical axis) is a smooth, continuous function (black curve) of the assumed level of uniform swing (horizontal axis). The labeled points on the RMSE curve indicate the performance of various aggregate-level forecasting models (most of which presented their predictions in the summer 2012 PS symposium (Campbell 2012)), when interpreted as generating state-level predictions via the uniform swing assumption.

the RMSE generated by an oracle constrained to using uniform swing.[5]

In table 1, I compare the performance of other state-level forecasts to uniform swing. Nate Silver's point predictions had the smallest RMSE of the set considered here; the 1.98 point RMSE of Silver's predictions is just 85% of the best RMSE that a uniform swing model could have produced in 2012. Harry Enten did well too by this criterion, also outperforming the best that uniform swing could have done. Josh Putnam produced predictions in just 32 states, but also outperformed the best that uniform swing could have done in those states, at least with respect to RMSE.

Drew Linzer is similar to me on the RMSE performance metric, closely followed by DeSart and Holbrook. Wang and Ferguson do not fare well in this comparison: a relatively wide set of uniform swing assumptions would have generated state-by-state predictions outperforming theirs on RMSE.

This might seems like rather unimpressive predictive performance, at least for some of the poll-averaging models. On one hand, a model as naïve as uniform swing could have outperformed my model (and some of those like it), which incorporated thousands of polls right up until election morning. On the other hand, while *there exist* uniform-swing models with superior RMSE performance to poll-averaging models, knowing precisely what level of uniform swing to use ex ante in making predictions is more difficult. Note that the ensemble of aggregate forecasting models computed by Montgomery et al. did not outperform my model

Table 1

Predictive Performance (Root Mean Square Error), State-Level Forecasts of Obama 2012 Two-Party Vote Share, Selected Forecasters Using Poll Data

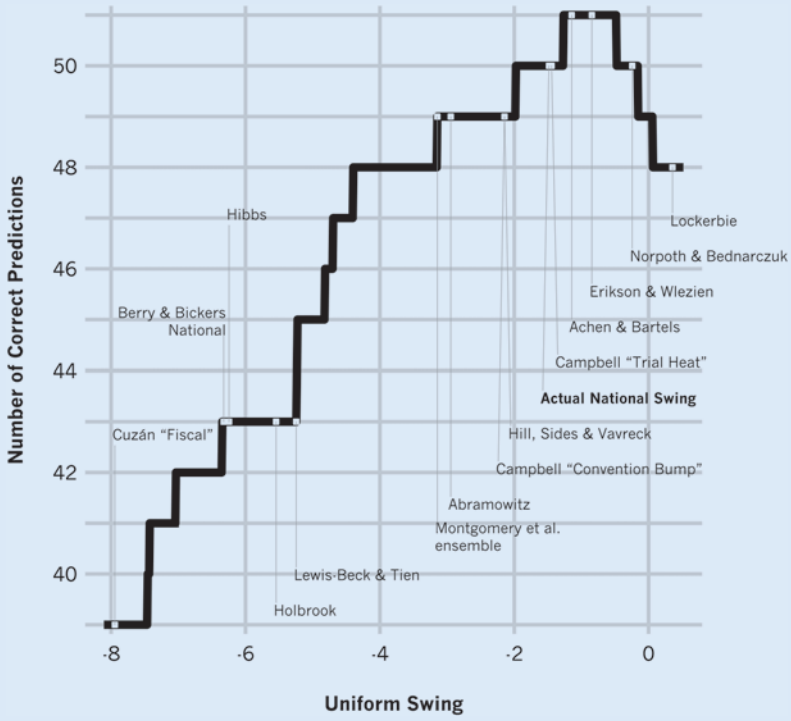
FORECAST	NUMBER OF STATES PREDICTED	RMSE	MINIMUM RMSE POSSIBLE, UNIFORM SWING	RMSE AS PERCENTAGE OF MINIMUM RMSE
Josh Putnam	32	2.11	2.30	92
Simon Jackman	46	2.25	2.21	102
Drew Linzer	50	2.42	2.35	103
Wang & Ferguson	51	2.82	2.33	121
DeSart & Holbrook	51	2.46	2.33	105
Harry Enten	51	2.19	2.33	94
Nate Silver	51	1.98	2.33	85

The minimum level of RMSE attainable under a uniform swing model is also reported, along with the percentage of the actual RMSE to the lower bound attainable under uniform swing. Two-party predictions from http://www.gwern.net/docs/2012-election-statemargin. csv, except Nate Silver's predictions from http://fivethirtyeight.blogs.

https://doi.org/10.1017/S1049096514000109 Published online by Cambridge University Press

Figure 4

Predictive Performance of Uniform Swing, Number of States Called Correctly, 2012 US Presidential Election



on RMSE or with respect to the number of states correctly called.

Uniform swing has become a more accurate "first approximation" to state-level election outcomes in recent years. Accordingly, all serious efforts to forecast state-level election outcomes condition on the recent political history of any given state. Considerable skill, effort, and data is required to improve the predictions made by translating the output of the better, national-level forecasting models to the state level via uniform swing. As in so many research programs, 90% of the effort delivers the last 10% of the results, since a reasonable estimate of the national-level swing plus the uniform swing assumption gets us most of the way there. ■

NOTES

1. E.g., Felix Salmon, "When quants tell stories": http://blogs.reuters.com/felix-salmon/2012/11/07/when-quants-tell-stories/; Tom Bartlett, "The Poll Quants Won the Election," *The Chronicle of Higher Education*: http://chronicle.com/blogs/percolator/the-poll-quants-won-the-election/31722.
2. See http://www.washingtonpost.com/blogs/wonkblog/post/create-your-own-election/2012/04/24/gIQAuaOIeT_blog.html for details.
3. Achen and Bartels model the margin between the two major party candidates, which I remap to be a prediction of Obama's share of the two-party vote. The 2012 prediction from the Achen and Bartels model appeared in a post by Bartels on the *The Monkey Cage* blog: http://themonkeycage.org/2013/01/08/obama-toes-the-line
4. See http://www.huffingtonpost.com/simon-jackman for details
5. Every forecaster has at least one call they wish they could take back. For me, it is Hawaii, where I underestimated Obama vote share by more than 6.7 points. RMSE punishes larger errors. If I added Hawaii to the list of states with relatively scant amounts of polling data where I did not produce a point estimate, my RMSE would have dropped to 2.04, handily outperforming any uniform swing model and many others, too.

REFERENCES

Achen, Christopher H., and Larry M. Bartels. 2004. "Musical Chairs: Pocketbook Voting and the Limits of Democratic Accountability." Typescript. Department of Politics, Princeton University. https://my.vanderbilt.edu/larrybartels/files/2011/12/musical-chairs.pdf

Bartels, Larry M. 1998. "Electoral Continuity and Change, 1868–1996." *Electoral Studies* 17 (3): 301–26. http://dx.doi.org/10.1016/S0261-3794(98)00035-3

Berry, Michael J., and Kenneth N. Bickers. 2012. "Forecasting the 2012 Presidential Election with State-Level Economic Indicators." *PS: Political Science and Politics* 45 (4): 669–74. http://dx.doi.org/10.1017/S1049096512000984

Campbell, James E., editor. 2012. "Forecasting the 2012 American Elections." *PS: Political Science and Politics* 45 (4) 610–74.

Linzer, Drew A. 2013. "Dynamic Bayesian Forecasting of Presidential Elections in the States." *Journal of the American Statistical Association* 108 (1): 124–34. http://dx.doi.org/10.1080/01621459.2012.737735

Montgomery, Jacob M., Florian M. Hollenbach, and Michael D. Ward. 2012. "Ensemble Predictions of the 2012 US Presidential Election." *PS: Political Science and Politics* 45 (4): 651–54. http://dx.doi.org/10.1017/S1049096512000959

https://doi.org/10.1017/S1049096514000109 Published online by Cambridge University Press