

Issues in Presidential Election Forecasting: Election Margins, Incumbency, and Model Credibility

James E. Campbell, *University at Buffalo, SUNY*

There is no shortage of interesting issues surrounding election forecasting. From a long list of fascinating matters worthy of rumination, I have selected three. The first concerns the consequences of shrinking presidential vote margins. The second deals with the incumbency advantage and its implications for model estimation and democratic accountability. The third addresses often neglected criteria for assessing the credibility of forecasting models.

ELECTION MARGINS

Since the late 1980s, presidential elections have been particularly close. From the late 1940s to the early 1980s, a period covering many elections used in estimating most forecasting equations, the two-party popular presidential vote varied from a high of about 62% to a low of 38%, a 24-point range. In the seven presidential elections since Reagan's 1984 landslide, however, no candidate has received as much as 55% of the two-party popular vote, a 10-point spread. The standard deviation of the two-party vote in the 10 elections from 1948 to 1984 was 6.8 percentage points. In the post-1984 period, it has been only 3.1 percentage points.

Assuming that the more limited range of outcomes in the last seven elections is not a fluke, but a consequence of changed underlying political realities—greater polarization of the public and parties and a more competitively balanced realigned party system (“sorted” for the squeamish) (Abramowitz 2010, Abramowitz and Saunders 1998, 2008; Campbell 2006, 2010)—the hyper-competitiveness of post-1984 presidential elections may have some consequences for the general accuracy of the models as well as their specification and estimation. On the one hand, one would expect narrower vote margins would increase forecast accuracy. Forecasts within the new 55–45 limits should be more accurate than within the old 62–38 limits. This narrower band of plausible forecasts may also increase expectations of greater accuracy.

On the other hand, applying models estimated using mostly wide-range elections to post-1984 narrow-range elections presents a problem: a unit change in a predictor is likely to have had a larger impact on the vote in earlier wide-range elections and, thus, a larger coefficient than would be applicable in more recent vote-constrained elections. The normal fix, a host of interaction terms, is not feasible with so few elections.

This problem is likely to be greater for some forecasting models than others. To the extent that a forecast is based on

opinion predictors, either preference polls or approval ratings, polarization probably has already been taken into account. Preference polls and approval ratings should be affected by the same polarization that affects the vote. This cannot be said, however, for other predictors such as the economy and incumbency. With each additional election from the polarized party era included in updated estimations, the impact of the vote margin change should lessen. But for now, to the extent that models do not already take this change into account or adapt to do so, we might expect elections to be a bit closer than predicted.

An unfortunate irony in this is that this problem has probably blurred differences between strong and weak models by strengthening the apparent performance of weak models and weakening the performance of strong models. Consider a rubbish-model, essentially noise. It looked good by goodness-of-fit statistics, but that was only luck. In use, the predictions of bad models tend toward the mean vote (52.0% since 1948). Since the polarized realigned system has moved the playing field in the same direction, bad models will tend to be more accurate than they would have been otherwise. In contrast, strong models that have not taken polarization into account in their specification will produce forecasts more appropriate to earlier higher margin elections. Their errors will be larger than they would have been otherwise. In effect, without adaptations to post-1984 hyper-competitiveness (and poll based models have built-in adaptations to this, at least partially), weak models appear stronger and strong models appear weaker than they might otherwise have been, at least until we grow out of the problem with the inclusion of more post-1984 elections.

PRESIDENTIAL INCUMBENCY

One of the clearest contributions to voting behavior research from election forecasting is in illuminating the impact of presidential incumbency (Campbell 2000, 2008). The effect of incumbency has been reflected explicitly in Norpoth (2000), Abramowitz (1988), Fair (1988), and others and implicitly through other variables (approval ratings and preference polls) in many other forecasting models.

Based on our forecasting experience, the incumbency advantage appears to be more complex than often thought. The effect of incumbency is not captured, not even remotely, by a dichotomous variable for a sitting president. It has as much to do with the tenure of the party in office as with the person in office. It also is enmeshed with the standards that voters use in retrospective voting. Voters cut some in-party candidates more slack than

Table 1

Election Results for the In-Party Presidential Candidate, 1900–2012

IN-PARTY CANDIDATE AND ELECTION TYPE	TWO-PARTY POPULAR VOTE PERCENTAGE		
	Won	Lost	Mean Vote
Not Incumbent, Open Seat	4	5	49.3
Incumbent, First-Party Term	11	1	56.3
Incumbent, Not First-Party Term	4	4	49.1

others. In-party nonincumbents (successor candidates) receive only partial credit or blame for their party's past performance. For incumbents, retrospective evaluations of the record appear to depend a good bit on how long the party has occupied the White House.

How might the system function differently if everyone understood that putting a new party in the White House is tantamount to an eight-year commitment?

As several forecasting models suggest (Abramowitz 1988; Norpoth 2000) and as table 1 attests, the presidential incumbency advantage goes largely, if not exclusively, to first party-term incumbents (presidents whose party has occupied the White House for just one term) (Campbell 2013a, 2013b). Since 1900, there have been 12 first party-term incumbents. Eleven of the 12 won reelection. Only one (Carter) was defeated, and he entered the election with the economy in steep decline. Barring abject failure, first party-term incumbents are virtually assured of a second term. Other in-party candidates (successor candidates in open seats and incumbents of parties in office for more than one term) seem to have no discernable advantage. The first party-term advantage appears to be so strong that all of the models should determine whether they adequately take it into account in some way.

First party-term presidents are in the enviable position of being able to credibly campaign either advocating stability if things are going well or advocating change if things are going poorly. Having been in office for just four years, these incumbents can still plausibly blame their predecessor for persisting problems (Campbell 2008). They are credited for their successes, but can evade a good deal of the blame for their failures. In effect, retrospective evaluations may be asymmetric for first party-term incumbents.

The 2012 election is a textbook example. Although a good case could have been made that the incumbent's policies had led to a weak economy or, at least, had failed to set things back on course, a majority of voters placed more of the blame on the president four years out of office rather than the president in office those four years. Several late campaign polls and the exit polls indicated that a slight majority blamed former President Bush for the weak economy and only just over a third blamed President Obama (Campbell 2013b, 26).

The incumbency advantage, both in the partial credit or blame assigned to successor candidates and the big first party-term advantage, raises some interesting questions about electoral accountability and interpreting elections. Should voters hold successor candidates more accountable for the national conditions left by their party? Are the standards of responsibility for first party-term incumbents so lax that they can evade accountability in all but the most egregious cases? How might the system function differently if everyone understood that putting a new party in the White House is tantamount to an eight-year commitment?

MODEL CREDIBILITY

To the uninitiated, election forecasting is often caricatured as a barefoot empiricist's atheoretical search for the best fitting equation, the highest R-square. While there may be some truth to this critique in a few cases, in general, it is off base. This misperception is unfortunately fed by the all too often exclusive attention paid to a model's fit to past elections. I have no argument with judging forecasting models by their fit to past

elections, that is essential, but I take issue with that being the sole criterion. Accuracy standards are important, but as Lewis-Beck observed (2005), they are not the whole story. Many factors add to or detract from the credibility of a forecasting model beyond its out-of-sample goodness-of-fit statistics. Unfortunately, these are not neatly captured by a single statistic or index.¹ As a result, these factors are often neglected. Somehow these other elements should be considered in recognizing a model and its forecast as being credible. It should be observed that the neglect of credibility criteria beyond model fit statistics is by no means a problem limited to forecasting.

Beyond a model's general accuracy, five criteria are important to a model's credibility. The list augments that originally assembled by Lewis-Beck (2005), sans his index. The first criterion is transparency. Unless we know exactly what has gone into a forecast, it should have about the same credibility as fortune-telling. The ultimate transparency is the availability of the data and replicability of the forecast model's estimates.

The second criterion is the simplicity and logic of the predictors. Models that have straightforward measurements and sensible weights for predictors are more credible than those with complex or even Rube Goldbergesque indices.

The third metric is model stability and the track record of an unrevised model over a series of elections. An unchanged and fairly accurate forecasting model should be considered more credible than a one-hit wonder or frequently tweaked model. Model revisions are necessary from time to time, but consumers have a right to be a bit wary of models that are frequently "new and improved."

Relatedly, the fourth criterion is whether a model is supported by closely corresponding companion models. These are models in the same vein as the central model, but are differently timed or

configured in their predictor variables and are “well behaved” or have expected coefficient differences. These companion models provide evidence of a model’s robustness and bring additional independently measured data to the forecast. For example, the Trial-Heat and Economy Model was estimated using successive presidential preference polls at different points in a campaign (Campbell and Wink 1990). As expected, as the forecast moved closer to Election Day, the contribution of the preference poll to the prediction increased and the contribution of the economic indicator decreased. This pattern of effects is exactly what one would expect as the state of the economy gradually becomes fully incorporated into the voters’ preferences.

Last, but not least, forecasting models are more credible if they are consistent with existing empirical explanatory findings or, at least, not inconsistent with explanatory research. Forecasting models need not be based entirely on empirical theory, they have different goals, but their credibility should suffer if contradicted by empirical findings. To the extent that forecasting models are supported by empirically confirmed theory, credibility is substantially enhanced. For example, both prior and subsequent research on campaign effects (Campbell 2008; Erikson and Wlezien 2012; Lazarsfeld 1944), economic effects (Campbell 2008; Campbell, Dettrey, and Yin 2010; Holbrook 2008; Lewis-Beck 1988; Nadeau and Lewis-Beck 2001; Norpoth 2002; Vavrek 2009), and incumbency (Campbell 2008; Mayhew 2008; Weisberg 2002) corroborate the Trial-Heat and Economy Model. A forecast model well integrated into empirically supported explanatory theory at both the micro and macro levels should be accorded greater credibility than one lacking that corroboration.

OUTLOOK ON 2016

Although directed at the long-term development of forecasting, the three issues presented here have more immediate implications as well. In looking toward 2016, the hyper-competitive realigned and polarized partisan context as well as its open-seat status set the stage for another very close election, closer perhaps than some of the forecasts will indicate. In terms of model credibility, although forecast consumers cannot conduct a credibility review of each forecast, they can assess its plausibility. Beyond generating a prediction, forecasters should explain why their models work. If the explanation of a forecast seems too convoluted or contrived to be believable, it probably is. ■

NOTES

1. Lewis-Beck (2005, 154) constructed a “quality index” consisting of four elements: accuracy, lead time, parsimony, and reproducibility. However, the scoring of the elements and the formula for combining them could easily have been different, and this would have produced different evaluations.

REFERENCES

- Abramowitz, Alan I. 1988. “An Improved Model for Predicting Presidential Outcomes.” *PS: Political Science and Politics* 4: 843–47.
- . 2010. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. New Haven, CT: Yale University Press.
- Abramowitz, Alan I., and Kyle L. Saunders. 1998. “Ideological Realignment in the American Electorate.” *Journal of Politics* 60: 634–52.
- . 2008. “Is Polarization a Myth?” *Journal of Politics* 70: 542–55.
- Campbell, James E. 2000. “The Science of Forecasting Presidential Elections.” In *Before the Vote: Forecasting American National Elections*, ed. James E. Campbell and James C. Garand. Thousand Oaks, CA: Sage.
- . 2006. “Party Systems and Realignment in the United States, 1868–2004.” *Social Science History* 30: 359–6.
- . 2008. *The American Campaign: U.S. Presidential Campaigns and the National Vote*, 2nd ed. College Station: Texas A&M University Press.
- . 2010. “Explaining Politics, Not Polls: Examining Macropartisanship with Recalibrated NES Data,” *Public Opinion Quarterly* 74: 616–42.
- . 2013a. “A First Party-Term Incumbent Survives: The Fundamentals of 2012.” In *Barack Obama and the New America: The 2012 Election and the Changing Face of Politics*, ed. Larry J. Sabato. Lanham, MD: Rowman and Littlefield.
- . 2013b. “The Miserable Presidential Election of 2012: A First Party-Term Incumbent Survives,” *The Forum: A Journal of Applied Research in Contemporary Politics* 10 (4): 20–28.
- Campbell, James E., Bryan J. Dettrey, and Hongxing Yin. 2010. “The Theory of Conditional Retrospective Voting: Does the Presidential Record Matter Less in Open Seat Elections?” *Journal of Politics*, 72: 1083–95.
- Campbell, James E., and Kenneth A. Wink. 1990. “Trial-Heat Forecasts of the Presidential Vote,” *American Politics Quarterly* 18: 251–69.
- Erikson, Robert S., and Christopher Wlezien. 2012. *The Timeline of Presidential Elections How Campaigns Do (and Do Not) Matter*. Chicago: University of Chicago Press.
- Fair, Ray C. 1988. “The Effect of Economic Events on Votes for President: 1984 Update.” *Political Behavior* 10: 168–79.
- Holbrook, Thomas M. 2008. “Incumbency, National Conditions, and the 2008 Presidential Election.” *PS: Political Science and Politics* 41: 709–12.
- Lazarsfeld, Paul F. 1944. “The Election is Over.” *Public Opinion Quarterly* 8: 317–30.
- Lewis-Beck, Michael S. 1988. *Economics and Elections*. Ann Arbor: University of Michigan Press.
- . 2005. “Election Forecasting: Principles and Practice.” *British Journal of Politics and International Relations* 7: 145–64.
- Mayhew, David R. 2008. “Incumbency Advantage in U.S. Presidential Elections: The Historical Record,” *Political Science Quarterly* 123 (2): 201–28.
- Nadeau, Richard, and Michael S. Lewis-Beck. 2001. “National Economic Voting in U.S. Presidential Elections,” *Journal of Politics* 63: 159–81.
- Norpoth, Helmut. 2000. “Of Time and Candidates: A Forecast for 1996.” in *Before the Vote: Forecasting American National Elections*, ed. James E. Campbell and James C. Garand. Thousand Oaks, CA: Sage.
- . 2002. “On a Short-Leash: Term Limits and the Economic Voter.” In *Economic Voting*, ed. Han Dorussen and Michael Taylor. Oxford: Routledge.
- Vavreck, Lynn. 2009. *The Message Matters: The Economy and Presidential Campaigns*. Princeton, NJ: Princeton University Press.
- Weisberg, Herbert F. 2002. “Partisanship and Incumbency in Presidential Elections,” *Political Behavior* 24: 339–60.