International Journal of Forecasting 31 (2015) 895–897

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Editorial

Forecasting US presidential elections: New approaches (an introduction)

We are pleased to introduce this Special Section on US presidential election forecasting, with its focus on the closely-watched 2012 contest between Barack Obama and John McCain. Forecasters bore down on the race as never before, and much was learned. The *International Journal of Forecasting* (IJF) has an abiding interest in the task of election forecasting. Besides the ongoing publication of individual papers, they have published three separate special sections or issues on the forecasting of elections around the world. Most recently, the forecasting of elections in "neglected democracies" was addressed (Bélanger & Lewis-Beck, 2012). Prior to that, attention was turned to European election forecasting (Jérôme & Lewis-Beck, 2010), and earlier still, US presidential elections were given a special issue (Campbell & Lewis-Beck, 2008). Thus, the papers at hand provide an opportunity to update our knowledge, particularly with respect to the pivotal US case.

Election forecasting is a natural outgrowth of social scientific efforts to understand why elections turn out the way they do. It integrates research into voting behavior, public opinion, electoral systems and political economy, and increasingly makes use of many cutting-edge methodological techniques in the social sciences. When producing quantitative election forecasts, we are putting our understanding of these fields to the test: if our theories are correct, and our data are well-measured, we should be able to make accurate predictions about the future, with well-calibrated statements of uncertainty.

The rigor of quantitative models is one of their virtues. In the run-up to an election, political commentators frequently make ad-hoc arguments about which candidate should be considered "ahead" or "behind" in their chances of winning, and why. These arguments are often reasonable, intuitively appealing, and even plausible. However, the problem for news consumers is that they lack any systematic, historical basis for evaluating whether or not the arguments are actually true.

In contrast, statistical election models are transparent about their underlying assumptions, the data from which they are extrapolating, and how those factors combine to generate predictions. If and when different models produce divergent election forecasts, we can tell which factors account for the discrepancies. Alternatively, when models with different assumptions lead to similar forecasts, it can give us confidence in the predictions. By focusing our attention on a small set of key predictive elements, the models help us to avoid being misled or distracted by events during a campaign that seem like they "should" matter for an election outcome, but cannot be verified empirically to have had much impact.

Quantitative election models also make it possible to investigate how the predictions would be affected by changes in either the assumptions of a model or measures of current conditions. How sensitive are our expectations to the assumptions we have made about the data? We can even judge the quality of the forecasting models themselves from the theories they are based upon, the specification of the models, and the selection of data being used.

Finally, by basing models on mathematical equations, researchers can collaborate to adjust and improve their forecasting models over time. Indeed, the collection of papers contained in this special section is reflective of a community of researchers building iteratively on previously established techniques. To forecasters, an election is not merely another data point to feed into their models, but an opportunity to test and refine the models themselves. If a quantitative election forecast turns out to be wrong, quantitative forecasters can track down and isolate the error so as to avoid repeating the same mistake in the future. As new and better data sources become available, we can gradually determine how useful that information is for identifying the correct election outcome as early as possible.

Quantitative election forecasting is in an exciting period. The state of the art is moving beyond its traditional dependence on static regression models fitted to small historical data sets. At the same time, there has been a surge of public interest in scientifically grounded election forecasting as a way of following and learning about political campaigns. This interest has grown both in response to new

http://dx.doi.org/10.1016/j.ijforecast.2015.03.004
0169-2070/

developments in academic research and thanks to the increased visibility of political "data journalism," especially online, which has brought statistical forecasting models into the media mainstream.

Why is the US case pivotal? First, and obviously, it is the world's most powerful democracy. Second, and less obviously, it is home to most of the methodological innovation in the field. Simply stated, the American case has become the principal testing ground for new ways of forecasting elections. These nine papers offer multiple tests of leading approaches, old and new, which can be grouped by whether they rely on polls, structural models, prediction markets, or other media. For an intermediate, non-technical introduction, as applied to the 2012 US presidential competition, the reader may wish to consult the collection of papers edited by Lewis-Beck and Stegmaier (2014a). However, for advanced, cutting edge treatment, the reader should consult the papers in this volume. Below, we introduce each of them in turn, by author name.

These efforts represent the best in "second generation" work, with an emphasis on sophisticated, dynamic forecasting of the Electoral College vote, drawing on large datasets from the American states. We begin with the use of polls as forecasting instruments, first considering the reliance on the traditional vote intention question, then looking at voter expectations. From there, we turn to structural models, starting with simple "fundamentals" models and moving to more elaborate ensemble models, including combinations of structural models with polling and prediction market data. Finally, we address new approaches to the measurement of public sentiment using non-probability polls and social media data. In each case, the authors examine both the strengths and the weaknesses of their chosen approach.

Polling results from a vote intention question (e.g., "Who would you vote for if the election were held today?") have long served as a mechanism for forecasting the winner. Early studies followed the poll results from one firm, usually Gallup. However, the dominant strategy at present is to aggregate the data from publicly released polls by multiple polling organizations. Sam Wang, of the Princeton Election Consortium, looks at vote intention polls from the American states, in order to forecast the winner using what he calls his electoral vote estimator. Fixing on the median poll for a time period, he claims a very high accuracy, with an error of less than one percentage point. He has applied his "meta-analysis" to the 2004, 2008, and 2012 elections. Putnam also works this vein, in his application of graduated weighted polling averages, by state. His approach, which increases the weight given to polls closer to Election Day, forecast each Electoral College winner correctly, state-by-state, in 2012. Further, following elections back to the 2000 contest, he never got more than three states wrong.

A rival polling approach does not use a vote intention question. Instead, it finds predictive value in a vote expectation question (e.g., Who do you think will win the presidential election?). Murr terms this "citizen forecasting," and applies it to surveys conducted by the American National Election Study (ANES), state-by-state. For 2012, his methodology forecasted 49 of the 51 states correctly (including DC). Looking back further, the method predicted 8 of the last 11 presidential elections correctly. Murr attributes the success of this approach to the "wisdom of the crowd" empirics, which are based on Condorcet's jury theorem.

Three papers consider methods for improving fundamentals-based forecasting models by combining the results across model specifications or data sources. A typical example of a "structural" election model for forecasting the national presidential election vote from macro-political economy variables might be expressed as vote $= f$ (*economic growth, presidential popularity*). Montgomery et al. describe a technique for overcoming the small sample sizes and the instability of any single model specification by combining forecasting models using calibrated ensemble Bayesian model averaging (EBMA). The approach of Graefe et al. also combines the results of several distinct structural models to arrive at a final forecast. They propose calculating a simple average of the constituent forecasts, which, as they demonstrate, can be even more accurate than the more technical EBMA. Rothschild, instead of combining different scores from a single method (such as structural models), combines single scores from three different methods—aggregate polling, structural models, and prediction markets. He calculates a daily score for each, over a 130-day pre-election period, then combines them into a single daily forecast. He finds that each data source can contribute incrementally to the accuracy of Electoral College forecasts.

Lauderdale and Linzer assess the power of structural models for forecasting presidential elections, whether interpreted individually or not. They demonstrate that although typical regression-based models are derived from voting theory, they nevertheless tend to understate the uncertainty in their forecasts. Introducing a Bayesian forecasting model that takes large numbers of predictors at both the national and state levels, they find that – given the amounts of historical data currently available – the 95% prediction intervals around presidential election forecasts should be approximately $+/-10\%$ at the state level, and $+/-7\%$ at the national level, which is considerably wider than what is commonly reported.

Part of the challenge in generating accurate election forecasts is the collection of data on predictive factors that are both well-measured and systematically related to election outcomes. For some approaches, the measures themselves are biased or highly noisy; in other cases, quality measures are available but have minimal or inconsistent predictive power. Rothschild addresses this dilemma briefly in his discussion of whether or not prediction markets contain any information beyond the polls themselves (when properly exploited). Likewise, as Wang notes, the number of public polls is dwindling, and methodological challenges in the modern polling industry raise questions about whether aggregating polls is enough to eliminate any bias and reduce the variability in the underlying data, whether due to declining response rates, pollster "house effects," or other factors.

As the polling industry experiments with survey methodologies that deviate from simple probability sampling, we expect the question of poll quality to be a recurring one. Telephone surveys, even when based on random draws from lists of phone numbers, routinely need serious weighting in order to be made "representative." Online surveys, in which respondents are often "opt-in" volunteers, pose even greater problems of inference. On this score, however, Wei Wang et al. are optimistic. Their paper demonstrates that, with proper model-based adjustments, even patently non-representative samples can be made to generate accurate forecasts. They base their argument on an analysis of an opt-in survey conducted on the Xbox gaming platform, which ran continuously for 45 days before the 2012 election. The raw survey data were biased heavily in terms of age and gender, but by combining multilevel model estimates with 2008 exit poll data, the authors' results were shown to predict the 2012 election outcome successfully.

Some researchers who are less optimistic about the continued, effective use of polls for forecasting elections have begun to explore the potential value of social media data. In his paper, Huberty expresses considerable skepticism about this approach. His stark conclusion is that "All known attempts at election forecasting with social media have failed." Huberty's own attempt, reported here, is no exception. He investigated Twitter mentions of candidates in various 2010 House of Representatives races. From these tweets, he built a forecasting model which he tested against the 2012 contests. Tellingly, it could do no better than the simple rule: predict that an incumbent will win. According to Huberty, this is due to the absence of control over the data-generating process. Again, we are back to sampling problems.

Conclusions

The strategy of poll aggregation, by state, appears to have yielded highly accurate US presidential election forecasts, both for 2012 and for recent previous elections. The success of the structural models is mixed, but they clearly work better when they are combined. Both strategies face sampling issues. For polls, the worry is over the declining quantity and quality of the surveys. For the structural models, the worry is over the small number of presidential elections available, e.g., $N = 16$ (1952–2012). An emerging strategy here, as was mentioned by Rothschild, is the combination of approaches. Elsewhere, we have explored a combination strategy in "synthetic" models. We start within the framework of a structural model, then add aggregated polling data, with updated forecasts being offered as the election approaches. (For US presidential elections, see Linzer, 2013; for European elections, see Lewis-Beck & Dassonneville, 2015.)

The forecasting efforts discussed in this special issue are statistically complex, some extremely so. Could the same accuracy be achieved with simpler models? As Lauderdale and Linzer note, the simple Incumbency Rule (the incumbent will be re-elected) is hard to improve upon, if analysts confine themselves to the world of structural models. Huberty also observed that the Incumbency Rule works extremely well for forecasting House elections. Elsewhere, Abramowitz (2014) reported that, at least for 2012, an incumbency rule would have predicted the Electoral College vote winner in 48 of the 50 states (i.e., simply predict that the party with the leading vote share in the state in 2008 would win in 2012).

Of course, accuracy is not everything, especially when it is confined to one contest. The models explored here are based on strong statistical and voting theory, and would be expected to perform better than simple rules-of-thumb, especially in the long-run. For evaluating forecasting instruments, we elsewhere offered the following criteria: accuracy, parsimony, transparency and lead (Lewis-Beck & Stegmaier, 2014b).

When examined in this light, the complex models offer a tradeoff. While they may achieve a greater accuracy, they incur costs in parsimony, transparency, and lead. Given the big-data requirements and the modeling uncertainties that face US presidential election forecasting in the 21st century, it may be that parsimony and transparency are hard to attain in practice (although they remain worthy goals). However, the criterion of lead remains, and is difficult to dismiss. Forecasting is about seeing into the future. When the election looms far away and we forecast the outcome accurately, that is impressive. However, if no accurate forecast can be made until on the heels of Election Day itself, the usefulness of the forecasting method is reduced. A next step in US presidential election forecasting research would be to focus on optimizing the lead time of our forecasts.

References

Abramowitz, A. (2014). Long-term trends and short-term forecasts: the transformation of US presidential elections in an age of polarization. *PS: Political Science and Politics*, *47*, 289–292.

Bélanger, É., & Lewis-Beck, M. S. (Eds.) (2012). Election forecasting in neglected democracies (special section). *International Journal of Forecasting*, *28*(4), 767–892.

Campbell, J., & Lewis-Beck, M. S. (Eds.) (2008). Presidential election forecasting (special issue). *International Journal of Forecasting*, *24*(2), 189–322.

Jérôme, B., & Lewis-Beck, M. S. (Eds.) (2010). European election forecasting (special section). *International Journal of Forecasting*, *25*(1), 19–115.

Lewis-Beck, M. S., & Dassonneville, R. (2015). Forecasting elections in Europe: synthetic models. *Research and Politics*, *2*(1), http://dx.doi.org/10.1177/2053168014565128.

Lewis-Beck, M. S., & Stegmaier, M. (Eds.) (2014a). United States presidential election forecasting (special section). *PS: Political Science and Politics*, *47*(2), 284–347.

Lewis-Beck, M. S., & Stegmaier, M. (2014b). United States presidential election forecasting: an introduction. *PS: Political Science and Politics*, *47*(2), 284–288.

Linzer, D. A. (2013). Dynamic Bayesian forecasting of presidential elections in the states. *Journal of the American Statistical Association*, *108*(201), 124–134.

Drew Linzer
Department of Political Science, Emory University, United States

Michael S. Lewis-Beck*
Department of Political Science, University of Iowa, United States
E-mail address: michael-lewis-beck@uiowa.edu.

* Corresponding editor.