International Journal of Forecasting 31 (2015) 910–915

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A simple approach to projecting the electoral college

Joshua T. Putnam

Department of Government and Justice Studies, Appalachian State University, Boone, NC 28608, United States

ARTICLE INFO

Keywords:
Electoral votes
Public opinion polling
Presidential elections
United States
Weighted average

ABSTRACT

The following research note examines the utility of a simpler method of projecting the winners of the various states within the United States Electoral College system. While more advanced models may be able to center on state-level presidential winners earlier in an election year, those models, among others, continue to be confounded by states where the lead is small and/or not clear. This research will demonstrate that over the course of the elections from 2000 to 2012, a simple weighted average identifies state winners as well as the more complex models.

1. Introduction

For all the hype, talk, campaigning, and money spent, the 2012 United States presidential election was a rather pedestrian affair. Through the lens of the media, every tamale not properly shucked, every ill-advised tank-themed photo op taken, every Wendy's slogan borrowed on the campaign trail, or every unearthed-turned-viral hidden video was *the* "game change" that was going to make or break the election for one or both major party candidates at some point during the campaign season. Off that roller coaster ride, however, the fundamentals – roughly, presidential approval and some measure of the state of the US economy – often chart a different, steadier path toward election day (Abramowitz, 1988). In fact, in 2012, for all the noise that the various campaign events represented, the snapshot provided very early by the extant polling at the state level held very steady throughout.

This was true across a number of forecasting models that were employed, from very rudimentary polling averages to more complex Bayesian models, some of which accounted for additional variables beyond the state-level polling, while reducing the uncertainty of the predictions through simulations (see Linzer, 2013). Across the array of models, the picture was clear: the handful of swing states all favored President Obama, with the exception of North Carolina. In mapping out the Electoral College, the president had and held a 332-206 electoral vote advantage from June through election day. The one piece of the map that fluctuated across the various forecasts with any regularity was Florida.

In that scenario, then – the one where Florida is the only state to flip from one candidate to the other, and, in some cases, back again – those 29 electoral votes were neither decisive nor determinative. They were superfluous to the effort of Obama's reelection campaign to cobble together a coalition of states that summed to 270 electoral votes or more. In light of that fact, it was not much of a race at all. The president had, and maintained throughout the summer and fall campaign, leads in enough states to add up to more than 300 electoral votes, depending on Florida. In many respects, the 2012 US presidential election was a steady state. Despite fluctuations in the state-level polling, the picture always remained virtually the same in the aggregate. Obama always maintained some consistent cushion during the fall campaign.

However, that is not what was detailed through the news media over the summer and into the fall of the general election campaign. Nor did it follow the normal pattern witnessed in the lead-up to the three previous

E-mail address: frontloading.hq@gmail.com.

http://dx.doi.org/10.1016/j.ijforecast.2015.01.002
0169-2070/

presidential elections of the 21st century. The 2012 campaign saw little movement from the summer onward in the projected electoral vote count based on state-level polling. The previous three cycles, on the other hand, ended in November with different projections from those earlier in the campaign, in June. The 2000–2008 cycles had higher state-level polling volatility over the course of the campaign, but the movement across states was typically in the same direction and was still limited to just the handful of competitive states. To the extent that the electoral vote count fluctuated, it was a function of the small number of battleground states switching from one candidate's column to the other candidate's. That type of movement was absent from the aggregation of 2012 polling. There were fluctuations across the various survey snapshots taken in 2012, but those changes were muted relative to previous cycles.

Together, those four elections have provided forecasters with a wide array of conditions under which to project electoral college outcomes. At their most basic, the four elections are a reasonable cross-section of presidential elections during the polarized era, given that, collectively, they represent an open seat election following a two-term Democratic incumbent (2000), one with a Republican incumbent (2004), an open seat contest following a term-limited Republican incumbent (2008), and a reelection campaign involving a Democratic incumbent (2012). Not only do the four elections account for different types of incumbency and partisan control of the White House over this period, they also coincide with an era that has seen an overall proliferation of publicly released polling data.

However, the publicly released polls can fluctuate wildly over time and among the various polling firms surveying registered and likely voters across the country. At a bare minimum, simple polling averages and other more complex methods of aggregation and analysis can smooth out the crests and troughs in the data to give a clearer, more stable picture/forecast of where a presidential race stands at any given point in time during an election year. This paper will establish the rationale behind one of these methods, the graduated weighted polling average (GWPA), and describe its mechanics, discuss the reasons for inter-model convergence with the GWPA in 2012, and, finally, examine the effectiveness of averages as a baseline presidential election projection over the period 2000–2012.

2. A simple electoral college projection

The graduated weighted polling average (GWPA) works on the assumption that survey data are an adequate proxy of the state of the presidential race. Polls account for and are a reasonable reflection of the oft-cited fundamentals in the race for the White House, as well as of campaign effects and other events. Furthermore, the GWPA treats all polling data as "good" polling data; that is, good in the sense that each new survey provides an update on the actual state of the general election race in a given state at any point in the campaign. The snapshots vary in their accuracy, but are more powerful – more sound – when they are aggregated and averaged over time. On the whole, newer information is more valuable in the battleground states, but also helps to identify which states are the swing states over the course of any election cycle. For instance, in 2012, Michigan saw the average margin between the candidates widen over time, making it less of a Romney target in the process. Conversely, just four years earlier, North Carolina became more competitive as the campaign continued, putting the Tar Heel state on the Obama campaign radar.

Thus, any information is valuable to the graduated weighted polling average. However, the recency of the data is also of value. The timing serves as the only "variable" that is accounted for in the projection model directly. The rest are subsumed in the actual polling data or left unaccounted for in the projection.

The polling data are paramount, and are only corrected for by a decay function that attempts to account parsimoniously for how old the information is. In the context of the graduated weighted polling average for a state, the candidates' levels of support (or "vote shares") in the most recent poll are given a full, undiscounted weight, while an older survey is multiplied by a decay function. That discount is calculated by dividing the day of the year in which the poll was last in the field by the day of the year at which the campaign currently is. For example, election day 2012 was on November 6, the 311th day of the year. A poll that was last in the field in Arizona on January 9 – the ninth day of the election year – as a Behavior Research Center poll was, would be discounted significantly in the average, due to its age on election day (de Berge, 2012a). That is, it would have been reduced by a decay function of 0.029 (or 9 / 311). A poll closer to the day of the election – like the early October survey of Arizona by the same firm – would have been discounted in the GWPA at a less severe rate of reduction. This later Behavior Research Center survey, being last in the field on the 284th day of the election year, would have retained a little more than 91% of the value of the candidates' "vote shares" in an election day average of all weighted surveys (de Berge, 2012b).

This is a very gradual decay function. In addition, the discount is being utilized as part of an average. The former means that the "vote shares" in the polls decay very slowly the older they get, while the latter indicates that, in the absence of outliers, the aggregation of polling data over time will translate into a slowly reacting gauge of aggregated public opinion on the race in a state. The slow reaction time can be viewed as a drawback of the GWPA approach, as it fails to keep pace with changes in the race. From another perspective, however, the slow reaction time can also mean that when a change occurs – specifically, a state's average switching from favoring one candidate to the other – it is evidence of a meaningful and lasting change in the state of the overall race.

These conflicting perspectives highlight the advantages and disadvantages of the GWPA relative to the extant, alternative models. The 2012 presidential election cycle offers one such lens through which to examine the various prediction models. There are at least two overarching and interrelated lessons that can be gleaned from 2012.

(1) First of all, despite the fact that a number of models predicted the winner of the race for the White House and the distribution of electoral college votes for each candidate correctly early on, and showed little if any