

Forecasting presidential elections in the United States

It is widely believed that U.S. presidential elections represent a referendum on the policies and achievements of the incumbent party. Pioneering studies by Kramer (1972) and Fair (1978, 1982, 1988, 1996) found empirical support for this proposition; specifically, both concluded that more rapid growth of real output produced gains for the party of the incumbent president. Following Kramer and Fair, others have investigated empirical links between election outcomes and economic performance using a variety of techniques, specifications, and data sets. These empirical relationships are often referred to as “vote functions.”

If a vote function accurately describes how changes in economic conditions determine changes in political support, one can use knowledge of prevailing conditions to predict outcomes in an upcoming election. Fair’s model has frequently been cited for its use in forecasting. While it is natural to use vote functions in this manner, a useful distinction can be made between most election forecasting models and empirically estimated vote functions. The primary purpose of a vote function is to describe the behavior of voters, i.e., to test alternative theories or hypotheses about voter behavior. A good forecasting model cannot ignore knowledge of voter behavior, but its primary purpose is to use all available information to make the best possible forecast. To make the distinction clear, Fair’s vote function expressly excluded pre-election polls as explanatory variables. In 1992, Fair’s model predicted an easy win by incumbent George Bush, even as pre-election polls made it clear that this outcome was unlikely. Adding a poll as an explanatory variable in Fair’s vote function would presumably have improved its forecast, but would not have provided any fundamental explanation for voters’

dissatisfaction with the incumbent. In the remainder of this entry, attention to the behavioral issues studied by Fair and Kramer is deferred; we instead focus on the production of accurate forecasts.

One might ask why election forecasting, apart from associated behavioral issues, would be of academic interest – accurate election forecasts might be important for candidates and their campaign staff, but societal welfare gains associated with small reductions in election uncertainty are likely to be small. There are several good reasons for investing some effort in forecasting, however. First, forecasting effort is likely to produce spillover benefits related to the more fundamental issues associated with investigating voter behavior; the separation of forecasting issues from behavioral issues is somewhat illusory. For example, Gellman and King (1993) use findings about the forecasting performance of polls to draw inferences about how and when voters make their choices over candidates. Second, the specific forecasting problem presented by the institution of the Electoral College is an unusual one, so the application itself is likely to lead to innovative statistical methods. Finally, the entertainment value associated with forecasting activities cannot be dismissed, given public and media interest focused on the horse-race aspects of the campaign.

Among the simplest forecasting models are several that predict national two-party vote shares using time series data and parsimonious sets of explanatory variables. Campbell and Wink (1990) use just two predictor variables, a trial-heat poll and second quarter GDP growth in the year of the election. Lewis-Beck and Rice (1992) use a similar specification, but add variables capturing recent partisan trends. Both Campbell and Wink and Lewis-Beck and Rice report that their models produce accurate forecasts by

Labor Day, roughly 60 days prior to the election. Specifically, out-of-sample forecast errors for the Campbell-Wink model had a mean absolute error of just 1.3% over the 1948-1992 period (Campbell, 1996). Other contributions in this genre include Abramowitz (1992, 1996), Lewis-Beck and Tien (1996), Holbrook (1996), and Erickson and Wlezien (1996). Notably, Hibb's two-variable "bread and peace" voting model predicts well despite its behavioral focus and the absence of any poll-based measures (Hibbs, 2000).

Although the accuracy of these models is admirable, predicting shares of the popular vote should not be the principal objective when the election winner is selected in the Electoral College. Electoral College votes are cast according to outcomes in the individual states; consequently, good forecasts of the Electoral College should be derived from vote forecasts for the states. Several models, including those developed by Rosenstone (1983), Holbrook (1991) and Campbell (1992) have taken this approach. All of these models examine election outcomes across both states and time, using a mixture of national- and state-level variables as explanatory variables. Campbell lists three limitations of the Rosenstone model: (1) values for some explanatory variables are not actually available prior to the election, (2) measurement of some explanatory variables requires subjective assessments by "experts," and (3) the model includes a very large number of explanatory variables. Holbrook's model avoids these limitations, but fails to achieve the same forecasting accuracy as Rosenstone's.

Campbell's (1992) model of state election outcomes aspires to be comparable to Holbrook in parsimony and operationality, and comparable to Rosenstone in forecasting accuracy. Following the example of his national vote-share model, Campbell includes an

early-September national trial-heat poll and second quarter GDP growth as explanatory variables. The model also includes prior state deviations from national voting outcomes and other state- and region-specific indicators of partisan strength and economic performance. Although it is difficult to explain some shifting partisan alignments over time, Campbell's model ultimately does rival that of Rosenstone's in terms of within-sample forecasting accuracy both for vote shares and Electoral College outcomes. At this is written, Campbell's model remains "state-of-the-art" in forecasting presidential elections at the state level.

Future work on econometric approaches to presidential election forecasting is likely to proceed in several directions. First, in recent years trial-heat polls for individual states have become available with increasing frequency. Since 1988, state-specific polls have been tracked and archived by the *Hotline Weekly* and its successor *National Journal*. Given that polls have demonstrated value for forecasting at short horizons before elections, it seems very likely that state-specific polls will be useful for state-specific forecasts. Holbrook and DeSart (1999) provide early evidence that this is the case. Second, it would be useful to present forecasts in terms of underlying probabilities rather than point estimates (Beck, 2000). In the literature, most assessments of accuracy are reported in terms of regression standard errors or mean absolute values of regression residuals. Given the complexity introduced by the Electoral College and the likelihood that forecast errors across states are not independent, such statistics tell us little about candidate win probabilities. A third issue concerns the timing of forecasts. Campbell's model, as many others, is designed to produce a forecast in early September, utilizing a trial-heat poll at that date. However, there is no reason forecasters should be restricted to

this time horizon. Forecasts could be produced repeatedly at different pre-election horizons, using the best data available at a given time. Chappell and Brown (1999) and Campbell (1996) and describe methods for producing forecasts at differing pre-election time horizons.

While most scholars continue to use of econometric models for forecasting elections, there is an important alternative. Researchers operating the Iowa Electronic Markets have instead used data generated by market outcomes (Forsythe, Nelson, Neuman, and Wright, 1992). In the Iowa electronic markets, traders buy and sell shares whose ultimate value is determined by the election outcome. Depending upon the type of shares being traded, market prices leading up to the election provide a measure of either a candidate's probability of winning or his expected vote share. If traders efficiently absorb available information about candidates' prospects, the market should give an accurate and continuously updated forecast of the election outcome. In practice, Iowa market forecasts have performed well, better than unadjusted point estimates from the major polls.

Further, evidence indicates that market prices are not simply driven by the latest polls. At the very least, the Iowa market poses an interesting challenge to analysts employing econometric forecasting models: if you can produce a better forecast, why not use it to make money in the Iowa?

References

- Abramowitz, A. (1996). Bill and Al's excellent adventure: forecasting the 1996 presidential election. *American Politics Quarterly*, 24: 434-442.
- Abramowitz, A. (1992). An improved model for predicting presidential outcomes. *PS: Political Science and Politics*. 21: 843-847.
- Beck, N. (2000). Evaluating forecasts and forecasting models of the 1996 presidential election. In Campbell, J. and Garand, J., eds., *Before the Vote: Forecasting American National Elections*. Thousand Oaks, CA: Sage Publications: 161-168.
- Brown, L., and Chappell, H. (1999) Forecasting presidential elections using history and polls. *International Journal of Forecasting* 15: 127-135.
- Campbell, J., and Wink, K. Trial-heat forecasts of the presidential vote. *American Politics Quarterly* 18: 251-269.
- Campbell, J. (1996). Polls and votes: the trial heat presidential election forecasting model, certainty, and political campaigns. *American Politics Quarterly* 24: 408-433.
- Erikson, R., and Wlezien, C. (1996). Temporal horizons and presidential election forecasts. *American Politics Quarterly* 24: 492-505.
- Fair, R. (1996). Econometrics and presidential elections. *The Journal of Economic Perspectives* 3: 89-102.
- Fair, R. (1978). The effect of economic events on votes for president. *Review of Economics and Statistics* 60: 159-173.
- Fair, R. (1982). The effect of economic events on votes for president: 1980 results. *Review of Economics and Statistics* 64: 322-325.
- Fair, R. (1988). The effect of economic events on votes for president: 1984 update. *Political Behavior*, 10: 168-179.
- Forsythe, R., Nelson, F., Neumann, G., and Wright, J. (1992). Anatomy of an Experimental Political Stock Market. *American Economic Review* 82:1142-1161.
- Kramer, G. (1971). Short-term fluctuations in U.S. voting behavior. 1896-1964. *American Political Science Review* 65: 131-143.

- Gelman, A., and King, G. (1993). Why are American presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science*, 23: 409-451.
- Hibbs, D. (2000). Bread and peace voting in U.S. presidential elections. *Public Choice* 104: 149-180.
- Holbrook, T. (1996). Reading the political tea leaves: a forecasting model of contemporary presidential elections. *American Politics Quarterly* 24: 434-442.
- Lewis-Beck, M. and Rice, T. (1992). *Forecasting Elections*. Washington, D.C.: Congressional Quarterly Press.
- Lewis-Beck, M., and Tien, C. (1996). The Future in Forecasting: Prospective Presidential Models. *American Politics Quarterly* 24: 468-492.
- Rosenstone, S. (1983) *Forecasting Presidential Elections*. New Haven: Yale University Press.