

The PollyVote Forecast for the 2016 American Presidential Election

Andreas Graefe, *Columbia University*

Randall J. Jones, Jr., *University of Central Oklahoma*

J. Scott Armstrong, *Wharton School of the University of Pennsylvania*

Alfred G. Cuzán, *University of West Florida*

INTRODUCTION

The PollyVote is an evidence-based formula designed to forecast election outcomes, using both well-established methods and innovations. Forecasting error is reduced by combining forecasts within and across different methods, equally weighted. Following this rule, the PollyVote has accurately forecast the outcome of the last three presidential elections by as much as a year in advance of Election Day. Updated twice a week in 2004 (Cuzán, Armstrong, and Jones 2005a; 2005b) and in subsequent elections at least once daily, at no time has the PollyVote called the election for any other than the winner. Moreover, on average across the past six elections, the PollyVote forecast has been more accurate than any of its component methods (Graefe et al. 2014a; 2014b).

In the sections that follow, we successively elucidate the combination principle used in the PollyVote, summarize the methods incorporated into it, review its performance in forecasting presidential elections, issue a forecast for 2016, and conclude with remarks on the nature of the PollyVote.

COMBINING FORECASTS

As mentioned, the PollyVote technique combines forecasts from different methods. One reason for combining is that it is difficult to determine *a priori* which individual method will provide the best forecasts for a given election or for a given day prior to an election (Armstrong 2001). Every election is held in a different context and has its idiosyncrasies. As a result, a method that worked well in the past might not work in future elections. A method that worked well 100 days before an election might not work as well even 30 days before the same election. Combining helps resolve the problem of method selection by incorporating many methods. Each method usually relies on different information than other methods, so combining includes more information than any individual method. Both systematic and random errors of individual forecasts tend to cancel out in the aggregate, particularly when the individual forecasts draw on different information, bracket the true value being predicted, and are uncorrelated (Graefe et al. 2014b). Not surprisingly, combining forecasts from different methods that rely on different information has become a well-established method of reducing forecast error. This is one of the major

findings in forecasting research over the past half century (Armstrong, Green, and Graefe 2015).

EQUAL WEIGHTS

Research has shown that, apart from being easy to understand, a combining procedure that uses equally weighted components often outperforms more complex approaches which aim to estimate “optimal” weights for the components. These findings apply both to weighting the predictor variables in linear models and to combining forecasts from different models (Cuzán and Bundrick 2009; Graefe 2015b, Graefe et al. 2015). One reason for the strong performance of equal weights is that the accuracy of individual component forecasts varies over time and may be impacted by exogenous effects. Another more technical reason is error in estimating the weights. In general, the simple average will be more accurate than estimated “optimal” weights if two conditions are met: first, the combination is based on a large number of individual forecasts and, second, the optimal weights are close to equality. In such situations, each forecast has a small weight, and the simple average provides an efficient tradeoff against the error that arises from the estimation of weights (Graefe et al. 2015b).

THE POLLYVOTE FORMULA

The PollyVote combines forecasts within and across six different component methods in the following manner. A combined forecast for each component method is calculated. Then equal weights are applied to compute the simple average across the component forecasts. For example, we average the point forecasts of the various regression models. The forecast of each model is weighted the same; thus equal weights. That average becomes the forecast for the models’ component. We also average the results of recent polls, which becomes the trial heat polls component forecast. Similarly, the forecasts within the remaining component categories are averaged. The results of such averaging “within” all six components are, in turn, then averaged. This averaging “across” the six components is the PollyVote forecast.

When using equal weights, this does not mean that every forecast used in calculating the PollyVote is weighted equally with every other forecast. There is only one prediction market that predicts the national popular vote, so there is one forecast available for that component method. On the other hand,

there are several regression models, each producing its own forecast. If we used a simple average of all available forecasts, that would over-represent models and grossly underrepresent prediction markets. In short, we use equal weights when averaging forecasts within each component category, and we use equal weights when averaging the six combined forecasts. Thus we seek to equalize the impact of each constituent method rather than each individual forecast.

THE COMPONENTS OF THE POLLYVOTE

The 2016 PollyVote is derived by averaging forecasts within and across six different component methods:

1. Trial heat polls
2. Prediction markets
3. Regression models
4. Expert judgment
5. Index models
6. Citizen forecasts

The first four methods listed above comprised the original specification of the PollyVote used in 2004 (Cuzán, Armstrong,

“Who do you think will win the US presidential election, Donald Trump or Hillary Clinton?” The aggregate responses are then used to predict the election winner.

Though often overlooked, these citizen forecasts are highly accurate predictors of election outcomes (Graefe 2014). In 89% of 217 surveys administered between 1932 and 2012, a majority of respondents correctly predicted the winner. Regressing the incumbent share of the two-party vote on the percent of respondents who expect the incumbent party ticket to win accounts for two-thirds of the variance. Moreover, in the last 100 days of the previous seven presidential elections, vote expectations provided more accurate forecasts than vote intention polls, prediction markets, econometric models, and expert judgment. Compared to a typical poll, for example, vote expectations reduced the forecast error on average by about 50%. Other tests also have found this method to be successful in increasing forecast accuracy (Graefe 2015a).

In deriving a forecast using this component, we translate the results of vote expectation surveys into a two-party vote share prediction using the vote equation estimated by Graefe (2014). Then, we calculate the combined component forecast

Moreover, in the last 100 days of the previous seven presidential elections, vote expectations provided more accurate forecasts than vote intention polls, prediction markets, econometric models, and expert judgment.

and Jones 2005a; 2005b) and 2008 (Graefe et al. 2009). In 2012, index models were added to the formula (Graefe et al. 2014a), and citizen forecasts have been included for the first time in this year's forecast. Each component method has been shown in research findings to be an appropriate and accurate election predictor.

Polls—Vote Intention Surveys

Vote intention surveys—trial heat polls—are most prevalent and highly visible in the news media coverage. The method asks respondents a variation of this question: *“If the election for President were held today, for whom would you vote: Donald Trump, the Republican, or Hillary Clinton, the Democrat?”* The PollyVote relies on several poll aggregators, each of which collects and aggregates results of individual polls. In order to calculate its combined poll component, the PollyVote averages the forecasts across the different poll aggregators. On September 7, PollyVote's combined poll component predicted the Democratic nominee Hillary Clinton would receive 52.1% of the two-party popular vote.

Citizen Forecasts—Vote Expectation Surveys

Vote expectation surveys—or citizen forecasts—are the newest addition to the PollyVote. Vote expectation surveys ask respondents who they expect to win the election, rather than asking people for whom they themselves intend to vote (Lewis-Beck and Skalaban 1989). A typical question might be:

by exponential smoothing. On September 7, PollyVote's citizen forecast component predicted Hillary Clinton to win 51.9% of the two-party vote.

Prediction Markets

Prediction markets are another expression of people's expectations of who will win. Yet, instead of asking a representative sample for their opinion, prediction markets are open for anyone to participate. And, people reveal their opinion by betting money on the election outcome. The resulting betting odds can then be interpreted as the market forecast. Most available markets provide probability forecasts for the candidates' likelihood to win. However, the PollyVote requires a national popular vote share prediction. We know of only one prediction market, the University of Iowa's Iowa Electronic Market (IEM), which provides such information. Graefe (2016) reviewed prediction market accuracy for providing vote-share forecasts for elections in different countries. He found that prediction markets tend to outperform forecasts made by experts, as well as forecasts based on quantitative models and trial-heat polls, although compared to citizen forecasts the evidence was mixed. The PollyVote uses the IEM's daily market prices but calculates one-week rolling averages to diminish short-term fluctuations. On September 7, the IEM one-week rolling average predicted Clinton to gain 52.7% of the two-party vote.

Expert Judgment

The PollyVote includes the judgment of prominent academics (and in 2004 some practitioners, as well) knowledgeable of American politics. Experts have been shown to be more accurate than polls or the IEM early in the election season, when the election is still nine months to a year or more in the future (Jones and Cuzán 2013). In 2016, a panel of 15 experts formed by the PollyVote team has been polled monthly, and the mean forecast is incorporated into the PollyVote. On September 7, the panel of experts expected Clinton to garner 53.5% of the two-party vote.

Index Methods

Benjamin Franklin may have been the original inventor of what he called “moral algebra” (Franklin 1956). He advised that before making a decision to do or not to do something of importance, one should list all possible considerations involved in the decision, assign a positive or negative value to each, weight them according to their relative importance, and sum them. If the result was positive, one was to proceed with the decision, but desist if the total was negative. Franklin’s technique was a form of index or checklist. Indexes are a

As of September 7, 62 days before the election, the PollyVote predicts a narrow win in the two-party vote for Democrat Hillary Clinton, 52.4% vs. 47.6% for Donald Trump.

Regression Models

For the past several presidential election cycles at least a dozen political scientists and economists have computed regression equations to forecast the election results (Campbell 2013; Jones 2008). Many of the models use economic data through the second quarter of the election year, the first official estimate of which becomes available in late July. Forecasts from those models are made shortly after that. There are exceptions, however. The predictions of some models are available well before then, even two years ahead of the election, while at least one is delayed until the first polls after Labor Day are released (Campbell 2013). As these forecasts become available, they are averaged into a combined regression model component and incorporated into the PollyVote. On September 7, this component forecast Clinton to receive 49.6% of the two-party vote, which makes it the only component to predict a Trump victory.

component of the PollyVote and are appropriate in situations in which (a) a large number of variables are important and (b) there is good prior knowledge about the directional effect of each variable on the phenomenon of interest (Armstrong and Cuzán 2006; Graefe 2015b).

In the context of election forecasting, indexes are typically constructed based on ratings of specific characteristics of candidates or events. Ratings can be made by experts or members of the public (e.g., based on survey data) and typically cover factors such as the candidates’ biographic information, leadership skills, or issue-handling competences (Graefe 2013; Armstrong and Graefe 2011), as well as exogenous effects, such as economic performance or the presence of a third party (Lichtman 2005). Point forecasts of an election are provided by inserting current data into an equation specified by regressing the vote on the respective index scores.

Currently forecasts from five index models are averaged to calculate the PollyVote’s combined index model component. On September 7, this component predicted Clinton to achieve 53.7% of the popular two-party vote, which, among our component forecasts, is the most optimistic for Clinton.

THE POLLYVOTE TRACK RECORD AND 2016 FORECAST

As shown in table 1, the PollyVote has been highly accurate in forecasting the last six presidential elections, three times retrospectively and three times

Table 1

PollyVote Forecast Accuracy, 1992–2012: Mean Absolute Error (MAE) at 100, 60, and 30 days before the election

Year	Incumbent Vote ^a	PollyVote Forecast by Horizon Days to Election Day			MAE ^b
		100	60	30	
1992 ^c	46.5	54.0	49.6	47.9	3.9
1996 ^c	54.7	56.9	55.6	56.2	1.7
2000 ^c	50.3	50.7	52.4	52.4	1.5
2004 ^d	51.2	52.8	51.5	52.0	0.9
2008 ^d	46.3	48.6	48.1	47.6	1.8
2012 ^d	51.9	51.7	51.6	51.8	0.02
MAE at 100, 60, or 30 days to election, all years.		2.4	1.4	1.2	1.65
MAE, at 100, 60, or 30 days to election, last three years.		1.4	0.8	0.7	0.91

Notes:

^aSource: Dave Leip’s “Atlas of U.S. Presidential Elections” (uselectionatlas.org).

^bMAE across the three horizons, by year.

^cPost-dictions, i.e., these “forecasts” were calculated retrospectively, with only three components, the only ones available then: polls, the IEM, and econometric models.

^dPredictions, i.e., these were forecasts in real time with four components, the previously mentioned three plus a panel of experts.

prospectively, in real time. On average, across three horizons in the election cycle (100, 60, and 30 days before the election), the Mean Absolute Error (MAE) of the forecasts for all six elections is less than 2 percentage points. In the last three elections, in which true *ex ante* forecasts were made, the MAE is even lower, less than 1 percentage point.

As of September 7, 62 days before the election, the PollyVote predicts a victory in the two-party vote for Democrat Hillary Clinton, 52.4% vs. 47.6% for Donald Trump. Needless to say, this prediction may change as new information becomes available. That said, the PollyVote predictions have been remarkably stable over the course of the past three elections.

CONCLUSION

The PollyVote draws information generated by others with different methods and averages all forecasts within and across methods to make a final forecast. With the exception of the expert panel's forecast and the index models, the PollyVote does not add any new information to the mix. Neither does it compete with any individual model. Rather, it aggregates all relevant information generated by all models and methods in a pre-specified and simple manner.

The PollyVote is educational and was founded as a means for comparing forecasting methods in real time. At PollyVote.com, the visitor will find not only the current day's forecast, but also all data used to generate the PollyVote forecasts for all presidential elections since 1992, as well as links to previous papers and articles associated with this work. ■

REFERENCES

- Armstrong, J. Scott. 2001. "Combining Forecasts." In *Principles of Forecasting: A Handbook for Researchers and Practitioners*. J. Scott Armstrong, Editor. New York: Springer, 417–39.
- Armstrong, J. Scott and Alfred G. Cuzán. 2006. "Index Methods for Forecasting: An Application to the American Presidential Elections." *Foresight: International Journal of Applied Forecasting* 3: 10–13.
- Armstrong, J. Scott and Andreas Graefe. 2011. "Predicting Elections from Biographical Information about Candidates: A Test of the Index Method." *Journal of Business Research* 64 (7): 699–706.
- Armstrong, J. Scott, Kesten C. Green, and Andreas Graefe. 2015. "Golden Rule of Forecasting: Be Conservative." *Journal of Business Research* 68 (8): 1717–31.
- Campbell, James E. 2013. "Recap: Forecasting the 2012 Election," *PS: Political Science & Politics* 46 (1): 37.
- Cuzán, Alfred G., J. Scott Armstrong, and Randall J. Jones Jr. 2005a. "How we Computed the PollyVote." *Foresight: The International Journal of Applied Forecasting* 1 (1): 51–2.
- . 2005b. "The PollyVote: Applying the Combination Principle in Forecasting to the 2004 Presidential Election." Paper presented at the 2005 International Symposium on Forecasting, San Antonio.
- Cuzán, Alfred G. and Charles M. Bundrick. 2009. "Predicting Presidential Elections with Equally Weighted Regressors in Fair's Equation and the Fiscal Model." *Political Analysis* 17 (3): 333–40.
- Franklin, Benjamin. 1956. "Benjamin Franklin's 1772 letter to Joseph Priestley." Available: <http://www.procon.org/view/background-resource.php?resourceID=1474>.
- Graefe, Andreas. 2013. "Issue and Leader Voting in US Presidential Elections." *Electoral Studies* 32 (4): 644–57.
- . 2014. "Accuracy of Vote Expectation Surveys in Forecasting Elections." *Public Opinion Quarterly* 78 (S1): 204–32.
- . 2015a. "Accuracy Gains of Adding Vote Expectation Surveys to a Combined Forecast of US Presidential Election Outcomes." *Research & Politics* 2 (1): 1–5.
- . 2015b. "Improving Forecasts Using Equally Weighted Predictors." *Journal of Business Research* 68 (8): 1792–99.
- . 2016. "Political Markets." Forthcoming (subject to changes) in the *SAGE Handbook of Electoral Behavior*. Available: https://www.researchgate.net/publication/292615991_Political_markets.
- Graefe, Andreas, J. Scott Armstrong, Randall J. Jones Jr., and Alfred G. Cuzán. 2014a. "Combining Forecasts: An Application to Elections." *International Journal of Forecasting* 30 (1): 43–54.
- . 2014b. "Accuracy of Combined Forecasts for the 2012 Presidential Elections: The PollyVote." *PS: Political Science & Politics* 47 (2): 427–31.
- . 2009. "Combined Forecasts of the 2008 Election: The PollyVote." *Foresight: The International Journal of Applied Forecasting* 12: 41–42.
- Graefe, Andreas, Helmut Küchenhoff, Veronika Stierle, and Bernard Riedl. 2015. "Limitations of Ensemble Bayesian Model Averaging for Forecasting Social Science Problems." *International Journal of Forecasting* 31 (3): 943–951.
- Jones, Randall J., Jr. 2008. "The State Of Presidential Election Forecasting: The 2004 Experience." *International Journal of Forecasting* 24 (2): 310–21.
- Jones, Randall J., Jr. and Alfred G. Cuzán. 2013. "Expert Judgment in Forecasting American Presidential Elections: A Preliminary Evaluation." Presented at the 2013 meeting of the American Political Science Association, Chicago.
- Lewis-Beck, Michael S. and Andrew Skalaban. 1989. "Citizen Forecasting: Can Voters See into the Future?" *British Journal of Political Science* 19 (1): 146–153.
- Lichtman, Alan J. 2005. "The Keys To The White House: An Index Forecast For 2008." *International Journal of Forecasting* 24 (2): 301–09.