2018 State Legislative Election Forecasts

APPENDIX: Details of Forecasting Model

Carl E. Klarner, *Klarnerpolitics*

As in the earlier forecasts, state-years where one party had an overwhelming majority in the prior election are excluded from analysis, such as states in the one-party South. Unlike the prior analyses, odd-year elections are excluded, as the relationship between national conditions and election outcomes may be much different in these elections.

Numerous eccentricities that only occur for some state legislative elections conspire to substantially increase prediction error when all taken together. One example is when candidates switch parties--especially when they are sitting legislators--this is associated with a marked increase in vote share for the party they switched to. Party switching incumbents appear in about 1 out of 300 elections.

District level variables measuring whether there is a "repeat contest" between the same winner and loser in a single-member district have been added, along with an interaction between this variable and prior vote share: prior vote share has a greater impact in the context of a repeat contest. Similar interactions for "repeat winner" and "repeat loser" have also been added.

The fact that the incumbency effect has gone down over time was also modeled by including dummy variables representing periods, and interacted with incumbency.

"Supplemental lagged variables" for vote share, incumbency, etc., have also been added. These are generated from 1) nested or identical districts between the chambers, such as in Arizona, Nevada, Ohio, Washington or Wisconsin, 2) post-multimember districts and 3) alternating seats (the lagged value from two years ago for seats with four-year staggered terms, such as in the West Virginia State Senate). These are especially useful when standard lagged values are missing because of redistricting.

Because the model presented here is a model of change, lagged independent variables for all the variables should be included, as was done in the 2010 model. However, for the national variables, that brings about a large number of year-level independent variables relative to the number of years, which is only 25. First, doing so violates widely agreed on practices for modeling hierarchical phenomena. One general rule is that the number of high level predictors should not be greater than X times the number of cases in the level in question (). Next, studies of forecasting indicate that when sample size is small, assuming "unit effects" will yield more accurate forecasts than fitting the data to the sample. Fitting the data to the sample will cause random error to unduly effect the estimated impact of X on Y. As a result, a reduction of the four primary national level predictor variables to just one was conducted.

The four variables in question were z-scored to put them on the same scale. Next, a weight of "4" was given to Congressional vote intention, a weight of "3" was given to "midterm penalty," a weight of "2" was given to presidential approval, and a weight of "1" was given to change in real disposable income. The intuition behind this is that Congressional vote intention asks people directly how they will vote, and so it is given the highest weight, although the question is about voting for the U.S. House, not state legislative voting. Midterm penalty is given the second highest most weight as there is broad empirical support for the notion that the electorate engages in "ideological balancing" from a large number of contexts (thermostatic models, gubernatorial penalty, and U.S. House elections).

Presidential approval is given the third most weight as approval towards a specific president often does not translate into support for a particular party's state legislative candidate. This is especially the case if a president is not up for reelection, at least for U.S. House elections (Holbrook 2013), and so the weight of presidential approval is changed from "2" to "1" when the president is not up for reelection. It may very well be that a president who is not associated strongly in the public mind with one party or the other may also weaken the relationship between presidential approval and voting behavior. This may very well be the case for how voters perceive Trump.

Last, the state of the national economy is given the lowest weight, as the relationship between how people vote for state legislature and the state of the national economy does

not appear to be strong, once Congressional vote intention, the midterm penalty and presidential approval are taken into account.

2018 VALUES

Battle for control of the North Carolina legislature is a key focus of the 2018 elections. However, North Carolina's state senate and house maps were redrawn between the 2016 and 2018 elections. To deal with this, the correspondence between 2018 districts and 2016 precincts was used to generate lagged values for the 2018 elections. However, because of the high level of uncontested elections in 2016, precinct level vote share for 12 statewide offices (including president, governor, secretary of state, etc.) was used to predict state legislative vote share in competitive elections. Ballot roll-off in contested 2016 state legislative elections was similarly estimated. From this, lagged vote share at the state legislative district level was estimated.

A small number of incumbents who are running in the 2018 elections are neither Democrats or Republicans. It is assumed they will win reelection (once elected, non-major party candidates very often win, see Chamberlain and Klarner 2016), and it is assumed they will caucus with the same party as before the election.

SIMULATIONS

The heterogeneous nature of prediction error is also taken into account. Analyses indicate that prediction error varies greatly depending on what district level

information is available. The largest expected errors are those that involve missing values because of redistricting. At the other extreme, "repeat contests" see the lowest error. Numerous other situations exist in between, with one common example being uncontested prior elections, although the latter situations are not that damaging because they tend to occur in safe seats. The simulations that follow the generation of point predictions from the model take these varying district level error rates into account.

When Republicans caucus with Democrats, or vice versa, that is also taken into account. Coin flips are used in the two cases of such Republicans who are facing opposition in their primaries. The New York Senate's Independent Democratic Caucus has been disbanded, but a 50 percent chance of them breaking from the Democrats again was utilized in the forecast.[1]

---

[1] Probability based on assessment by expert on New York politics.