

Combining Forecasts for U.S. Presidential Elections: The PollyVote

Andreas Graefe

Alfred G. Cuzà

Randall J. Jones, Jr.

J. Scott Armstrong

10/30/09

Abstract. In the PollyVote, we evaluated the combination principle to forecast the five U.S. presidential elections between 1992 and 2008. We combined forecasts from three or four different component methods: trial heat polls, the Iowa Electronic Markets (IEM), quantitative models and, in the 2004 and 2008 contests, periodic surveys of experts on American politics. The forecasts were combined within as well as across components. On average, combining within components reduced forecast error -- and increased predictive accuracy -- by 17% to 40%. Combining across components led to additional error reductions ranging from 7% to 68%, depending on the forecast horizon. In addition, across all five elections, the PollyVote predicted the correct election winner on all but 4 out of 957 days. The gains from applying the combination principle to election forecasting were much larger than those obtained in other fields.

Introduction

In both the 2004 and 2008 U.S. presidential campaigns we operated a web site (www.pollyvote.com) that made frequent forecasts of the election outcome. These forecasts were generated by combining forecasts within each of four methods and then combining the forecasts across these four methods. The forecast methods included campaign (trial heat) polls, vote-share contract prices from the Iowa Electronic Market, surveys of experts who were mostly political scientists specializing in American politics (but not election forecasters), and several quantitative models, most well-known for forecasting presidential elections. To further test the value of this approach beyond our experience in 2004 and 2008, we made retrospective forecasts for the 1992, 1996, and the 2000 elections. In the following sections we describe 1) the theory underlying the technique of combining forecasts, 2) previous applications of this technique in other contexts, 3) the design and implementation of our approach to combining in this application to elections, and 4) the gains in forecast accuracy that resulted from combining.

The combination principle in forecasting

Combining forecasts can reduce error in several ways. A combined forecast is likely to be more accurate than a typical forecast of an individual component, because biases associated with the data and methods used in various forecasts are likely to differ. Hence, their forecast errors are likely to be uncorrelated and perhaps offsetting. In addition, combined forecasts make use of more information than any one component forecast. More information provides a more complete picture of influences affecting the future. In probability terms, because the “sample” of information underlying a combined forecast is larger than that of a single forecast, it is likely that the combined forecast is more accurate than that coming from any single included source. Mathematically, the combined forecast will always be at least as accurate as the typical individual forecast.

These expectations are supported by empirical evidence. In a meta-analysis of 30 studies, Armstrong (2001) showed that combining forecasts reduced forecast error by about 12% when compared with the typical error of the components. In addition, combining forecasts never reduced forecast accuracy, and substantially lowered the risk of large

forecast errors. Often, the combined forecast was more accurate than forecasts from the best individual method. Many of these studies were based on combining only two methods, and most of the combinations were derived from similar methods (such as judgmental forecasts). With every additional forecast, the accuracy of the combined forecast normally improves, although at a slower rate. Based on this, Armstrong (2001) recommended using as many as five methods, and combining forecasts mechanically, according to a predetermined procedure. Furthermore, the findings suggest that forecasts should be weighted equally, unless there is strong prior evidence that supports differential weights.

Under ideal conditions, the gains from combining can be expected substantially to exceed the 12% error reduction reported in Armstrong's meta-analysis. In addition, gains are likely to be greater when forecast uncertainty is high. Thus, combining is especially useful for long forecast horizons.

Although the combination principle has been shown to reduce error, its application to forecasting and decision-making is not widespread. Larrick and Soll (2006) concluded from a series of experiments that combining is in limited use because “[p]eople lack the intuition for it, and rarely have an opportunity to learn it” (2006:111). In this paper, we demonstrate the value of combining when applied to election forecasting, perhaps thereby encouraging wider use of this simple but powerful technique.

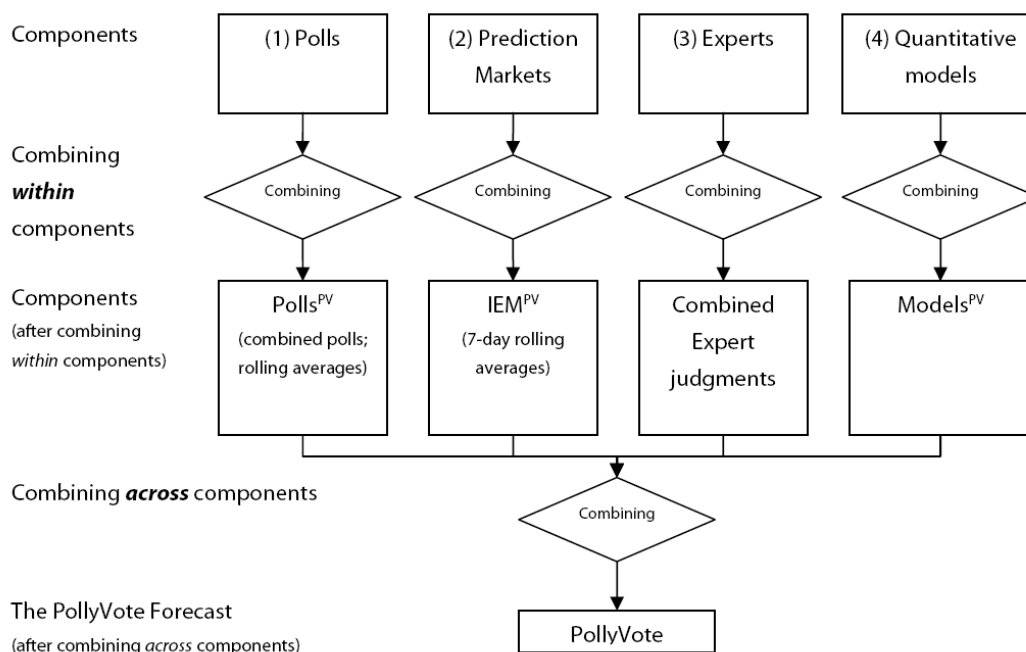
Construction of the PollyVote

Our presidential election forecasts predict the percent of the two-party vote garnered by the incumbent party candidate. We call these forecasts the PollyVote – “pol” for political and “poly” for many methods. The process of combining forecasts is simple—we average within and across the component forecasts, weighing them all equally. We used this technique in generating our *ex ante* forecasts for the 2004 and 2008 elections, as reported in real time at www.pollyvote.com. This also was the technique used in creating our retrospective forecasts for 1992, 1996, and 2000.

Again, as shown in Figure 1, the PollyVote takes information from four different component forecasting methods: (1) polls, (2) prediction markets, (3) surveys of

American politics experts, and (4) quantitative models. To calculate the PollyVote, forecasts are combined in a two-stage process. First, forecasts are combined within each method. For example, in 2008 we used the average of recent polls as reported at RealClearPolitics.com (RCP). Second, these averaged forecasts from each of the four methods are in turn averaged. The latter process of combining across the component methods produces the PollyVote. Thus, the PollyVote's final 2008 forecast of November 3 was the average of the following: (1) the RCP average of polls, 46.1% for McCain; the median of the experts' predictions in the final survey, 47.5%; the average price of the last week's Iowa Market vote-share contract, 46.7%; and the average forecast of the regression models, 47.1%. This average of averages was 46.8% for McCain, the PollyVote forecast. (McCain received 46.3% of the 2-party vote, so the PollyVote forecast error was 0.5%.)

Figure 1: Construction of the PollyVote for Forecasting the Incumbent Share of the Two-Party Vote --



Combining within components

Campaign polls

Campaign – or trial heat – polls reveal voter support for candidates in an election. Although survey results are not predictions – only assessments of current opinion or “snapshots” – consumers of polls routinely interpret them as forecasts and project the results to Election Day. However, the information generated by any single poll is unreliable for forecasting the outcome of the election, especially early in the campaign. In both 2004 and 2008, polls conducted by reputable survey organizations at about the same time revealed considerable variation in results.

One way to mitigate the problem of the variation among polls is to combine polls that have been published at about the same time by calculating their mean or median. For example, Gott and Colley (2008) identified the median poll taken in the previous 30 days to predict the 2004 electoral vote. They correctly forecasted the winner of every state but one. This successful performance was repeated in 2008, when they again missed only one state.

In the PollyVote, polls are combined by calculating rolling averages of the incumbent party’s share of the two-party vote in recently published polls. In 2004, as well as retrospectively for the three elections from 1992 to 2000, we calculated the rolling averages of the three most recent polls. (In case more than three polls were published per day, all polls of that day were averaged. That way, we obtained one poll-based forecast per day.) In 2008, we relied on the daily updated RCP poll average from RealClearPolitics.com. The poll-based component forecast is hereafter referred to as Polls^{PV}.

Table 1: Incumbent Share of Two-Party Vote: Last RCP Poll Average Prior to Election Day, 2008

Poll	November	McCain two party	Absolute error
Marist	3	45.3	1.1
Battleground	2 - 3	48.2	1.2
Rasmussen Reports	1 - 3	46.9	0.6
Reuters/C-SPAN/Zogby	1 - 3	44.3	2.0
IBD/TIPP	1 - 3	45.8	0.5
FOX News	1 - 2	46.2	0.1
NBC News/Wall St. Jrnl	1 - 2	45.7	0.6
Gallup	10/31 - 2	44.4	1.9
Diageo/Hotline	10/31 - 2	47.4	1.0
CBS News	10/31 - 2	45.2	1.2
ABC News/Wash Post	10/30 - 2	45.4	1.0
Ipsos/McClatchy	10/30 - 2	46.5	0.1
CNN/Opinion Research	10/30 - 1	46.5	0.1
Pew Research	10/29 - 1	46.9	0.6
Error of typical individual poll (MAE)			0.90
RCP poll average	10/29 - 11/03	46.07	0.25
Error reduction by RCP compared to typical poll			72%

The value of combining polls to reduce their forecast error can be illustrated by again referring to the last RCP poll average prior to the 2008 election (calculated as the mean of the 14 individual polls conducted between October 29 and November 3, 2008). As evident in Table 1, even though all of these polls were conducted shortly before Election Day, their individual errors varied substantially. In fact, the range in error was from 0.1 to 2.0 percentage points. The mean absolute error (MAE) across these 14 polls can be considered the error of a typical individual poll, which was 0.9%. In other words, if one had randomly relied on an individual poll, the forecast error on average would have been 0.9%. By comparison, the RCP average of the 14 poll results missed the election outcome by only 0.25%. Thus, as expected, the forecast of the combined polls was more accurate than the forecast of the typical individual poll. Compared to the typical poll, the reduction in forecast error from combining polls was 72% ($[0.9 - 0.25] / 0.9 * 100$).

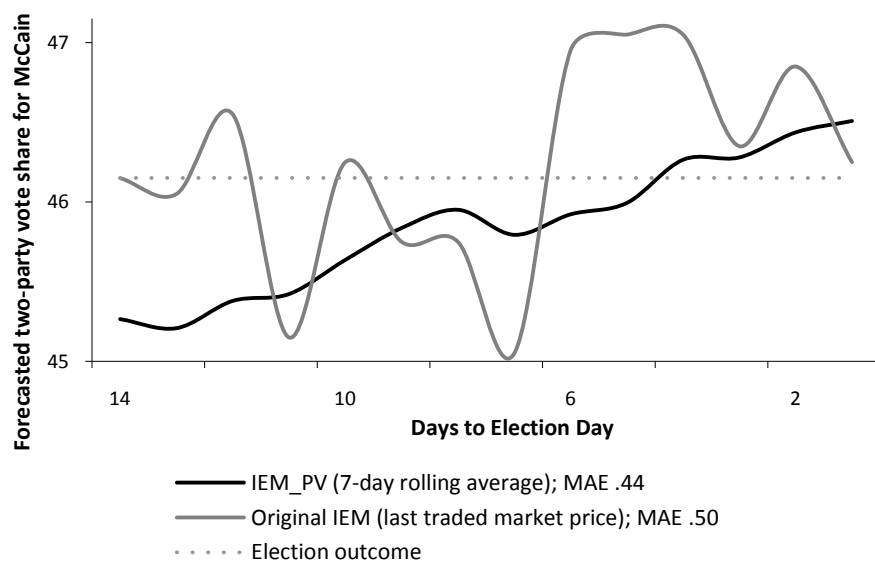
Prediction markets

Betting markets to predict election outcomes are not new and have an interesting history. Rhode and Strumpf (2004) studied betting markets that existed for the 15 presidential

elections from 1884 through 1940 and found that these markets “did a remarkable job forecasting elections in an era before scientific polling” (2004:127). At times trading activity in the betting markets was higher than in the stock exchanges on Wall Street. During some election campaigns, newspapers such as the *New York Times* reported market prices almost on a daily basis. Nonetheless, with increasing availability of other forms of gambling and the rise of opinion polls, presidential betting markets disappeared after 1940.

Betting markets for elections reappeared in 1988 when the Iowa Electronic Market (IEM) was launched as a futures market in which contracts were traded on the outcome of the presidential election that year. Initially, the IEM, commonly viewed as a "prediction market," provided more accurate election forecasts than traditional opinion polls. In analyzing 964 polls for the five presidential elections from 1988 to 2004, Berg et al. (2008) found that IEM market forecasts were closer to the actual election results 74% of the time. However, this advantage seems to disappear when comparing the market forecasts to damped polls. In analyzing data from the same elections, Erikson and Wlezien (2008) found that polls which had been combined and damped were more accurate than both the winner-take-all and the vote-share IEM markets.

Figure 2: Original IEM and IEM^{PV} Forecasts of the Incumbent’s Share of the Two-Party Vote (14 days prior to Election Day, 2008)



The IEM component of the PollyVote is based on average daily trading prices of the vote-share contract for the incumbent party candidate. Forecasts for 1992 through 2004 were generated by calculating 7-day rolling averages of the average daily trading price. Forecasts for 2008 were based on the price of the last trade of the day, also calculated as 7-day rolling averages. We refer to this combination of original IEM forecasts as IEM^{PV}. Calculating the IEM^{PV} forecasts as rolling averages was expected to moderate overreactions of the market due to information cascades, which can cause unexpected positive or negative spikes in prices. Figure 2 illustrates this for the last two weeks prior to Election Day in 2008. The mean absolute error of the original IEM forecasts was 0.5 percentage points. By comparison, IEM^{PV} was smoother, leading to a MAE of 0.44 percentage points – or an error reduction of 14%. Note that this method of combining, unlike that used for the other components, does not ensure that the combined forecast will be better than the latest trade.

Expert forecasts

In 2004 and 2008 we formed a panel of experts in American politics, mostly academics, and contacted them periodically for their estimates of the incumbent's share of the two-party vote on Election Day. We deliberately excluded experts who had made forecasts from regression models, because that method is represented as a separate component in the PollyVote. In 2004 and for initial surveys for the 2008 election, the expert survey was conducted using the Delphi method. Accordingly, we asked participants to provide reasons for their forecasts and then conducted a second round of the survey (Jones, Armstrong, and Cuzán 2007). Later, after finding that forecasts were rarely adjusted as a result of the second round, we simplified the process, which became a one-round expert survey without feedback. For the retrospective analyses of the three elections from 1992 to 2004, no expert forecasts were available.

The results of the surveys are provided in Appendix 2. (Appendices can be accessed online at <http://tinyurl.com/pollyvoteappendix>. Also, the full dataset is available online at <http://tinyurl.com/pollyvotedata>.) Interestingly, in neither election year did the median expert prediction of the incumbent share in the two-party vote vary more than 1.0% from one survey to the next. Indeed, this was the most stable component of the

PollyVote, moderating the turbulence in the polls.

Quantitative models

Over the last several election cycles, economists and political scientists have used regression models to estimate the impact of certain variables on the outcome of U.S. presidential elections. Then, they would often use these models to provide a forecast of the two-party vote received by the incumbent party candidate in the next election. As summarized by Jones and Cuzán (2008), most models include between two and five variables. They typically have an indicator of economic conditions and a measure of public opinion.

The track record of quantitative models is mixed. Among the best-known and historically better performing models are those by Abramowitz, Campbell, Erikson and Wlezien, and Fair (Jones and Cuzán 2008). The structure of these models has remained relatively unchanged over time. Most other models, however, have undergone significant revision since their first appearance, which has usually occurred after a forecast has been wide of the mark.

In 2008, the PollyVote incorporated forecasts from 16 quantitative models. Ten models were used in 2004. For our retrospective analyses of the elections in 1992, 1996, and 2000, we included the forecasts of 4, 8, and 9 models, respectively. Forecasts for most models were available in July and August, and several were updated in mid- or late-October, as revised data became available. These forecasts are reported in Appendix 1. Not surprisingly, in four of the five elections the average of the models' forecasts resulted in lower error than the typical individual model forecast.

Accuracy gains from combining *within* components

Table 2 summarizes the reduction in forecast error that was achieved across five presidential elections, 1992-2008, by combining forecasts within each component method of the PollyVote. Error reduction for each component was calculated as follows:

- *Polls^{PV}*: The error of the last combined poll forecast prior to Election Day was compared to the error of the typical individual poll for that period.
- *IEM^{PV}*: The error of the actual IEM prices for each of the last 7 days prior to

Election Day was compared to the error of the 7-day rolling averages of the IEM (i.e. the IEMPV) for the same period.

- *Experts*: The error of the median expert forecast from the last survey prior to Election Day was compared to the error of the typical individual expert forecast.
- *Models^{PV}*: The error of the last combined forecast from regression models prior to Election Day was compared to the error of the typical individual model's last forecast.

Across all five elections, combining forecasts reduced prediction error by an average 17% for experts, 25% for quantitative models, 35% for polls and 40% for the IEM.

Table 2: Percent Reduction in Error of Combined Forecasts vs. Typical Forecasts

	Retrospective			Ex ante		Mean
	1992	1996	2000	2004	2008	
Polls ^{PV}	0	26	33	46	72	35
IEM ^{PV}	78	22	-8	42	65	40
Models ^{PV}	5	40	0	13	58	23
Experts	-	-	-	39	-6	17

Accuracy gains from combining across components

In addition to the error reduction that resulted from combining within forecasting methods, further gains in accuracy can likely be achieved by combining forecasts from different methods that draw on different data. Combining across methods has been shown to be particularly successful in improving forecast accuracy where forecast errors from different methods are negatively correlated or uncorrelated. If forecast errors are positively correlated, combining is still useful when correlation coefficients are low. For the last 93 days prior to Election Day, in most cases the errors of the component forecasts were only moderately correlated (cf. Table 3), so combining across methods did increase forecast accuracy. Also, note that there is no pattern of correlations between components across years. For the elections in 2004 and 2008, data from all four components was available for 93 days prior to Election Day. The retrospective analysis from 1992 to 2000 did not include expert forecasts.

Although the correlations were not strong in either year, in 2008 the most correlated errors were those of the IEM and the polls, and in 2004 those of the models and the polls. In both years, the experts' errors were uncorrelated with errors of the polls, as they were in 2008 with those of the IEM quotes. This led us to speculate about experts' seeming immunity from disturbances. The fact that they were surveyed on a quarterly or monthly basis no doubt allows for this more serene perspective.

Table 3: Correlations of Errors, 93 Days Prior to Election Day

Election year	Polls^{PV} & IEM^{PV}	Polls^{PV} & Models^{PV}	Polls^{PV} & Experts	IEM^{PV} & Models^{PV}	IEM^{PV} & Experts	Experts & Models^{PV}
2008	.60	.22	.05	.39	.01	.22
2004	.06	.61	-.01	.43	.36	.68
2000	.19	.11	-	.61	-	-
1996	.35	.04	-	.49	-	-
1992	-0.5	-.42	-	.79	-	-

Furthermore, we found that in each of the four elections, the combined forecast of the quantitative models consistently over-predicted the incumbent's actual vote share. Only on one out of 465 days (93 days per election) did the Models^{PV} forecast under-predict the incumbent's vote share, whereas the remaining three components under- or over-predicted the vote shares about 50% of the time. Given that the Models^{PV} consistently over-predicted, bracketing occurred whenever one of the remaining components under-predicted. In these cases the PollyVote was more accurate than the typical forecast drawn randomly from each component. For example, on the last day prior to the election, the Models^{PV} forecast over-predicted the outcome in each of the five election years. By comparison, Polls^{PV} under-predicted four times and IEM^{PV} three times. For the two elections for which we obtained expert forecasts, the expert forecasts once over-predicted and once under-predicted the election outcome.

Table 4 shows the mean absolute error (MAE) for the PollyVote forecast and for each combined component forecast for three time periods: the Election Eve forecast (1 day prior), 7 days prior to Election Day, and 93 days before the election. Across all five elections, the PollyVote yielded a lower MAE than each of its component methods in all three time horizons. For the full 93 days prior to Election Day, errors were particularly

small compared to Polls^{PV} and Models^{PV}, although Polls^{PV} became more accurate closer to the election. Among the four components, IEM^{PV} was most accurate. (Note that the experts' forecasts were only available for the last two elections, 2004 and 2008.) See Appendices 4 and 5 for MAEs and the error reduction of the PollyVote relative to each component forecast per election year.

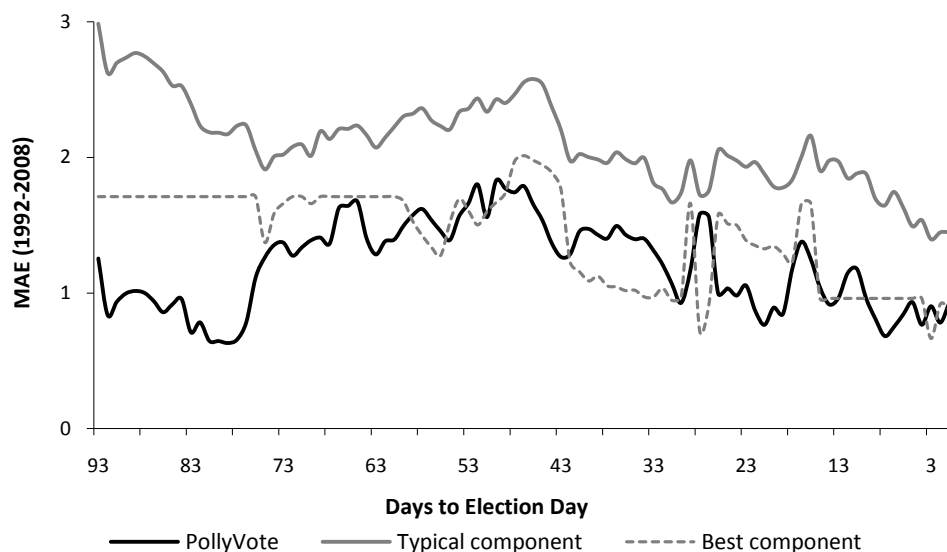
Table 4: MAE of the PollyVote and Its Components, 1992-2008

Time horizon (days to Election Day)	PollyVote	Polls^{PV}	IEM^{PV}	Models^{PV}	Experts*
1	0.9	1.1	1.1	2.6	1.0
7	0.8	1.5	1.0	2.6	1.0
93	1.2	2.8	1.3	2.7	1.7

*only for the two elections in 2004 and 2008

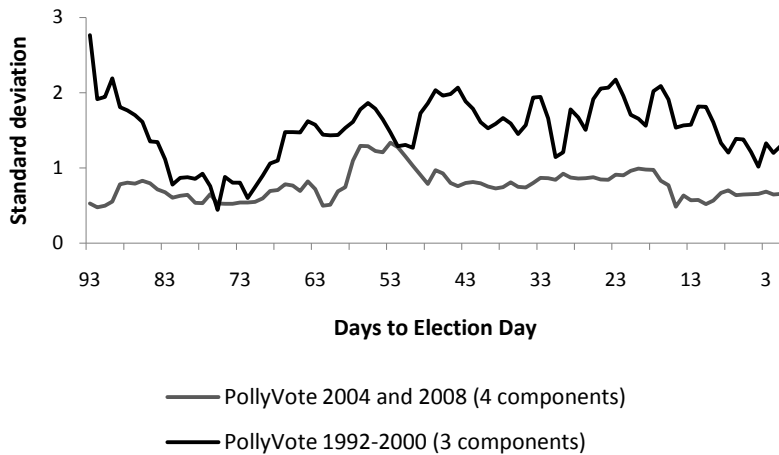
Figure 3 compares the mean absolute error of the PollyVote to a typical component and to the most accurate (“best”) component for the last five elections, 1992 to 2008. On each of the 93 days prior to Election Day, the PollyVote error was substantially smaller than the typical error of its components. Furthermore, the PollyVote was often more accurate than its most accurate component, the performance of which is plotted in the figure as the "best component" line. Normally, the combined forecast is more accurate than its best component only in ideal conditions. That is, if the individual component forecasts bracket the true value in a way that no component forecast is more accurate than the combined forecast. Within the 93-day time horizon, the PollyVote was more accurate than its best component on 78 days in 1992, 15 days in 1996, 38 days in 2000, 35 days in 2004, and 5 days in 2008. On average, the PollyVote forecast was more accurate than its best component on 37% of all days across the five elections. Not surprisingly, the PollyVote was more accurate than the best component for long-term forecasts during the first third of the 93 day forecast period, when uncertainty was highest.

Figure 3: Mean absolute errors of the PollyVote, its typical component, and its best component



Combined forecasts create problems for the traditional estimates of uncertainty favored by statisticians. However, these problems can be avoided by using empirical estimates of uncertainty based on out-of-sample forecasts. To assess out-of-sample accuracy, we calculated the standard deviation for the five elections for which we made PollyVote forecasts. The uncertainty is expected to vary according to the forecast horizon. Figure 4 shows the average standard deviation of the PollyVote with all four components (as implemented in the last two elections) as well as for the three elections from 1992 to 2008, when no expert forecasts were available. Uncertainty was higher for the elections in which only three components were available. As expected, uncertainty was low closer to Election Day.

Figure 4: Average standard deviation of the PollyVote with four and three components



Predicting the election winner

Thus far we have assessed the capability of the PollyVote and its component methods to make point forecasts of the vote share of the incumbent party’s candidate. We now turn to an evaluation of these forecasting techniques in predicting the election winner. How well can they predict the winner, not only shortly before Election Day, but far in advance? To make this evaluation we counted the number of days on which an approach correctly predicted the winner of the election. We refer to that result as the "hit rate."

Table 5: Hit Rate (%) of the PollyVote and Its Components

	1992 (n=93)	1996 (n=93)	2000 (n=93)	2004 (n=246)	2008 (n=432)	Total (=957)
PollyVote	100	100	89	100	100	99
Polls ^{PV}	100	100	46	58	84	77
IEM ^{PV}	75	100	37	92	97	88
Models ^{PV}	0	100	100	100	100	90
Experts (combined)	-	-	-	13*	100	89**

* n=93; ** n=525

Table 5 shows the hit rate for the PollyVote and its components for each of the last five elections, as well as for all elections combined. Over the five elections, comparable forecasts were possible for 957 days. The PollyVote's hit rate was 99%. That is, on 99% of the days included in this study, the PollyVote correctly predicted the election winner. No

method that was a component of the PollyVote matched this performance.

The PollyVote versus uncombined polls and IEM forecasts

In the preceding analysis we compared the performance of the PollyVote relative to its components, which are comprised of combined forecasts produced by each respective method. In this final section, we compare the PollyVote with *uncombined* polls and original *uncombined* IEM forecasts across the forecasting horizons examined previously. We limit the analysis to polls and IEM forecasts because they provide many more data points across longer periods than exist for the other two methods, the regression models and the expert surveys.

Table 6: MAE of the PollyVote, the Typical Poll and the Original IEM, 1992-2008

Time horizon (days to Election Day)	PollyVote	Typical poll	IEM
1	0.9	2.0	1.9
7	0.8	2.7	1.4
93	1.2	3.2	1.4

Table 6 shows the mean absolute error of the PollyVote, the typical poll and the typical IEM trading price for the five elections that we have studied, 1992 through 2008. In each of the three time horizons, the PollyVote was most accurate. (See Appendices 6 and 7 for the MAEs and the error reduction of the PollyVote relative to the typical poll and the original IEM forecasts per election year.) Also, as shown in Figure 5, on each of the 93 days prior to Election Day, the PollyVote MAE was smaller than the MAE of the typical poll. On 55 days out of 93 days, the PollyVote was, on average, more accurate than the IEM. On the remaining 38 days, the IEM was more accurate. Not surprisingly, the original IEM did not achieve the PollyVote’s near perfect hit rate of 99%. Yet, IEM accuracy was still high; on 87% of all days the IEM predicted the winner correctly.

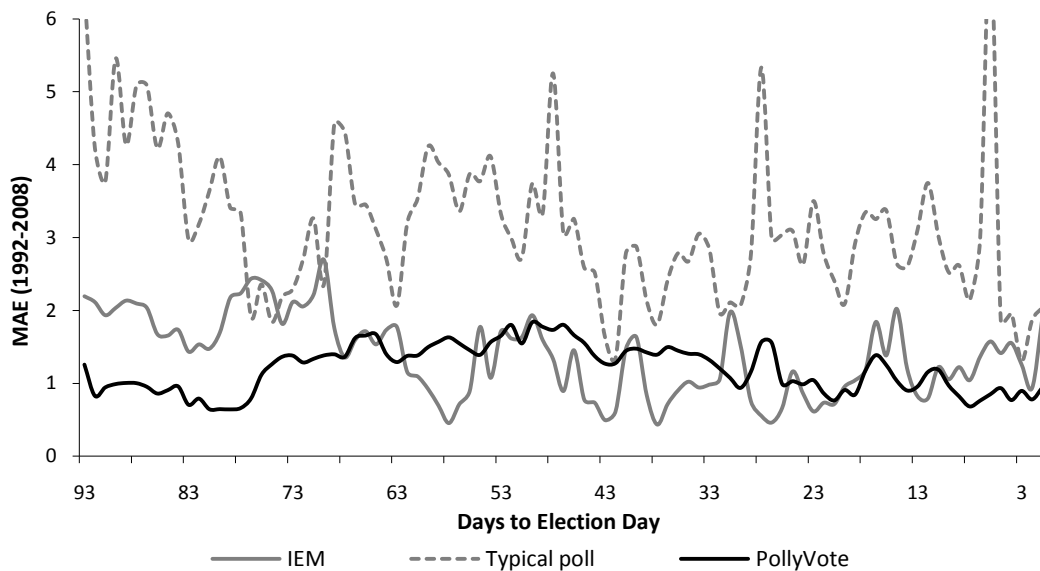
Discussion

In this study we have shown that combined forecasts reduced error of the typical individual forecast within each component method. Combining across methods led to additional gains in accuracy, compared to Polls^{PV}, expert forecasts, and quantitative

models. In addition, the PollyVote was more accurate than its best component, the IEM^{PV}, and substantially more accurate than the original uncombined IEM values.

The success of the PollyVote was achieved by simply averaging the forecasts derived from each component method, weighting each equally. Our results supported this procedure, since there was no component that was more accurate than the PollyVote across all five elections. The basic message is that one should not put much faith in individual forecasts. Combining forecasts within and across methods is preferable by far.

Figure 5: Mean absolute errors of the PollyVote, the IEM and the typical poll (five elections from 1992 to 2008)



On average the IEM^{PV} was clearly more accurate than each of the three other components. Thus, in the future perhaps greater weight could be assigned to the IEM^{PV}. However, care must be taken not to impact Polly’s performance in predicting the correct winner. With a hit rate of 88%, the IEM^{PV} only ranked third among the four components. The optimal weighting of the four components will be determined by further research. In addition, we will follow future advances in election forecasting, particularly regarding poll damping.

Further developments continue for election forecasting and they will be incorporated in

the next PollyVote. For example, the number of models continues to grow; different methods have been developed; and, as Erikson and Wlezien (2008) found, damping of polls offers an opportunity. As experience is gained, differential weighting of the methods might eventually be appropriate. However, the level of accuracy is limited by the measurement error in the dependent variable, the vote share. Mistakes and cheating are always present, although they are expected to be quite small in U.S. presidential elections. However, the greatest potential lies not in reducing accuracy, but in providing forecasts that can aid in policy decisions.

Conclusions

Election forecasting provides an ideal opportunity for combining forecasts because there are a number of data sources and several available forecasting methods. Furthermore, the methods produce forecast errors that are not highly correlated with one another. As a result, the error reductions were substantially larger than those that had been reported in prior research on combining.

The PollyVote method of combining produces accurate forecasts of U.S. presidential elections. This approach should be applicable to predicting other elections and, more generally, can be applied in many other contexts, as well. Given the methods available to forecasters, combining is probably the most cost efficient way to improve forecast accuracy and to prevent large errors.

References

- Armstrong, J. S. 2001. Combining forecasts. J. S. Armstrong, ed. *Principles of Forecasting: A Handbook for Researchers and Practitioners*, Kluwer, Norwell, 417-439.
- Berg, J. E., F. D. Nelson, T. A. Rietz. 2008. Prediction market accuracy in the long run. *Int J Forecasting* **24** 285-300.
- Erikson, R. S., C. Wlezien. 2008. Are political markets really superior to polls as election predictors? *Public Opin Quart* **72** 190-215.
- Gott, J. R., W. N. Colley. 2008. Median statistics in polling. *Math Comput Model* **48**

396-1408.

Jones, R. J., J. S. Armstrong, A. G. Cuzán. 2007. Forecasting elections using expert surveys: An application to U.S. presidential elections. *MPRA working paper No. 5301*.

Jones, R. J., A. G. Cuzán. 2008. Forecasting U.S. presidential elections: A brief review. *Foresight Int J Appl Forecasting* **2008** 29-34.

Larrick, R. P., J. B. Soll. 2006. Intuitions about combining opinions: Misappreciation of the averaging principle. *Manage Sci* **52** 111-127.

Rhode, P. W., K. S. Strumpf. 2004. Historical presidential betting markets. *J Econ Perspect* **18** 127-141.