

Forewarned Before Forecast: Presidential Election Forecasting Models and the 1992 Election¹

Jay P. Greene, *Harvard University*

Forecasting presidential elections has become a growth industry in political science. Econometric models used to predict the presidential election that were once viewed as part of "recreational political science" are now being taken seriously. The widespread belief that these models have been highly successful at predicting election outcomes has enticed a score of political scientists to propose new models hoping to share in the triumph of political science over the pundits and polls. Unfortunately, there has been very little critical examination of the models used to forecast presidential elections. A close review reveals that existing quantitative models are not useful predictors of presidential races. In addition, most of the new proposed models have adopted an approach that is unlikely to lead to better forecasts.

Problems with Precision

Of the many models that generated predictions for the 1992 presidential election, some picked Bill Clinton as the winner. James Campbell's model, for example, predicted that Clinton would win 61.7% of the *electoral college vote* (Campbell 1992).² Clinton actually won with 68.8%. Alan Abramowitz predicted that Clinton would get 53.7% of the two-party popular vote, while the actual figure was 53.4% (Abramowitz 1988). Other prominent modelers did not fare as well. Michael Lewis-Beck and Tom Rice's forecast was for a Bush victory with 57.5% of the *electoral college* (Lewis-Beck/Rice 1992). Ray Fair also picked the wrong winner, predicting that Bush would receive 55.7% of the two-party popular vote (Fair 1990)³ (see Figure 1).

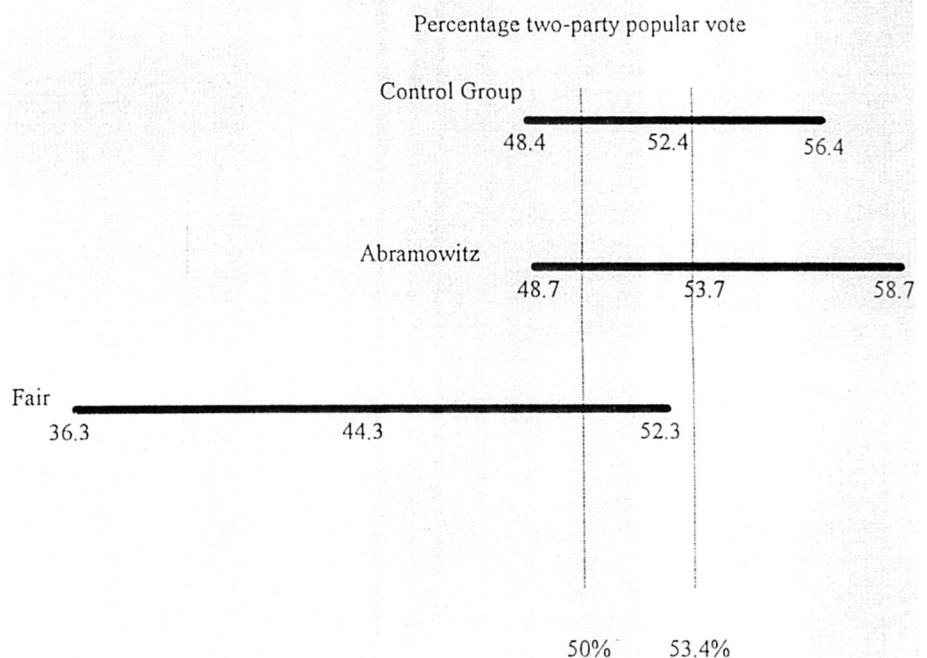
The standard for judging a class of models, however, should not be based solely on whether any model picks the right winner since, with

enough models, picking the winner and even predicting the margin of victory in an election is possible by chance alone. The real question is whether we can rely on the predictions models make. One measure of this reliability is the size of the confidence interval around the prediction of the model. Even political scientists not familiar with quantitative methods have come across confidence intervals. Opinion polls, for example, usually report their best estimate of the percentage of the population that answered in a certain way as well as a confidence interval around that estimate. A poll saying

that 40% answered "yes" with a 95% confidence interval of plus or minus three percent means that if the survey were to be repeated 100 times under the same conditions, we would expect that 95 trials would produce results within three percentage points of 40.

The predictions made by the authors of presidential election forecasting models are their best estimates of the outcome. Just as with opinion polls, it is possible and necessary to compute confidence intervals around these estimates as a measure of our uncertainty. While the modelers do not report con-

FIGURE 1
The Models' Best Estimates and 95% Confidence Intervals



The bars represent 95% confidence intervals for the predictions generated by the models of Alan Abramowitz and Ray Fair as well as by the control group. The number at the center of each bar is the best estimate of the Democratic share of the two-party popular vote for president in 1992. The numbers at the ends of each bar are the extreme predictions that fall within the confidence interval. The vertical line at 50% represents the difference between a Democratic and Republican victory. The vertical line at 53.4% represents the actual Democratic percentage of the two-party vote in 1992.