

# Combined Forecasts of the 2008 Election: The Pollyvote

ANDREAS GRAEFE, J. SCOTT ARMSTRONG, ALFRED G. CUZÁN, AND RANDALL J. JONES, JR.

## PREVIEW

At [PoliticalForecasting.com](#), better known as the Pollyvote, the authors combine forecasts from four sources: election polls, a panel of American political experts, the Iowa Electronic Market, and quantitative models. The day before the election, Polly predicted that the Republican ticket's share of the two-party vote would be 47.0%. The outcome was close at 46.6% (as of the end of November). In his Hot New Research column in this issue, Paul Goodwin discusses the benefits of combining forecasts. The success of the Pollyvote should further enhance interest in this approach to forecasting.

## INTRODUCTION

In this year's presidential election, as in 2004, the Pollyvote applied the evidence-based principle of combining all credible forecasts (Armstrong, 2001) to predict the election outcome. Pollyvote is calculated by averaging within and across four components, all weighted equally, to forecast the incumbent party's share of the two-party vote. The components were updated on a daily basis, or whenever new data became available, and included:

- Combined trial-heat polls (using the RCP poll average from [realclearpolitics.com](#))
- A seven-day rolling average of the vote-share contract prices on the Iowa Electronic Market (IEM)
- 16 quantitative models
- A survey of experts on American politics

## PERFORMANCE OF THE POLLYVOTE

Polly's performance was impressive. From August 2007 through the eve of the election, the Pollyvote

## THE POLLYVOTE TEAM



Andreas Graefe



J. Scott Armstrong



Alfred G. Cuzán



Randall J. Jones, Jr.

consistently predicted that Barack Obama would win the White House – even just following the conventions when combined polls, poll projections (such as [fivethirtyeight.com](#)), and prediction markets indicated at times that John McCain was ahead.

The same was true in 2004, when Polly consistently predicted George Bush as the winner, despite John Kerry's short-term lead in polls and markets. This year's final Polly forecast, issued on the day before the election, missed the actual outcome by 0.4 percentage points. Across the entire forecast horizon, the mean absolute error (MAE) was 1.6 percentage points. By comparison, the corresponding percentage point errors in 2004 were 0.3 and 0.5, respectively.

Comparing the Pollyvote with two other closely followed indicators, Real Clear Politics' average on election eve was off by 0.5 percentage points, and by 1.8 percentage points across the entire forecast horizon. The 'original' IEM (without calculating 7-day rolling averages), was off by 0.2 and 1.7, respectively. The RCP wrongly predicted John McCain as the winner on 41 days, and the IEM did so on 10 days.

Interestingly, the performance of the Pollyvote components was different in 2008, compared with 2004. Ranked in terms of most-to-least-accurate across the entire forecast horizon, the 2004 ranking was the IEM's most accurate, followed by the polls, the experts, and the quantitative models. This year, again over the entire forecasting horizon, the models led in

accuracy, followed by the experts, the IEM and the polls. The finding that the combined Pollyvote forecasts for the two elections were almost equally accurate supports the decision to weight the components equally, rather than differentially.

In a change from the previous presidential election, this year the Pollyvote incorporated *damping* to reduce measurement error in polls. This technique makes forecasts more conservative in situations involving high uncertainty. Applying it in 2008 seemed appropriate, because polls have been found to overestimate support for the front-runner, especially early in the campaign (Campbell, 1996). Campbell provides a damping formula, which we used to discount the polls' spread between the candidates, proportionate to the time

The Pollyvote was designed to demonstrate the power of combining forecasts. Many forecasters overlook the combining principle, even though more than thirty studies have shown that it greatly improves forecast accuracy.

remaining until election day. The longer the time until the election, the larger the discount applied to the front-runner's margin.

Measured over the entire forecast horizon, the MAE for the damped

polls was 2.7 percentage points vs. 1.8 for the original RCP average. The overall Pollyvote MAE increased from 1.3 to 1.6. From this result, which ran contrary to expectations, we conclude that further analysis is necessary to more effectively apply damping in election forecasting.

### THE POWER OF COMBINING

The number of quantitative models utilized in the Pollyvote increased in 2008 to 16 from the 10 used in 2004. Some of the new models brought new methods and data into the mix. For example, Polly added three models that use a segmentation approach by aggregating state-level polls, and two others that employ an index method. One of the latter, the PollyIssues model, represents an innovation: it assumes that voters choose the candidate they believe will better handle the country's problems (Graefe & Armstrong, 2008).

Adding additional models constructed by different methods may have been responsible for the superior performance of the quantitative model component this year. As has been shown by Armstrong (2001), combining forecasts is particularly valuable if you use methods that differ substantially and draw from different sources of information.

The Pollyvote was designed to demonstrate the power of combining forecasts. Combining yields a forecast error which is never larger, and normally is substantially smaller, than the error of the typical forecasts of the components. Still, many forecasters overlook the combining principle, even though more than thirty studies have shown that it greatly improves forecast accuracy. A large part of the problem could be that combining defies intuition. As demonstrated by Larrick and Soll (2006) in a clever series of experiments, a majority of highly intelligent people did not understand the value of combining. As a result, combining is not used nearly as much as it should be in forecasting. People simply think that they can forecast better on their own.

### REFERENCES

- Armstrong, J.S. (2001). Combining forecasts. In: J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Norwell, MA: Kluwer Academic Publishers, 417-439.
- Campbell, J.E. (1996). Polls and votes: The trial-heat presidential election forecasting model, certainty, and political campaigns, *American Politics Quarterly*, 24, 408-433.
- Graefe, A. & Armstrong, J.S. (2008). *Forecasting elections from voters' perceptions of candidates' ability to handle issues*, Available at [http://www.forecastingprinciples.com/PollyVote/images/articles/index\\_us.pdf](http://www.forecastingprinciples.com/PollyVote/images/articles/index_us.pdf)
- Larrick, R.P. & Soll, J.B. (2006). Intuitions about combining opinions: Misappreciation of the averaging principle, *Management Science*, 52, 111-127.

**CONTACT**  
Andreas Graefe  
graefe@itas.fzk.de