A Dynamic, Hierarchical, Bayesian approach to Forecasting the 2014 US Senate Elections

Gianluca Baio, Roberto Cerina

March 3, 2015

Every two years, American politics gives its signature performance: nationwide elections. The spectacularization of politics has produced multi-million industries born out of thin air to influence, explain and predict the vote. Recent elections have witnessed a staggering rise in the popularity of "trial-heat" polls as the standard prediction measure of election-day voter behaviour, with the number of polls conducted and available to the public increasing by over 900% since 1989. At the same time, models based on structural economic conditions have had remarkable success in predicting Presidential elections, with Alan Abramowitzs "Time for Change" (TFC) model famously achieving margins of error as low as 1.5 percentage points [1].

Modelling Senate elections, however, poses two main problems: the relatively low number of polls per state, leading to potential higher impact of the "house effect" [4] (favouring one side, depending on the polling house) and lower relevance of the economy. The former is endemic of mid-term elections: the stakes are perceived (maybe erroneously) to be lower and coverage by pundits and pollsters is thus reduced. The effect of the economy is weaker due to the lack of precise "blame" to be directed at the incumbent [2]. Whilst the President is considered to be responsible and hence heavily rewarded or penalised for the economic conditions, the same cannot be said for a Senator; thus, models such as the TFC are less relevant at the state level.

With the aim of analysing the 2014 Senate elections, we extended a modelling framework developed for the 2012 Presidential election [6]. The model is based on the integration of different sources of evidence to account for both the long- (e.g. the economic conditions) and the short-term effects (e.g. swings in the voters mood, mainly accounted for by the polls) — see Figure 1.

We model the Republican two-party vote share depending on a state-level effect capturing the performance of individual candidates and a national effect capturing the campaign effect. Technically, each effect is modelled using a *reverse-random walk*, where on election
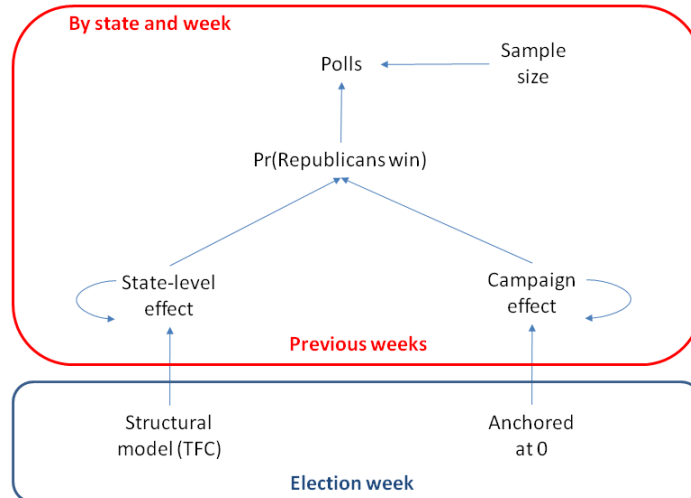
Figure 1: A graphical description of the model; for each state and week, the probability of a Republican win is informed by the observed polls and depends on the state-level and the campaign effects. These are modelled to account for the temporal correlation, with the values in the current week influencing the next ones. At election week, the campaign effect is assumed to wane at 0, while the state-level effect reverts to the long-term structural predictions.

day the state-level and the campaign effects are anchored at the structural forecast (a state-specific version of the TFC) and 0, respectively. Anchoring the state-level at 0 on election day encodes our belief that, at the end of the campaign, the candidates' efforts will become superfluous as most voters will have already made their mind up by that time.

Consequently, our forecast for election day will be some sort of compromise between the most recent polls and the historical predictions. When fitting the model close to the election week, the campaign effect will be already close to 0 and the state effect will not have much time to revert back to the historical forecast.Thus, the Republican election day two-party vote share will be determined largely by the polls. Conversely, when fitting the model months away from election day, the Republican share is largely determined by the structural forecast.

Because of the random walk specification, older polls are discounted in weight, but they leave behind their historical estimates of the state and campaign effects. This is interesting as it allows us to look at how voters behaviour evolved during the campaign. By comparing state and national effects we can assess how much this was due to campaign factors.

Figure 2 shows an overview of our model's performance. Out of the 33 races we monitor (excluding the uncontested Alabama race and the special elections in South Carolina and Oklahoma), only the forecast for North Carolina was on the wrong side of the 50% vote share line, suggesting that the model does a good job at predicting the "sign" of the election.

However, in key competitive races such as Virginia, Louisiana, Kentucky, Kansas, Georgia and Arkansas, the Republican vote share was under-estimated. We will see this was due to a skew in the polls, which led some of the predictions astray.

According to our model, the predicted probability of a Republican senate takeover was 94% by the end of election week. The most probable outcome is a Republican win of 19 seats — or a net gain of 7 seats, 1 more than they need to take over the Senate. The model assigns Harry Reid, the Democratic Senate majority leader, only a 6% chance of keeping his job.

In fact, the Republicans had a net gain of 8[1] seats, winning a total of 20 seats, which was rather improbable under our model.

A potential reason for such an under-estimation is perhaps that polls were heavily skewed in favour of the democrats [8]. As shown by Nate Silver on `fivethirtyeight.com` [13], polls bias is actually rather common and random, on average [11]. Sam Wang, at the Princeton Election Consortium, suggests an interesting hypothesis: polls were biased because Democrats did not turn up for the mid-terms. This is in line with political pundits, who argue that the "Obama coalition" does not turn up when the current President is not on the ballot [9].

We compared our results to other models using the Brier Score, a measure of accuracy for predictions of mutually exclusive, categorical outcomes [3], generally regarded as the standard in evaluating election predictions. Table 1 shows a ranking of predictors, as calculated by Sam Wang [2] and The Washington Post.

For our model, the computed Brier Score is 0.031, that is in the top 3. This is consistent with Figure 2, where the direction of the red wave was well captured.

Further research could focus on expanding our hierarchical specification to also include regional effects, such as a "South", "Great Lakes" or a "New England" effect on the Republican vote. The formal specification of appropriate distributions for these regional effects could help model common trends in regional voting blocks and give better prediction intervals. Other tweaks can be made to the distribution of the polling data and the structural forecast, to include third party candidates and independents. This would have been useful to avoid wrong estimations such as those made in the North Carolina race.

---

[1]The Louisiana "Jungle Primary" was modelled as a one off race between Democratic incumbent Senator Landrieu and the highest polling Republican contender, Bill Cassidy. At a later date, Cassidy defeated Landrieu in a run-off election to give the Republicans a total net gain of 9 seats.

[2]Sam Wang used a different scale from other predictors to evaluate his Brier Scores. Although on a different scale, it is reassuring to see that the ranking is consistent between the two columns.

Brier Score Calculations

Washington Post		Sam Wang (PEC)	
Daily Kos	0.024	Daily Kos	0.10
Washington Post	0.027	Washington Post	0.12
FiveThirtyEight	0.032	FiveThirtyEight	0.14
PredictWise	0.032	Betfair	0.14
Huff. Post	0.034	Huff. Post	0.14
Upshot	0.035	Upshot	0.15
Princeton	0.043	Princeton	0.18

Table 1: Ranking of pollsters by Brier Score, a measure of prediction accuracy. Our model would be third in the rankings as calculated by Sam Wang, worse than the Washington Post but better than `fivethirtyeight.com`. Drew Linzer, at the Daily Kos, takes gold again [5].

Elections are not random phenomena, but rather an expression of specific governing dynamics, which we can understand, explain and, eventually, engineer, at least to some extent. Whilst this is a frightening prospect for the future democracy, it promises to be an inexhaustible source of inspiration for the next generation of statisticians.

References

[1] Alan Abramowitz. Forecasting in a polarized era: The time for change model and the 2012 presidential election. *PS: Political Science & Politics*, 45(04):618–619, 2012.

[2] Christopher J Anderson. Economic voting and political context: a comparative perspective. *Electoral Studies*, 19(2):151–170, 2000.

[3] Glenn Wilson Brier. *Verification of a forecaster's confidence and the use of probability statements in weather forecasting*. US Department of Commerce, Weather Bureau, 1944.

[4] Andrew Gelman and Gary King. Why are american presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science*, 23(04): 409–451, 1993.

[5] David Jarman. Daily kos election outlook: One of those times when we hate to be right. Online, November 2014. URL `http://www.dailykos.com/story/2014/11/06/1342502/-Daily-Kos-Election-Outlook-One-of-those-times-when-we-hate-to-be-right`.

[6] Drew A Linzer. Dynamic bayesian forecasting of presidential elections in the states. *Journal of the American Statistical Association*, 108(501):124–134, 2013.

[7] Manu Raju and John Bresnahan. How sen. hagans husband won stimulus cash. Online, September 2014. URL `http://www.politico.com/story/2014/09/kay-hagan-husband-stimulus-cash-111339.html`.

[8] Douglas Rivers. Yougov poll performance in the 2014 senate elections. Online, November 2014. URL `http://today.yougov.com/news/2014/11/05/yougov-poll-performance-2014-senate-elections/`.

[9] Philip Rucker and Robert Costa. Battle for the senate: How the gop did it. Online, November 2014. URL `http://www.washingtonpost.com/politics/battle-for-the-senate-how-the-gop-did-it/2014/11/04/a8df6f7a-62c7-11e4-bb14-4cfea1e742d5_story.html`.

[10] Meredith Shiner. How the north carolina senate race became ground zero for what's wrong with 2014. Online, October 2014. URL `http://news.yahoo.com/how-north-carolina-s-senate-race-became-everything-that-s-wrong-with-201-193607993.html`.

[11] Nate Silver. The polls were skewed towards democrats. Online, November 2014. URL `http://fivethirtyeight.com/features/the-polls-were-skewed-toward-democrats/`.

[12] Karen Tumulty and Reid Wilson. Meet sean haugh, the libertarian pizza guy who may deliver a senate seat in n.c. Online, July 2014. URL `http://www.washingtonpost.com/politics/meet-sean-haugh-the-libertarian-`

pizza-guy-who-may-deliver-a-senate-seat-in-nc/2014/07/06/b321c03a-022f-11e4-8572-4b1b969b6322_story.html.

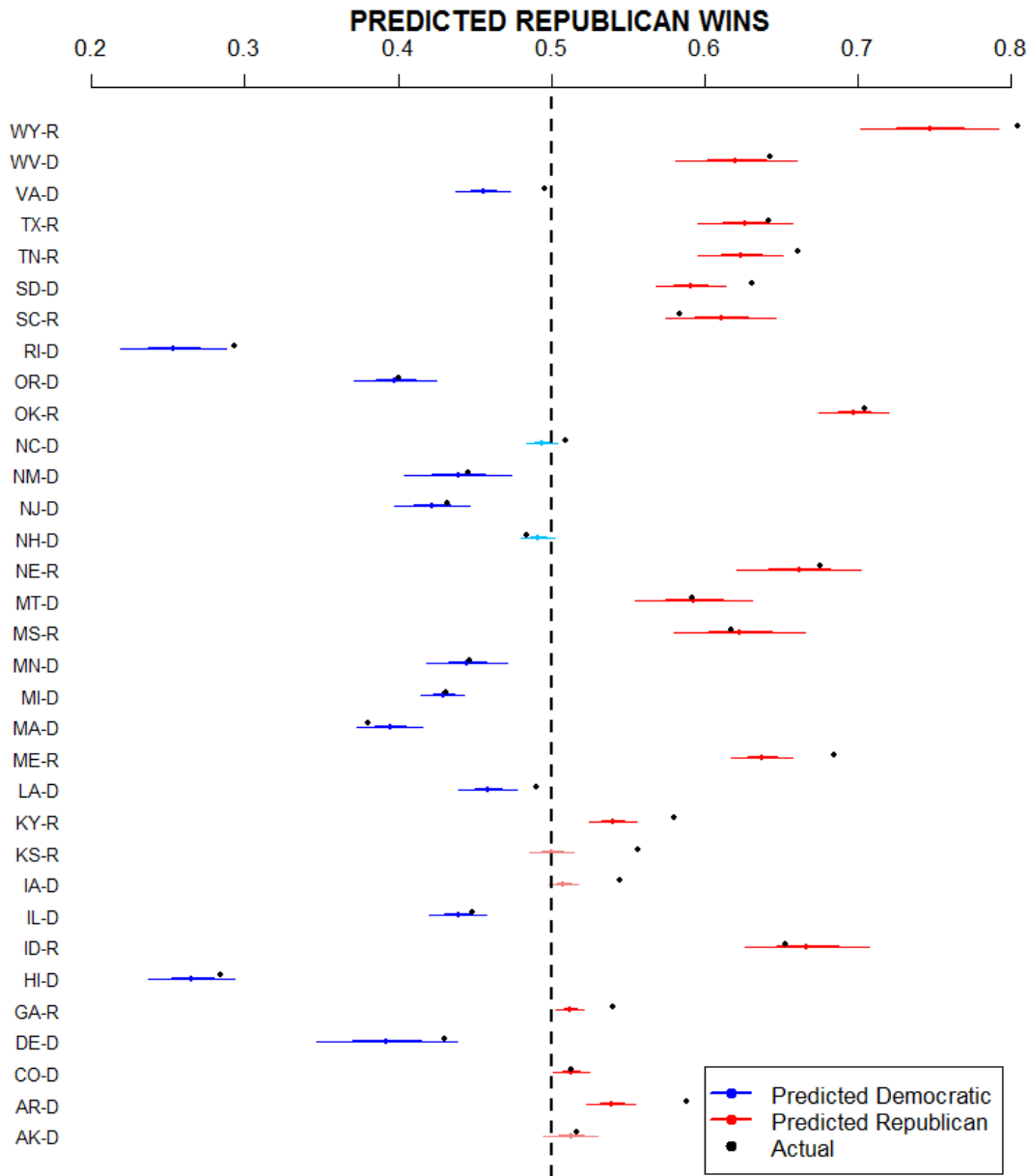[13] ESPN Internet Ventures. Fivethirtyeight. Online. URL http://fivethirtyeight.com/.

Figure 2: We compare the predicted two-party vote share distribution with the actual results. The prediction interval is reported as 2 s.d. around the predicted mean Republican vote share. A safe Republican seat is coloured red and is defined as the mean being at least 2 s.d. greater than the 0.5 cut-off; a likely Republican seat is coloured light red and is defined when the mean is larger than the 0.5 cut-off, but the lower tail touches the line. Democratic seats are defined in the same way, with a blue colour scale. The left axis contains the initials of the state in question and the initial of the incumbent party.