Election Forecasting: Too Far Out?

Will Jennings, University of Southampton
w.j.jennings@soton.ac.uk

Michael Lewis-Beck, University of Iowa
michael-lewis-beck@uiowa.edu

Christopher Wlezien, University of Texas at Austin
wlezien@austin.utexas.edu

International Journal of Forecasting, forthcoming

Abstract

We consider two criteria for evaluating election forecasts: accuracy (precision) and lead (distance from the event), specifically the trade-off between the two in poll-based forecasts. We evaluate how much "lead" still allows prediction of the election outcome. How much further back can we go, supposing we tolerate a little more error? Our analysis offers estimates of the "optimal" lead time for election forecasts, based on a dataset of over 26,000 vote intention polls from 338 elections in 44 countries between 1942 and 2014. We find that optimization of a forecast is possible, and typically occurs two to three months before the election, but can be influenced by the arrangement of political institutions. To demonstrate how our optimization guidelines perform in practice, we consider recent elections in the UK, the US and France.

Key words: polling, the timeline of elections, forecast, lead time, accuracy, cross-national, political institutions.

Forecasting has an ancient pedigree. The use of the word, “forecast,” as a noun and a verb, appears as early as the Wycliffe Bible of 1388 (Friedman, 2015). As a “verb,” to forecast means “to foresee,” or “to foretell” (Webster’s New Collegiate, 1961, 324). In the social sciences, the pioneers of forecasting appeared in economics, with their efforts to predict national aggregate outcomes, such as growth or inflation (Clements and Hendry, 1999). In political science, forecasting national aggregates, such as election outcomes, arrived later, in the early 1980s (Lewis-Beck and Rice, 1992). Currently, election forecasting has found a niche in the press and the public mind, not to mention the scholarly journals. Its popular appeal—“about seeing into the future”—seems very American; in the words of two US forecasters, “When the election looms far away and we forecast the outcome accurately, that is impressive” (Linzer and Lewis-Beck, 2015, 897). This declaration suggests two important criteria for evaluating an election forecast: *accuracy* (precision) and *lead* (distance from the event).

Forecasting instruments can be evaluated by other criteria, such as parsimony and transparency, but here we focus on these two—accuracy and lead—as applied to a widely practiced method for election forecasts, the utilization of vote intention polls (Lewis-Beck and Stegmaier, 2014). Our central argument can be stated simply. For the election forecaster, lead time represents the sine qua non. In the early days of election forecasting, models appeared that had no lead, i.e., they only generated “forecasts” ex post facto (for example, Rosenstone, 1983; Tufte, 1978). Thus, in the words of Lewis-Beck (2005, 151), “Accuracy is not a sufficient condition. Besides accuracy, a model must have lead, i.e., the forecast must be made before the event.” How much lead is needed? In principle, the more the better, and certainly some lead is better than none, although lead of just a few days risks triviality.

This tension between forecasting lead and vote intention polling has special relevance these days, in the face of apparent, or alleged, "forecasting failures," most strikingly witnessed in the run up to the 2015 British general election, and the 2016 United States presidential election (see, respectively, Fisher and Lewis-Beck, 2016; and Wright and Wright, 2018). These elections have, in fact, prompted serious soul-searching on the part of the professional polling community and public opinion researchers (see, respectively, Sturgis et al., 2016; Kennedy et al., 2018). That said, an expansive historical analysis reveals that the accuracy of pre-election polls really has not changed much over time, and actually might have improved a little (Jennings and Wlezien, 2018).

In what follows, we begin with actual examples of election forecasting lead, as practiced. Then, we explore the general question of how much election forecasting lead is wanted, needed, or even possible. As part of that discussion, we focus on the implications of the trade-off between lead and accuracy. We conclude by offering estimates of "optimal" lead time for an election forecast, drawing on a sample of democratic national elections that goes deep and wide. As examples of how our optimization guidelines perform, we examine the recent cases of the UK 2015 and 2017 general elections, the 2016 US presidential and congressional elections, and the first and second rounds of the 2017 French presidential election. Further, optimal forecasts typically occur two to three months before the contest, subject to some institutional variation.

EXAMPLES OF ELECTION FORECASTING LEAD IN PRACTICE

The clearest examples of lead time, in practice, come from the work of structural modelers, who typically forecast from explanatory equations containing aggregate variables, measured at a specific time point prior to the election. Look at a few scattered results, taking the most studied case, that of the United States, first. In one evaluation of 2004 US presidential election forecasts, seven different modeling teams were found to have a median lead time of 67 days, ranging from a low of 57 to a high of 278 (Lewis-Beck, 2005, 157). Or, in a more recent example, from the 2012 US presidential election, thirteen modelling teams offered forecasts, with a median of 99 days, and range of 57 to 299 days (Campbell, 2012, 612). Such exercises continued for the 2016 election, where there were eleven team forecasts, with a median lead of 78 days, and a range of 60 to 246 days (Campbell, 2016). Of course, more and more structural modelers are set on forecasting national elections in other democracies. The bulk of the work has been done in Britain or France, but the exercise has spread to smaller democracies as well (see, for example, the country studies collection by Bélanger and Lewis-Beck, 2012). In these investigations, three to six months has been a common lead time.

Ignoring the American case, scientific election forecasting has been most extensive in Britain. In the run up to the 2015 contest, twelve teams attempted to call the general election (Fisher and Lewis-Beck, 2016). Looking at the lead times in that collection, the median value was two months, the range from 10 days to 12 months. Distance, then, in certain of those efforts, held a proud place. As Lewis-Beck (2005, 151) sums it up, granting “the forecast a full horizon, says six months to a year, allows for an impressive performance.” But such long-term forecasts may come at the cost of less precision. What are the leading hypotheses about the impact of lead on accuracy? We turn to this question below.

HYPOTHESES RELATING LEAD AND ACCURACY

To begin, we propose three unconstrained hypotheses, the first of which seems to be largely uncontested: *H_1. The less lead, the more accuracy*. This expectation is implicit in most forecasting work, including Erikson and Wlezien (2012, 90), who observed that "the closer we get to the election, the more the polls tell us about the outcome." The idea is that as the campaign plays out, the fundamentals of the election become more and more clear from the polls.

Our second hypothesis posits earlier optima: *H_2. With optimal lead, the more accuracy*. The most notable proponents of this expectation may be Lewis-Beck and Rice (1992, 36-37), who tested H2 against the US and French cases, asserting that optimal accuracy can be obtained well before the election "in the summer months [in the US]… Interestingly, a similar pattern surfaces in French elections." With respect to Britain, a six-month lead argument has also been made (Whiteley et al. 2016). The idea here is that the fundamentals are evident before Election Day, perhaps even well in advance.

The third hypothesis represents an encompassing null: *H_3. Lead and accuracy are unrelated*. Interestingly, this has more support than one might think. Consider certain reviewers assessing the aforementioned forecasts of the 2015 British general election, who concluded "that making a call closer to election day did not really seem to improve accuracy" (Fisher and Lewis-Beck, 2016, 229).

The above hypotheses are rather simple, suggesting that lead either lacks a relationship to accuracy or, more likely, that it takes a first-order (linear) or second-order (fixed parabolic) form. But the hypotheses become more complex when we impose the constraint that, while both lead and accuracy are beneficial, when we have more lead we are likely to have less accuracy. In other words, "complex models offer a tradeoff," a lead-accuracy tradeoff (Linzer and Lewis-Beck, 2015, 897).

How do we achieve, in general, the optimal balance between accuracy and lead? Consider Erikson and Wlezien (2012, 83), who recognize that "polls at the end of the campaign are a good predictor of the vote – the correlation between the two is a near-perfect 0.98." Thus, polls on the eve of an election should work well, at least for the American case. But, they might be "too close" in some sense, when "the party's over" and forecasting is easy. In contrast, polls with more lead may generate more satisfaction, i.e., they may work well-enough to "satisfice." As Erikson and Wlezien (2012, 107) concede, when assessing the *leading candidate* during US presidential elections: "Surprisingly little changes from the conventions to the final week of the campaign." This implies an earlier optimum, where the fundamentals of the election are fairly clear in advance of elections.

There are various reasons, based on previous research, to expect early structuration (see Erikson and Wlezien, 2012). One possibility is that campaigns have little effect, and so preferences do not change much over the election cycle. Another possibility is that campaigns do have an effect but that the effects already are clear well before Election Day. Yet another possibility is that campaigns effects are ongoing, but their individual dynamics cancel out in the aggregate. There are other reasons, but determining those at work in our data is not the focus of this paper, which concentrates on their manifestation in the polls (and the vote).

There also are reasons to think that both lead and accuracy will depend on the particular characteristics of political and electoral institutions. Perhaps most notably, we expect preferences to be more fluid – and less predictable – in presidential systems, where voters possess less information about the attributes and positions of candidates, particularly well in advance of Election Day. By contrast, we expect electoral preferences to be more stable – and predictable – in parliamentary systems due at least in part to the centrality of political parties. As a result, we predict a longer lead-in time in parliamentary systems; indeed, there is reason to expect much the same in non-parliamentary legislative elections as well (see Jennings and Wlezien, 2016).

MODELLING THE ACCURACY-LEAD TRADEOFF

If we assume that increases in lead become increasingly costly, in terms of loss of precision, then we must ask: What is an acceptable amount of loss? Put another way, for every month of lead gained, how much less accuracy can be tolerated? To begin, suppose the following benchmark hypothesis of unit elasticity, where a one month gain in lead means a one percentage point loss of accuracy,

$$\text{Error} = f\ \{\text{ED} + \text{X}\}, \qquad (1)$$

where Error is measured in percentage points, and X is the number of days before Election Day (ED), where ED = 0. Then, an approximation of unit elasticity for one month (30 days), can be expressed as follows,

$$\text{Error}' = b_0 + 0.033\text{X}, \qquad (2)$$

showing that for a one month increase in time to the election the cost is about one percentage point (0.033 x 30 = 0.99). Of course, the relationship could be nonlinear rather than linear, not to mention other complications, which we take the opportunity to explore.

Below, as we start to examine our data, we soon see that this benchmark unit elasticity hypothesis should be rejected, for the estimated coefficient for X is quite small, something under 0.010 (less than a third of the unit elasticity calculation). We could go on to assess the relationship between polls and the vote across different elections on different days. How well do polls forecast the vote, day by day? Further, assuming this relationship changes, how much does it change in fact? To answer these questions, we employ the "timeline" method (especially see Wlezien and Erikson, 2002; Wlezien et al., 2017). Before taking on this challenge, however, we take a detour through the data.

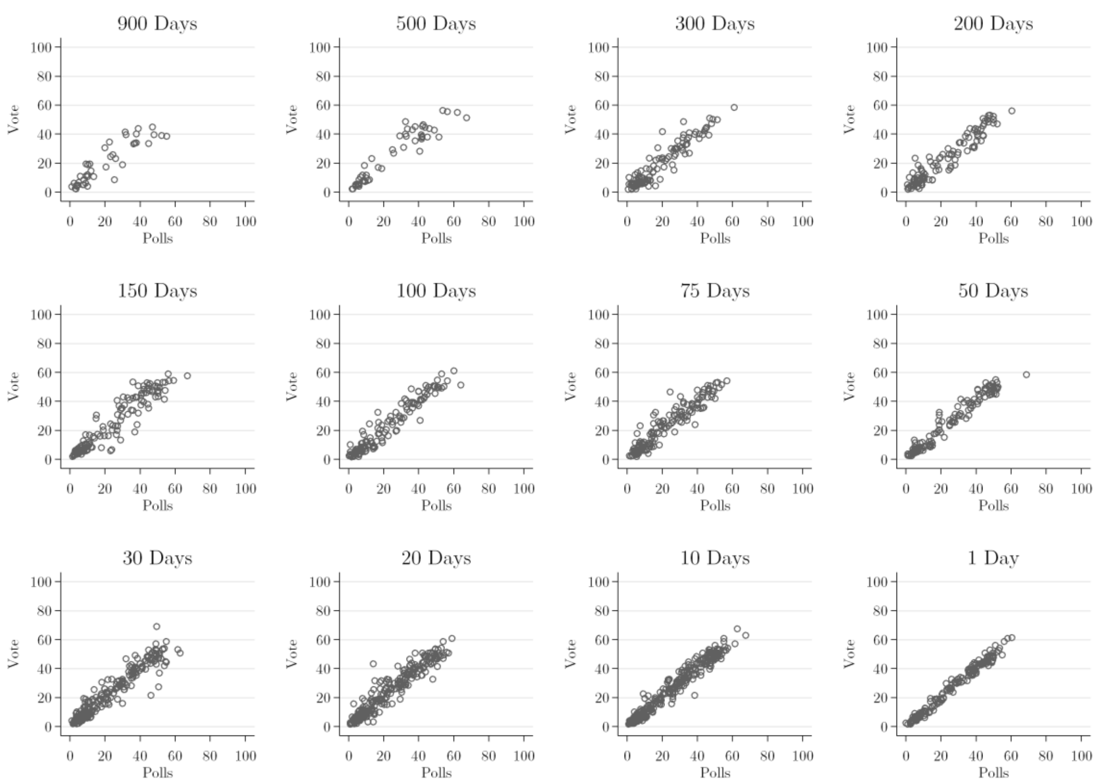
THE POLLING DATA

For our analyses, we use a comparative data set of over 26,000 pre-election polls measuring support for candidates or parties on a given day of the election cycle for presidential and election elections in 44 countries (see Jennings and Wlezien, 2016; 2018). These data consist of national "trial-heat" polls that typically ask citizens how they would vote "if the election were held today?" Sampling or weighting strategies of 'headline' vote intention polls (such as adjustments for likelihood to vote) are often not consistent over time, but may be the only available data, and thus these data reflect survey houses' best estimates of voter preferences at a given point in time. Since the survey fieldwork for most polls is conducted over multiple days, where possible we “date” each poll by the middle day of that period. For days when more than one poll result is recorded, we pool the results together into a single poll-of-polls, simply taking the naïve average of all available polls on each day in question. The final N of 26,324 polls thus produces 21,173 daily poll readings – with each reading capturing support for multiple candidates or parties (amounting to 95,401 poll-party readings in total). Towards the end of the election cycle, we frequently have nearly day-to-day measures of vote intentions.

In total, we have vote intention polls over the period between 1942 and 2014, relating to 338 elections in 44 countries, encompassing presidential elections in 23 countries and legislative elections in 30 countries – and covering 307 parties in total (excluding those parties or candidates that received a vote share of less than 1.5 per cent in a given election). On average, then, we have about 78 polls per election, a respectable number, although of course there exists variation country-to-country. A full list of the countries, ordered by political system and time period, is summarized in the Appendix Table A1.

It is useful as a first step to consider how these vote intention polls line up with the election vote share over time. Figure 1 reports a series of scatterplots where the poll share for a given party or candidate on a given day of the election cycle is plotted on the x-axis and the vote share on the y-axis; specifically from 900 days out in the top left-hand panel to one day out in the bottom right-hand panel. This illustrates how voters’ preferences evolve over the election timeline – and crucially how predictive the polls are of the final result at specific points in time. Nearly three years prior to Election Day, at 900 days out, there is a discernible pattern in the correspondence of polls and the vote though there is substantial variation too, suggestive of heteroscedasticity (with greater variance at higher poll-vote shares). By 500 days out the pattern has become slightly less dispersed, though still with some outlying cases, while by 200 days a linear pattern appears to emerge. At the end of the timeline, just a day before Election Day, linearity has crystallized, with polls exhibiting strong correspondence with the final result. It is clear, then, that the lead time matters, at least on the basis of this visual inspection. From eyeballing the data, substantial predictive power begins to manifest itself sometime after

Figure 1. Party Vote Share by Party Poll Share for Selected Days of the Election Cycle, Elections in 44 Countries 1942-2014



100 days, though much of the structure is in place even earlier, and with it the possibility of rather precise election forecasting. Below, we assess that degree of precision systematically.

THE TIMELINE METHOD

How can we consider the relationship between polls and the vote over the election timeline? The timeline method developed by Wlezien and Erikson (2002) treats poll data as a series of cross sections – across elections – for each day of the election cycle (also see Erikson and Wlezien, 2012). This enables us to assess how polls predict the final vote at different points in time, and in turn enables us to assess the optimal forecasting lead time. There are a number of possible measures for assessing prediction errors over the course of the election timeline (Wlezien et al. 2017). Firstly, there is the mean absolute error (MAE) from the difference between poll and vote shares. This is equal to the mean of the absolute error $\left|\boldsymbol{Poll_{jT}} - \boldsymbol{VOTE_j}\right|$ across *n* observations for parties and candidates *j*, where $\boldsymbol{Poll_{jT}}$ is the poll share on a particular day of the election timeline *T*, and $\boldsymbol{VOTE_j}$ is the vote share for the corresponding parties and candidates:

$$\boldsymbol{MAE_T = \frac{1}{n}\sum_{i=1}^{n}\left|Poll_{jT} - VOTE_j\right|.} \tag{3}$$

The statistic directly captures the match between the polls and votes at a given point in time. It allows us to assess the power of naïve forecasts from the polls to predict the vote.

Secondly, it also is possible to assess how well poll shares *predict* the vote. That is, we can model the vote share for party or candidate *j* in election *k* in country *m* using vote intentions in the polls on each day of the timeline:

$$\boldsymbol{VOTE_{jkm} = a_{jmT} + b_T Poll_{jkmT} + \varepsilon_{jkmT},} \tag{4}$$

where T designates the number of days before Election Day and a_{jmT} represents a separate intercept for each party or candidate j in country m. The latter are important because the level of electoral support can vary systematically across parties. If we wish to consider how polls predict the vote a year before Election Day, we would estimate this equation using polls from 365 days before each election, and repeat this using polls from 364 days in advance, and so on, up to Election Day. Here, the resulting estimates reveal how the polls and the vote line up over time. The regression coefficient b indicates what proportion of the poll margin on each day carries forward to Election Day. As the coefficient approaches 1.0 (and the intercept approaches 0), the poll margin provides an unbiased estimate of the final vote, and a declining MAE. (With a coefficient of 1.0 and an intercept of 0 the R-squared of the equation will tend toward 1.0 and the MAE toward 0.) Most important for forecasting purposes, the root mean squared error (RMSE) of the model offers a measure of fit, the degree to which the vote is *predictable* from current polls.[1] The RMSE is equal to the square root of the mean of the sum of squared errors:

---

[1] The *R-squared* is a useful indicator of fit when comparing parties (or candidates) where vote shares are approximately the same on average, as the statistic is standardized to the total observed variance. It is less useful, however, when comparing parties in different countries, and especially across countries, where the variances in vote shares differ (Lewis-Beck and Skalaban, 1990).

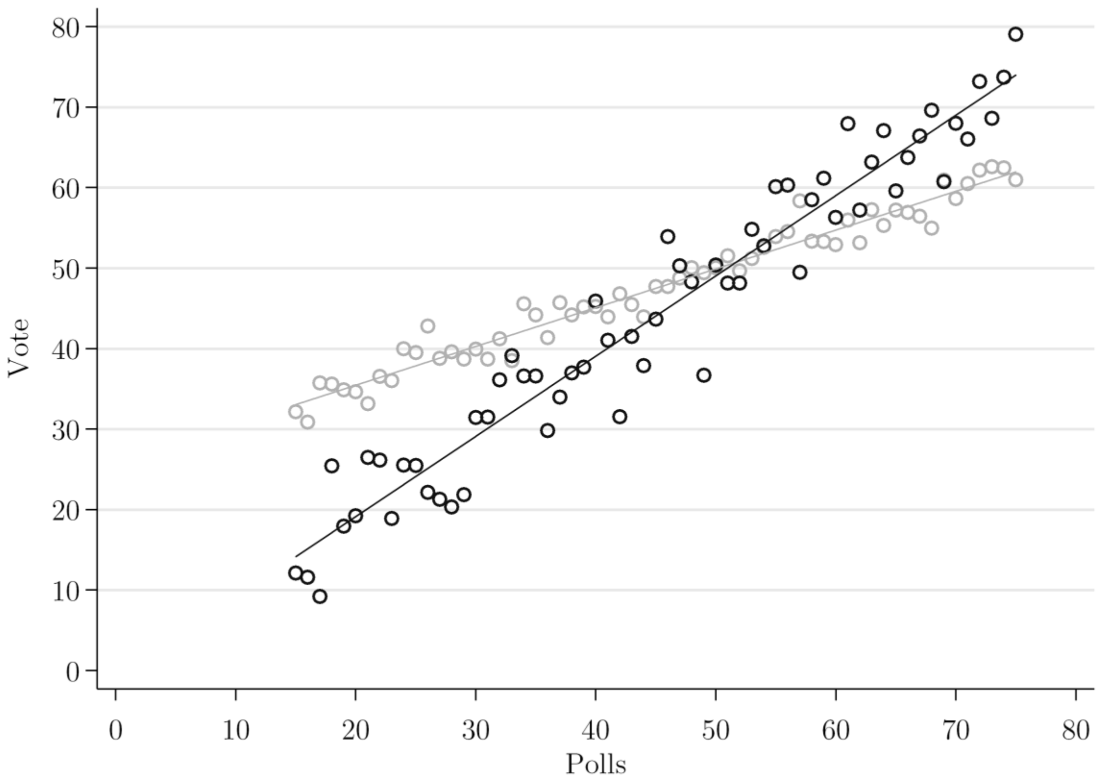
$$\boldsymbol{RMSE = \sqrt{\frac{1}{n}SSE}.} \tag{5}$$

For our assessment of the forecasting performance of polls, we consider *both* the MAE and RMSE, in part because they tap different things (also see Wlezien, et al 2017). To illustrate, consider how the two measures differ for the two sets of hypothetical observations depicted in Figure 2, where the poll share for parties or candidates is plotted on the x-axis and the vote share on the y-axis. Observe the two fitted lines, where one has a coefficient of about 1.0 (the black line), and the other a coefficient of .5 (the light grey line).[2] The poll shares in the first set of observations more directly match the actual vote, though not perfectly, as they are dispersed around the line-of-best-fit. For the other set, indicated in light grey, the poll and vote shares do not directly match but we still can effectively "predict" the latter from the former. Specifically, given that $b = 0.5$, indicated with light grey markers, poll leads are expected to shrink by about 50%. Because of the relationship, the RMSE from regressing polls on the final vote is lower (equal to 1.7) for our set of observations where $b = 0.5$ compared to that where $b = 1.0$ (where it is equal to 4.6). By contrast, the MAE is far higher (8.1 points by comparison with 3.7 points). The point is that the two measures reveal different things about the forecasting power of polls: the MAE indicates how well raw polls themselves match the vote whereas the RMSE indicates how well the polls predict the vote. Put another way, these two measures of fit

[2] The simulated data take the general form $VOTE = a + b_1 Poll + \varepsilon$, where in the first case $a = 0$, $b = 1.0$ and $\varepsilon \sim (\mu = 0, \sigma = 5)$ and in the second $a = 25$, $b = 0.5$ and $\varepsilon \sim (\mu = 0, \sigma = 2)$.

are more beneficial than one, providing different but complementary benchmarks: MAE offers a direct assessment of performance via comparison to the intuitive raw poll number; RMSE, based on a regression model with poll results as the independent variable, utilizes more systematically the information available in past polls, permitting more precise inference to the population vote share.[3]

Figure 2. Simulated data where *b* = 0.5 and *b* = 1.0



[3] By construction, observations of vote intentions for a given party or candidate within each poll are interdependent (as is the actual vote share within each national election). We do not explicitly incorporate this into our analysis, so as to allow more direct comparison of the MAE and RMSE and a direct assessment of total prediction errors.

In practice, pre-election polls are sometimes sparse and carried out at irregular intervals. For the last 200 days of our data-set, polls are missing on around 90% of days. How should we deal with these missing data? Using the raw poll data tends to make the MAE and RMSE estimates noisy, due partly to variation in the elections for which we have polls at a particular point in the timeline. It is possible, following Erikson and Wlezien (2012) and Jennings and Wlezien (2016), to linearly interpolate daily vote intention preferences using available polls. For any date of a given election cycle on which poll data are missing an estimate is calculated as the weighted average from the most recent date of polling and the next date of polling. Weights are in proportion to the closeness of the surrounding earlier or later poll. Given poll readings on days $t - \boldsymbol{\delta}$ and $t + \boldsymbol{\theta}$, the estimate of vote intentions for a particular day *t* is generated using the following formula:

$$\widehat{\boldsymbol{V}}_{\boldsymbol{t}} = \left\{\frac{[\boldsymbol{\delta} \times \boldsymbol{V}_{\boldsymbol{t}-\boldsymbol{\delta}} + \boldsymbol{\theta} \times \boldsymbol{V}_{\boldsymbol{t}+\boldsymbol{\theta}}]}{(\boldsymbol{\theta} + \boldsymbol{\delta})}\right\} + \boldsymbol{\varepsilon}. \quad (6)$$

As Wlezien et al. (2017) demonstrate, timeline estimates using raw data are substantially noisier, but still exhibit the same dynamics over time – with linear interpolation simply more clearly revealing the underlying trend. In addition to linear interpolation, multiple imputation techniques (Rubin, 1987) enable the incorporation of uncertainty (see Jennings and Wlezien, 2016; Wlezien et al., 2017). Here, a random component ε is introduced, based on the known poll variance drawn from a defined distribution $N(\mu, \sigma^2)$, to reflect uncertainty associated with the imputed values.[4] Multiple imputation averages the coefficients across the imputed data series and adjusts the standard errors to reflect noise due to imputation and residual variance.[5]

Additionally, bootstrapping is used to estimate the sampling distribution of our measure of goodness-of-fit – the RMSE – with the regression estimated for randomly drawn resamples (with replacement) of the data repeated 1,000 times for each day of the election cycle. From this we can observe the amount of uncertainty surrounding our estimates.

CALCULATING THE LEAD TIME

[4] Following Jennings and Wlezien (2016), we estimate the underlying variance of all polls in the data-set, once the country, party and election intercepts are controlled for (such that $\mu=0$, $\sigma=3.086$). Specifically, they estimate a regression of the poll share as a function of a separate intercept for each party or candidate j in election k in country m. The residuals of this equation provide our measure of underlying variance of the polls once the country-party-election equilibrium is taken into account:

$$\boldsymbol{Poll = a_{jkm} + \varepsilon}.$$

[5] Rubin (1987) shows that where γ is the rate of missing data, estimates based on m imputations have an efficiency that is approximately $(1+\frac{\gamma}{m})^{-1}$. Since polls are missing on around 90% of days, so we use 50 imputed data series in our analysis, which implies a relative efficiency of 0.98 compared to an infinite number of imputations.

For our analysis of lead time, we need a standard against which to evaluate forecasting performance. The performance at the very end of the cycle – when $T = 1$, the last day before Election Day on which polls are available – seems a logical and appropriate starting point. After all, we want to know whether forecasts going back in time are worse or better, or do not change. For this assessment, we identify the T at which the lower confidence interval of the daily estimates of MAE and RMSE is less than or equal to the final mean value on the last forecast day. That is, we use the point after which *all subsequent values* of the lower bound of the confidence interval are lower than the MAE/RMSE on $T = 1$. That crossing point, as a forecasting site, is optimal, in the sense that any further forecasting gains are statistically insignificant, rendering them trivial in that sense. Put bluntly, any forecasts after that time are not clearly better. (Using *the first point* at which the lower bound of the confidence interval becomes lower than the MAE/RMSE is, of course, another possible cut-point. Using early cut-points, however, may produce later periods where forecasts cease to be as accurate, i.e., registering false positives.)

AN APPLICATION: US PRESIDENTIAL ELECTIONS

Let us start with an example of US presidential elections between 1952 and 2012, which offers a rich data-base for illustrative purposes. The timeline estimates for the MAE are plotted in Figure 3 and for the RMSE in Figure 4 (with 95 per cent confidence intervals plotted as dotted lines). Here we can see that both measures trend downward over time, using poll readings from closer to Election Day. The patterns are different for the two measures, however, starting much higher and declining more sharply for the

MAE. This nicely reflects the difference between the two, where MAE taps errors between raw poll shares and the vote and the RMSE the errors between poll-based model predictions and the vote. Again, we see the value of each measure, respectively.

While MAE (see Figure 3) has a more intuitive interpretation, it generates considerably more error than its more complex counterpart, RMSE (see Figure 4). Only in the final days of the campaign do the two tend to converge on the true vote value,

Figure 3. MAE of poll and vote shares, US presidential elections 1952-2012

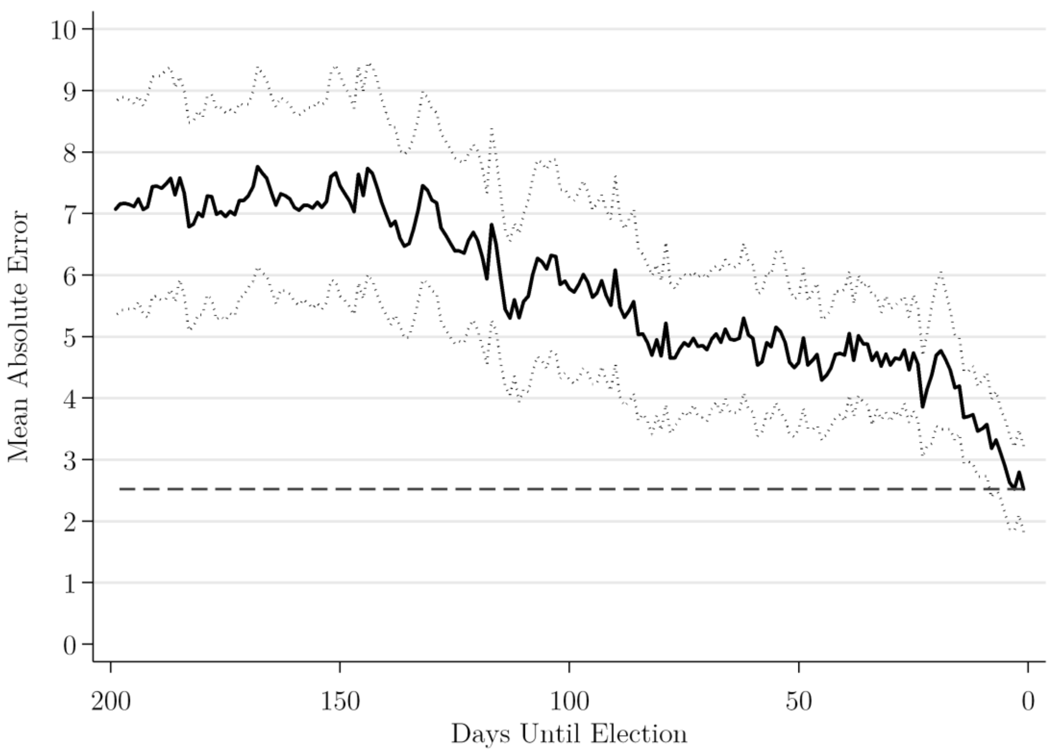
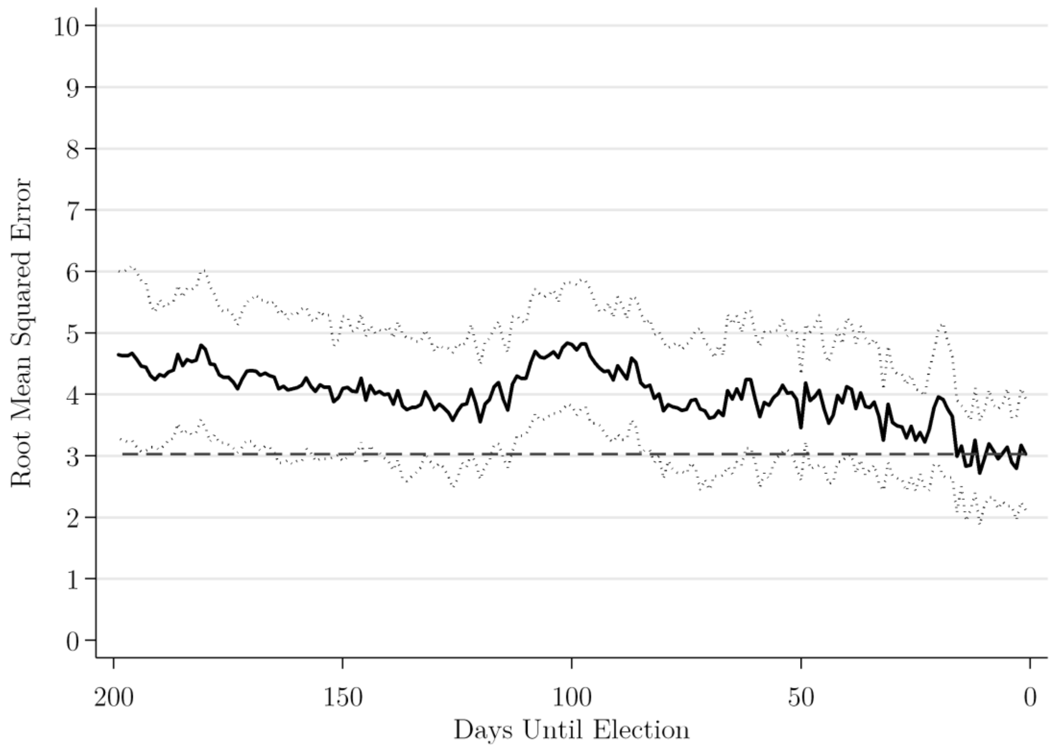


Figure 4. RMSE of poll and vote shares, US presidential elections 1952-2012



though even then polls and poll-based predictions contain a good amount of error. As Erikson and Wlezien (2012) show, while poll shares from early in the election cycle might not provide good naïve forecasts of the vote, they do offer a good amount of information about the ultimate outcome. As the campaign unfolds, and the coefficient systematically relating the poll shares and vote shares increases, the MAE and RMSE increasingly converge.[6] This is clear from the first row of Table 1, which reports the value of both the MAE and RMSE at T = 1, the final day of the election cycle. Here we

[6] When estimated from regressions of the vote on polls *without* party controls, the RMSE numbers are slightly larger than those from regressions including party controls that are reported in the text (per equation 4).

can see that the MAE at the end of the election timeline is 2.52 points while the RMSE is 3.03. The difference in the numbers is due to the calculation of the two measures, where the MAE averages raw (absolute) errors and the RMSE averages squared ones.

HYPOTHESES EVALUATION: US PRESIDENTIAL ELECTIONS

In sum, the patterns in Figures 3 and 4 reveal that the accuracy of forecasts improves the closer the poll readings are to Election Day. This supports our first hypothesis, H1: the less lead, the more accuracy. (It also automatically rejects our third hypothesis, H3: lead and accuracy are unrelated). But, what about our central alternative hypothesis? H2: with optimal lead, more accuracy. When is the optimal lead time? For this, recall that we begin by focusing on the point at which the lower confidence interval band for our measure for a particular day is less than the measure's value on $T = 1$ (and continues to be so for all subsequent days). These are shown in the second row of Table 1. In the case of MAE, the optimal lead based on this standard is 6 days, meaning that we cannot reliably do any better using polls from the last 6 days of the campaign. While the MAE measure is intuitive as a descriptive statistic, for inference the RMSE is preferred when forecasting, particularly across the points in the election timeline. Not surprisingly, the

Table 1. Lead time of predictions, U.S. presidential elections

	MAE	RMSE
MAE/RMSE at *T* = 1	**2.522**	**3.030**
Lower CI ≤ MAE/RMSE at *T* = 1	6 days	48 days
+ 0.5 points	14 days	86 days
+ 1.0 points	19 days	>200 days
+ 1.5 points	36 days	>200 days
+ 2.0 points	89 days	>200 days
+ 2.5 points	115 days	>200 days

RMSE provides an earlier, and more preferred, optimal estimate, specifically at 48 days. Polls from after this date offer us little *additional* predictive power.[7]

What happens if we opt for slightly more relaxed standards? That is, what if we accept a bit more error. In the third row (and rows below) in Table 1 we consider the optimal lead time, permitting additional increments of error in our prediction. Say that we might tolerate an additional 0.5 points error of the MAE in our poll-based forecast, then the earliest date on which the confidence intervals fall within our "range" increases from 6 to 14 days. Thus accepting a slight increase in prediction error extends the lead-in to two weeks. What happens when we tolerate the same additional amount of error (0.5) for the RMSE, our preferred measure? We see that our range increases to 86 days.

[7] The optimal lead time is only slightly shorter (36 days) when the RMSE is estimated without party controls (see Footnote 6), though still longer than that using the MAE.

Overall then, for the US case of presidential elections, the optimal lead time is two to three months. Put another way, forecasts that are closer in time to the election are unlikely to be better. Obviously, this finding has serious consequences for politicians, pollsters and also citizens, who are looking to cast informed votes.

COMPARATIVE APPLICATIONS: LEAD TIMES IN 44 COUNTRIES

Let us now consider the lead time and forecasting elections across different political systems – drawing on poll data for our entire data-set of 44 countries (including the United States). We focus on the final 200 days of the election timeline and limit our analyses (following Jennings and Wlezien, 2016) to cases where we have poll readings over the full period (hence ">200 days" is our maximum lead-in time). Table 2 reports the results for MAE across different election types – all elections, presidential elections, and legislative elections – with the value of MAE at $T = 1$ in the first row. The second row shows the time point at which the lower confidence interval of MAE is *less than* that value for all subsequent time points. Further rows show earlier time points, where a certain amount of additional prediction errors is tolerated.

When reading across the table, from left to right, across the first row, note first, for elections taken together, that the average prediction error of the vote share the day before the contest is about two points. Note, further, that the prediction error is about one point higher for presidential elections, as compared to legislative elections. (Observe, also, the average error for presidential elections generally is about the same as the average error for the US presidential elections in particular: compare row 1, Table 1 to

Table 2). How do these differences play out, in terms of lead time and optimal forecasting?

HYPOTHESES EVALUTION: ELECTIONS IN 44 COUNTRIES

For these set of comparative elections, we can confirm that the Table 2 findings support H1: the less lead, the more accuracy. By the time the final election poll results are released, the expected forecasting error across these democracies falls to about two or three points. Of course, as the forecasts are made at a greater distance from the contests, that error will tend to increase. Again, we return to our central question implied by H3. What is the optimal lead, in terms of the accuracy/distance tradeoff?

According to the MAE measure of Table 2, the optimal lead for this pool of countries is 6 days. That is, for the last week of the campaign, other poll readings cannot significantly add to the reliability of our forecast. If we increase our tolerance of the prediction error by 0.5 points, the lead time increases to almost one month (27 days), while further raising the tolerated error to 1.0 points gives a lead time of fully 93 days – around three months before the election. The patterns we observe across election types exhibit similarities in their overall trends, but with some important variations.

For our base criteria, where the value of the lower confidence interval is less than the MAE on Election Eve, the lead time is greater in presidential elections compared to legislative elections (10 days compared to 6 days). However, if the tolerated error is increased by just 0.5 points, the lead in time in legislative elections increases to 39 days, over double that for presidential elections (18 days). If it is increased a further 0.5 points (to 1.0), the optimal lead time becomes 137 days for legislative elections, whereas it is 57

Table 2. Lead time of MAE predictions

	All elections	Presidential elections	Legislative elections
MAE at *T* = 1	**1.999**	**3.117**	**1.843**
Lower CI ≤ MAE	6 days	10 days	6 days
+ 0.5 points	27 days	18 days	39 days
+ 1.0 points	93 days	57 days	137 days
+ 1.5 points	>200 days	92 days	>200 days
+ 2.0 points	>200 days	123 days	>200 days
+ 2.5 points	>200 days	145 days	>200 days

days for presidential elections. In sum, it is evident that poll readings provide us with a greater lead time in legislative elections. This may well be because voters cast their ballot based on party attachments, which are highly durable (as argued in Jennings and Wlezien 2016).[8]

The results for the RMSE measure, as presented in Table 3, offer a similar picture. For all elections, the base lead time is marginally greater than the MAE (16 days

[8] Note that there tend to be longer lead times for legislative elections in parliamentary compared to legislative elections in presidential (and semi-presidential) systems, consistent with the greater influence of party attachments on electoral preferences. See the Appendix Table A2.

rather than 6 days). However, if the tolerated error is increased by just 0.5 points this lead time is extended to over four months (i.e., 104 days), substantially longer than the

Table 3. Lead time of RMSE predictions

	All elections	**Presidential elections**	**Legislative elections**
RMSE at *T* = 1	**3.211**	**3.725**	**3.099**
Lower CI ≤ RMSE	16 days	30 days	19 days
+ 0.5 points	104 days	48 days	139 days
+ 1.0 points	>200 days	86 days	>200 days
+ 1.5 points	>200 days	>200 days	>200 days
+ 2.0 points	>200 days	>200 days	>200 days
+ 2.5 points	>200 days	>200 days	>200 days

equivalent lead-in for the MAE measure. Indeed, focusing only on legislative elections, the lead-in time increases to 139 days. Again, we see indications that the RMSE offers the better measure: about four or five months before the legislative election, its poll readings provide advance information about the final result.

OPTIMAL FORECASTING OF RECENT ELECTIONS: UK, US, FRANCE

Finally, as an informed experiment, we consider the impact of using our identified optimal lead times for recent elections in three leading democracies where forecasting via polling has been of central importance: the UK, the US and France. In particular, we examine the 2015 and 2017 UK general elections, the 2016 US presidential and house elections, and the first round of the 2017 French presidential election. These are three long-standing representative democracies where there is a tradition of forecasting, and, perhaps most importantly, the long history of polling provides sufficient observations to estimate country-specific lead times.[9] In those recent elections we focus on the two leading parties or candidates in terms of election vote, which facilitates cross-national comparison regarding the optimal lead time. To include parties or candidates receiving a smaller vote share would tend to reduce the average absolute error, impeding comparison across countries with different numbers of contestants. We do not consider the second (run-off) round of the French presidential election, as it is always held within just a few weeks of the first and the frequency of polling also is far less. To introduce the data, we plot (in Figures 5 to 9) the absolute error of the 7-day poll average (indicated with the black line) against the lower and upper confidence intervals of the MAE (marked with the dotted lines) in the respective set of historical elections, i.e., UK general elections 1945-2010, US presidential elections 1952-2012, US house elections 1942-2012, and French presidential elections 1965-2012. We employ comparisons of average errors (MAE), for purposes of illustration and interpretation. Voters' preferences, as can be seen, generally converged on the election result in the usual way.

For most elections the trend of absolute errors fits quite closely with the historical pattern – in particular for the 2015 UK general election, the 2016 US presidential election and 2017 the French presidential election. The exceptions are also interesting. The 2017

[9] As can be seen in Table A1, the US and France provide the longest time series for presidential elections and the UK offers the same for parliamentary elections.

UK general election was unprecedented in post-war elections, seeing a near twenty-point surge in the polls during the campaign for the opposition Labour Party, reflected in the substantial absolute error for most of the timeline before the polls converged on the result late on. (The 2016 elections for the US House of Representatives also saw late movement in the polls, which consistently had the Republican Party well behind, often below 40 per cent until late on, before it secured 49.1% of the vote on Election Day.)

More importantly for this analysis, we can assess the net gain or loss in prediction error between the optimal lead time (for a given set of elections, here calculated using the country-specific estimates of the MAE timeline) and $T = 1$. Table 4 reports the absolute error of the 7-day poll average at $T = 1$ for each election, alongside the absolute error (also of the 7-day poll average) at the optimal lead time (no increased error tolerance allowed). From this we can see how much the prediction error decreases between the optimal lead time and the day before Election Day. Interestingly, we find that, in general, using an optimal lead time of about two weeks, little can be gained from future forecasts.

Specifically, any net decline in prediction error is quite small (0.6 points on average). However, the particular points are spread out. At one extreme, the 2016 US presidential election, the error actually *increases* between $T = 6$ and the day before Election Day (by 0.61 points). At the other extreme, the 2017 French presidential election, the reduction from the identified lead time ($T = 18$) is 1.48 points (a value perhaps driven by the fact it was only the first round of the contest). But these values, even at the extremes, are substantively small. We thus can see how little information the final weeks of polling give us in prediction of the final result. Of course, these case studies, important as they are, remain case studies.

Figure 5. 2015 UK general election, 7 day poll average (Conservatives/Labour)

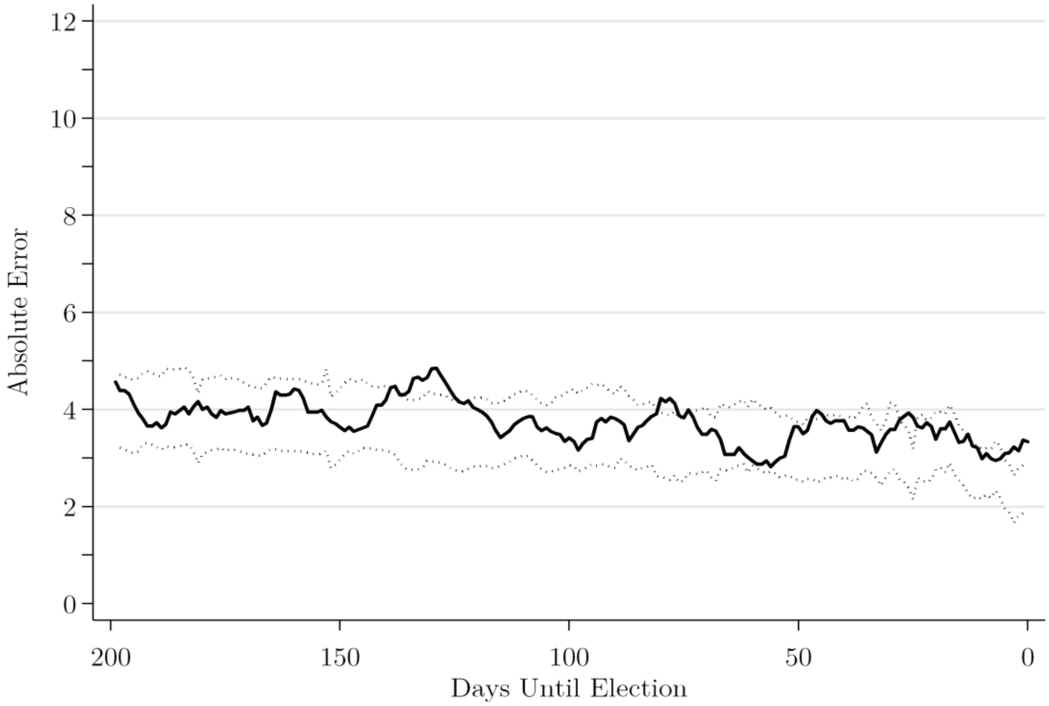


Figure 6. 2016 US presidential election, 7 day poll average (Trump/Clinton)

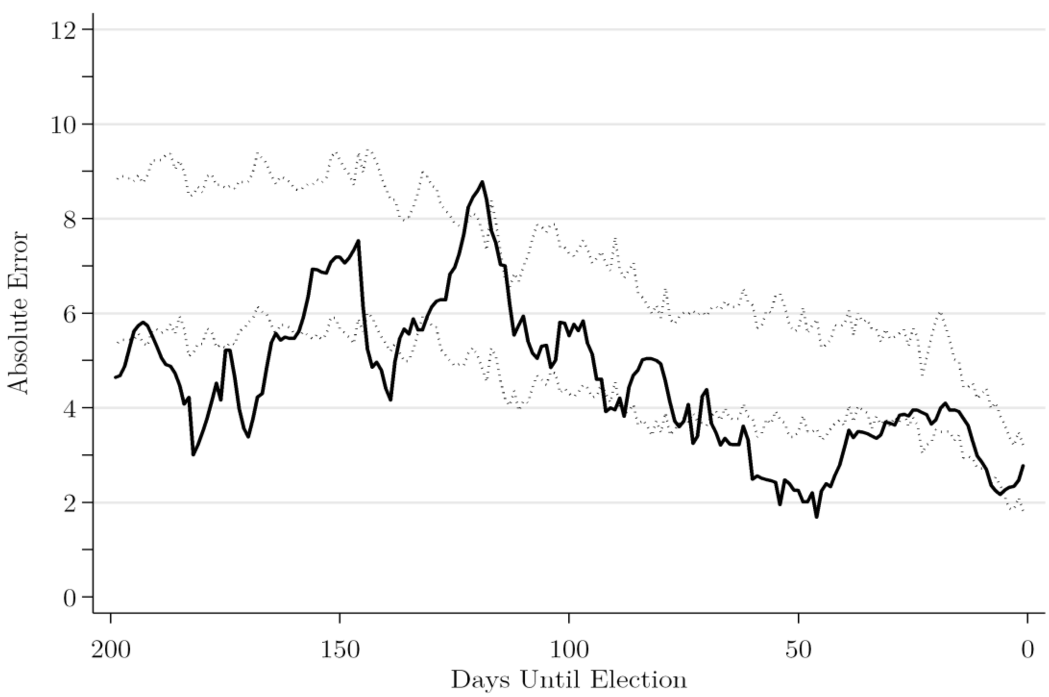


Figure 7. 2016 US House election, 7 day poll average (Republicans/Democrats)

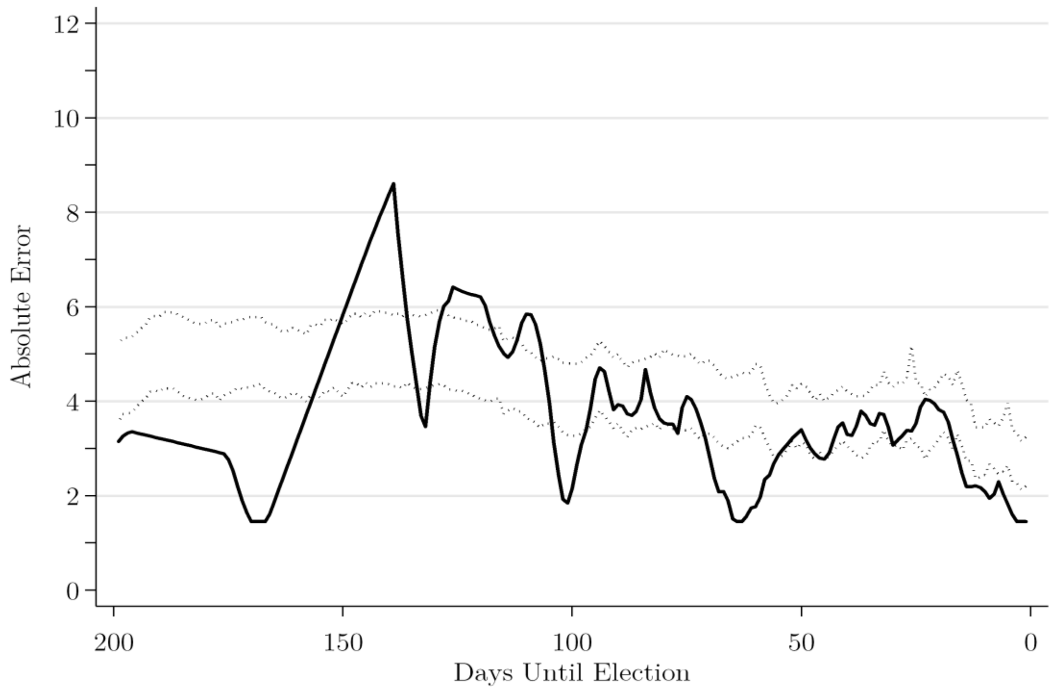


Figure 8. 2017 French presidential election, 1st Round, 7 day poll average (LePen/Macron)

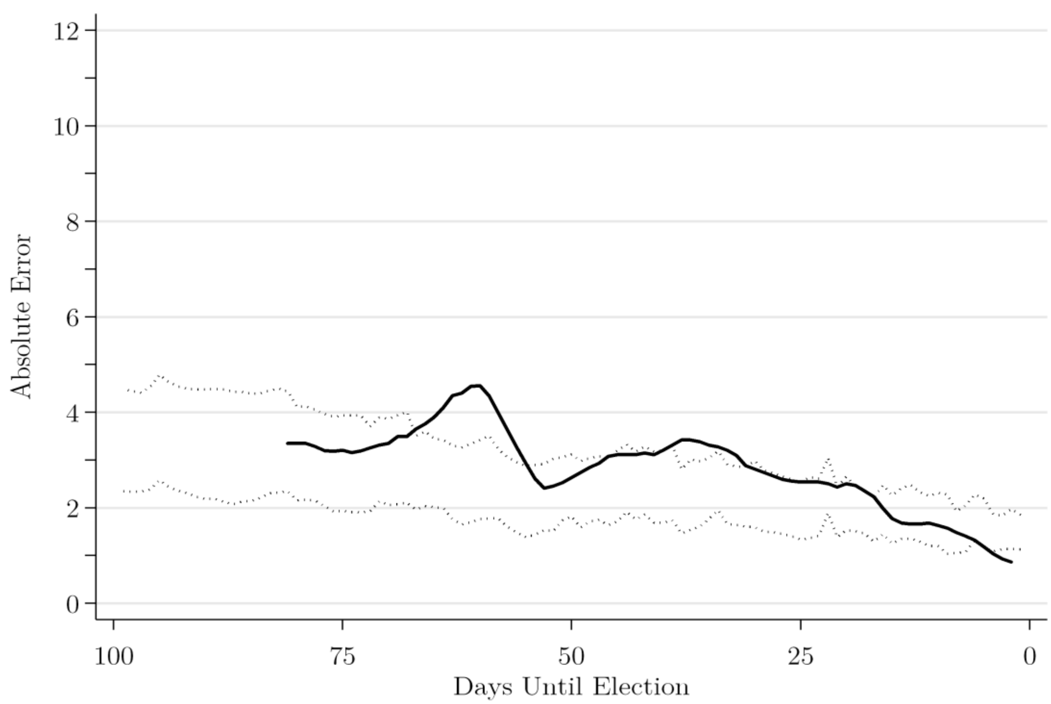


Figure 9. 2017 UK general election, 7 day poll average (Conservatives/Labour)

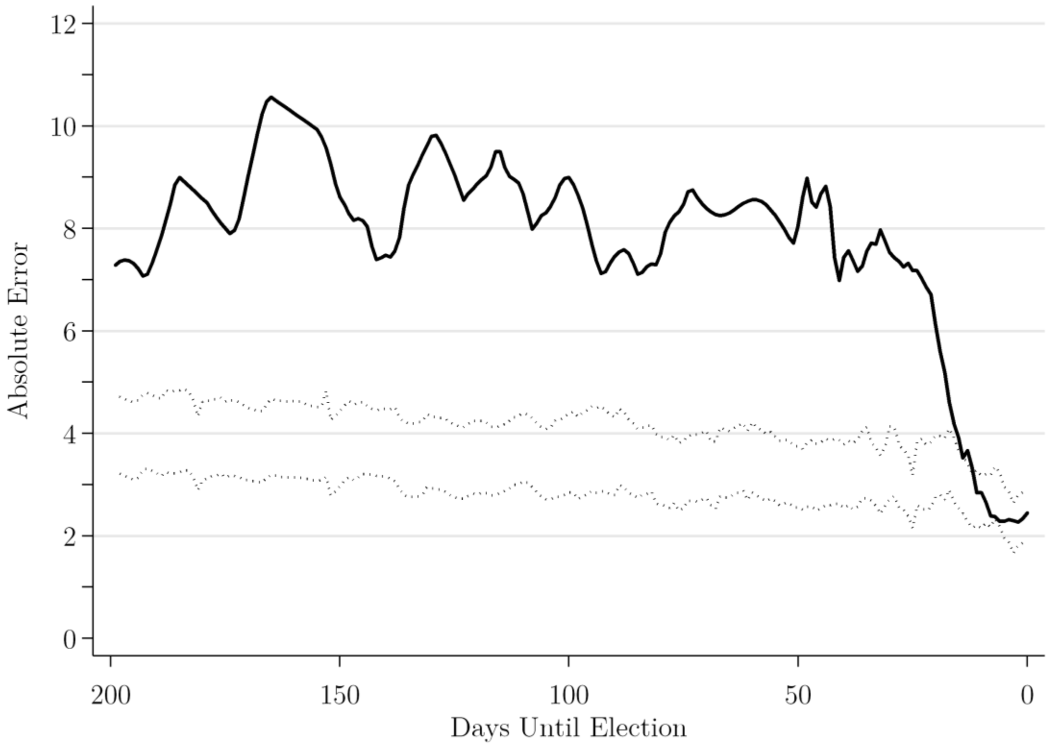


Table 4. Poll errors in US, UK and French elections, 2015-17

	Poll absolute error at $T = 1$	Optimal lead time T (based on national timeline estimates)	Poll absolute error at optimal lead time (T)	Net gain/loss in prediction error between lead time and $T = 1$
2015 UK general election	3.34	13 days	3.50	+0.16
2016 US presidential election	2.78	6 days	2.17	-0.61
2016 US House election	1.45	12 days	2.21	+0.76
2017 French presidential election, 1st Round	0.87	18 days	2.35*	+1.48
2017 UK general election	2.45	13 days	3.67	+1.22

Note: the absolute error is the MAE for the two main parties/candidates (i.e., Conservative/Labour Party, Trump/Clinton, Le Pen/Macron, Republican/Democratic Party).

*The earliest poll data we have for France 2017 is at $T = 80$ days

CONCLUSIONS

In judging the quality of an election forecast, there are different relevant criteria, with accuracy and lead perhaps the most important. Lead time forms an intrinsic part of the forecasting task, a sine qua non. However, as critical as lead may be, it has little value without accuracy. The careful forecaster's dilemma consists in balancing the two objectives, in order to achieve an optimal forecast (Lewis-Beck and Tien, 2016). Herein, we have experimented with different lead times for poll-based forecasts, in order to see what lead "costs" in terms of accuracy, going on to consider how big a cost can be tolerated. To carry out these experiments, we have relied on analyses of pre-election vote intention polls for 338 elections from 44 countries, up to 200 days before Election Day.

As a motivating example, we start with US presidential elections, a central case in election forecasting studies. For a baseline, we use the prediction error (e.g., mean absolute error) from the "final" forecast at *T-1*, the day before the contest (here MAE = 2.52 points). To begin, we denote the optimal lead time as the point at which (1) the baseline value is captured by the lower boundary of the 95 per cent confidence interval around the forecast timeline, and (2) that capture by the confidence interval is sustained until the election. In the case of US presidential elections, this is 6 days before the election, the point at which the predicted vote intention first falls within the confidence interval and after which never falls outside it. A forecast from this distance achieves maximal accuracy, in that later forecasts are very unlikely to show significantly more precision (i.e., the error band will almost certainly not be less 2.52 points). If we are willing to tolerate a bit less accuracy, e.g., 0.5 points less, we can maximize accuracy farther out (e.g., at 14 days), so buying more lead time (and a forecast that is less trivial).

The root mean squared error (RMSE), a measure of statistical fit for prediction accuracy from a regression model of prior polls (rather than simply the current MAE), allows us to obtain still more lead time, without sacrificing accuracy. First, it achieves the premier optimal forecast at a distance of 48 days. If we increase tolerance of inaccuracy by 0.5 points, we can optimize at 86 days out. In sum, we can do about as well predicting the winner roughly three months before the election, as one day before the election. Put another way, once we have reached that point in time, almost no more precision can be gained by waiting.

A question arises as to whether these results are unique to the United States case. In fact, they appear not too different in other democratic systems, as our analysis shows. For instance, with all presidential systems studied, the MAE at time *T-1* (at 3.12) is just over half a point greater than the US case, and the confidence interval captures that value once and for all at 10 days out, a few days more than for the US system. If we tolerate an additional 0.5 point error, that lead time rises to 18 days for presidential systems in general. Turning to legislative systems, they look remarkably like presidential systems at close range, in that the MAE the day before the election has approximately the same value. Presidential systems at first outperform legislative systems in terms of distance and forecasting accuracy; but that standing reverses as distance from the election increases, e.g., for the cost of 0.5 more points of error, legislative systems can forecast from 27 days out, while presidential systems can only manage a distance of 18 days out. This comparative increase in the ability to forecast in parliamentary systems suggests that voter preferences are more stable in these party systems, which is perhaps unsurprising.

As a test of the utility of the optimal lead time notion in election forecasting, we applied it to recent elections in leading Western democracies. For the US (2016), the UK (2015 and 2017), and France (2017), we embedded the vote intention trend (a seven-day moving average) within their historical 95 percent confidence intervals, then marked the optimal lead time, as before. For example, in the 2016 US presidential election, the MAE at *T-1* is equal to 2.78 points. On the basis of the historical confidence intervals, the optimal lead time is equal to 6 days, producing a forecasting error of 2.17 points, an error actually *lower* than the forecast the day before the election. Thus, in this case, the forecast from just under a week out was not only optimal historically, but optimal currently, in that it performed better that the final polls.

The foregoing result suggests, encouragingly, that the notion of optimal distance has value not only in itself, but in terms of precision. Turning to the 2015 United Kingdom general election, we see marginal accuracy gain comes from getting closer than the optimal distance (Wlezien et al 2013). In the 2016 US House elections, a gain in accuracy manifests itself, moving closer to the final polls, but it is slight, just less than one point. The same is true for the 2017 United Kingdom general election, where the gain is slightly greater. For the 2017 French presidential election (first round), the accuracy gain from the final polls asserts itself, registering almost 1.5 points.

What are lessons learned from these election forecasting exercises on accuracy and lead offered here? With respect to our initial hypotheses, several clear findings have emerged. First, in general, decreasing lead increases accuracy. That is, the closer the forecast is to the election itself, the more precise it will tend to be. Second, the gains from decreasing lead are not infinite. Forecasts that are too close to the election face two pitfalls, and both revolve around the possibility of triviality. That is, the forecast may be temporally trivial, e.g., we will know in days if not hours what the actual result will be. Or, the forecast may be statistically and substantively insignificant when compared to an earlier forecast. These possibilities help make our case for optimal election forecasting in general, and as applied to the set of elections studied here. Admittedly, our analysis has solely concerned poll-based forecasts, but structural modelers might adopt this approach to calibrate the optimal lead time for forecasts based on "fundamentals," commonly including economic conditions, or "synthetic" forecasts combining both polls and fundamentals (Lewis-Beck and Dassonneville, 2015). Indeed, the early structuration of voters' preferences suggests that *both* poll-based and fundamentals-based forecasts should be informative well out from Election Day (also see Wlezien and Erikson 2004).

With respect to the general question, it may be that the most accurate forecast is not the final forecast. Perhaps this helps account for why the rule of thumb of forecasting three months out persists. Our evidence for poll-based forecasts can be read to support such a rule of thumb, in that our optimal forecasting bands of two to three months (for presidential elections) and four to five months (for legislative elections) bracket the three-month rule. Further, even if the final forecast has more accuracy than a distant forecast, that accuracy gain might not be worth it—first because of its trivial magnitude and second because of its trivial distance from Election Day itself. There actually appears to exist an ideal point, where the trade-off between distance and accuracy is optimal, i.e., we achieve a substantively accurate forecast at a respectable distance from the election itself. Further research is needed to discover the limits of the optimal election forecasting strategy, and whether (and how) other contextual features and characteristics of political parties (and candidates) themselves condition those expectations.

ACKNOWLEDGMENTS

Earlier versions of this article were presented at the Annual Meetings of the American Political Science Association, Philadelphia, 2016, and San Francisco, 2017. We are grateful for comments we received at those meetings from numerous scholars, including Sabine Carey, Alfred Cuzan, Ruth Dassonneville, Jay DeSart, Andreas Graefe, Randall Jones, Matthew Lebo, Gerald Schneider, and Mary Stegmaier. We also are thankful for the feedback and guidance we received from the editors of the journal and the anonymous reviewers.

REFERENCES

Campbell, James E. 2012. "Forecasting the 2012 American National Elections." *PS: Political Science and Politics* 45(4): 610-613.

Campbell, James E. 2016. "Introduction." *PS: Political Science & Politics, 49*(4), 649-654. doi:10.1017/S1049096516001591

Clements, M.P. and D. F. Hendry. 1998. *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Erikson, Robert S. and Christopher Wlezien. 2012. *The Timeline of Presidential Elections: How Campaigns do (and do not) Matter*. Chicago: University of Chicago Press.

Fisher, Stephen and Michael S. Lewis-Beck (eds.). 2016. "Special Symposium on Forecasting the 2015 British General Election." *Electoral Studies* 41(March): 225-288.

Friedman, Benjamin. 2015. "Forecast." *Times Literary Supplement*, 12 June 2015, Number 5854, p. 6.

Honaker, James and Gary King. 2010. "What to do About Missing Values in Time Series Cross-Section Data." *American Journal of Political Science* 54(3): 561-581.

Jennings, Will and Christopher Wlezien. 2016. "The Timeline of Elections: A Comparative Perspective." *American Journal of Political Science* 60(1): 219-233.

Jennings, Will and Christopher Wlezien. 2018. "Election Polling Errors across Space and Time." *Nature Human Behavior* 2: 276-283.

Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua Clinton, Claire Durand, Charles Franklin, Kyley McGeeney, Lee Miringoff, Doug Rivers, Lydia Saad, Evans Witt, and Christopher Wlezien. 2018. "An Evaluation of 2016 Election Polls in the United States." *Public Opinion Quarterly* 82(1): 1-33.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1): 49-69.

Lall, Ranjit. 2016. "How Multiple Imputation Makes a Difference." *Political Analysis* 24(4): 414-433.

Lewis-Beck, Michael S. 2005. "Election Forecasting: Principles and Practice," *British Journal of Politics and International Relations* 7(2): 145-164.

Lewis-Beck, Michael S. and Éric Bélanger (eds.). 2012. "Election Forecasting in Neglected Democracies." *International Journal of Forecasting* 28(4): 767-768.

Lewis-Beck, Michael S. and Ruth Dassonneville. 2015. "Forecasting elections in Europe: Synthetic Models." *Research and Politics* https://doi.org/10.1177/2053168014565128

Lewis-Beck, Michael S. and Tom W. Rice. 1992. *Forecasting Elections*. Washington, D.C: CQ Press.

Lewis-Beck, Michael S., and Andrew Skalaban. 1990. "The R-Squared: Some Straight Talk." *Political Analysis* 2: 153-171.

Lewis-Beck, Michael S. and Mary Stegmaier. 2014. "US Presidential Election Forecasting." *PS: Political Science and Politics* 47(2): 284-288.

Lewis-Beck, Michael S. and Charles Tien. 2016. "Election Forecasting: The Long View." Oxford Handbooks Online. Oxford: Oxford University Press.

Linzer, Drew and Michael S. Lewis-Beck (eds.). 2015. "Forecasting US Presidential Elections: New Approaches (an introduction)." *International Journal of Forecasting* 31(3): 895-897.

Rosenstone, Steven J. 1983. *Forecasting Presidential Elections*. New Haven: Yale University Press.

Sturgis, Patrick, Nick Baker, Mario Callegaro, Stephen Fisher, Jane Green, Will Jennings, Jouni Kuha, Ben Lauderdale and Patten Smith. 2016. *Report of the Inquiry into the 2015 British general election opinion polls*. Market Research Society and British Polling Council, London.

Tufte, Edward R. 1978. *Political Control of the Economy*. Princeton, N.J: Princeton University Press.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.

Webster's New Collegiate Dictionary. 1961. *Webster's New Collegiate Dictionary*. Springfield, MA: Merriam and Webster.

Whiteley, Paul and Harold D. Clarke, David Sanders, and Marianne C. Stewart. 2016. "Forecasting the 2015 British General Elections: The Seats-Votes Model." *Electoral Studies* 41: 269-273.

Wlezien, Christopher and Robert S. Erikson. 2002. "The Timeline of Presidential Election Campaigns." *Journal of Politics* 64(4): 969-993.

Wlezien, Christopher and Robert S. Erikson. 2004. "The Fundamentals, the Polls, and the Presidential Vote." *PS: Political Science and Politics* 37(4): 747-751.

Wlezien, Christopher, Will Jennings, and Robert S. Erikson. 2017. "The 'Timeline' Methods of Studying Electoral Dynamics." *Electoral Studies* 48: 45-56.

Wlezien, Christopher, Will Jennings, Stephen Fisher, Robert Ford, and Mark Pickup. 2013. "Polls and the Vote in Britain." *Political Studies* 61(S1): 66-91.

Wright, Fred A. and Alec A. Wright. 2018. "How Surprising Was Trump's Victory? Evaluations of the 2016 U.S. Presidential Election and a New Poll Aggregation Model." *Electoral Studies* 54: 81-89.

APPENDIX

Table A1. Poll Data in 44 Countries, 1942-2014

Country	System	Election	Rule	First poll	Last election
Australia	Parliamentary	Legislative	SMDP	1943	2013
Belgium	Parliamentary	Legislative	PR	2004	2010
Bulgaria	Parliamentary	Legislative	PR	2009	2013
Canada	Parliamentary	Legislative	SMDP	1942	2011
Croatia	Parliamentary	Legislative	PR	2008	2011
		Presidential	Majority	2009	2010
Czech Republic	Parliamentary	Presidential	Majority	2012	2013
Denmark	Parliamentary	Legislative	PR	1960	2011
Finland	Parliamentary	Legislative	PR	2010	2011
Finland	Parliamentary	Presidential	Majority	2006	2012
Germany	Parliamentary	Legislative	PR	1961	2013
Greece	Parliamentary	Legislative	PR	2007	2012
Hungary	Parliamentary	Legislative	PR	2009	2014
Iceland	Parliamentary	Legislative	PR	2009	2013
		Presidential	Plurality	2012	2012
Ireland	Parliamentary	Legislative	PR	1974	2011
Italy	Parliamentary	Legislative	PR	2012	2013
Malta	Parliamentary	Legislative	SMDP	2012	2013
Netherlands	Parliamentary	Legislative	PR	1964	2012
New Zealand	Parliamentary	Legislative	SMDP/PR	1975	2014
Norway	Parliamentary	Legislative	PR	1964	2013
Poland	Parliamentary	Legislative	PR	2010	2011
		Presidential	Majority	2011	2011
Serbia	Parliamentary	Legislative	PR	2008	2012
Slovakia	Parliamentary	Legislative	PR	2010	2012

Slovenia	Parliamentary	Presidential	Majority	2012	2012
Spain	Parliamentary	Legislative	PR	1980	2011
Sweden	Parliamentary	Legislative	PR	1967	2014
Switzerland	Parliamentary	Legislative	PR	2010	2011
Turkey	Parliamentary	Legislative	PR	2010	2011
U.K.	Parliamentary	Legislative	SMDP	1943	2010
Argentina	Presidential	Presidential	Majority	2006	2011
Brazil	Presidential	Presidential	Majority	2002	2014
Chile	Presidential	Presidential	Majority	2008	2010
Colombia	Presidential	Presidential	Majority	2010	2010
Cyprus	Presidential	Presidential	Majority	2007	2013
Ecuador	Presidential	Presidential	Majority	2010	2013
Mexico	Presidential	Presidential	Plurality	2005	2012
Paraguay	Presidential	Presidential	Plurality	2013	2013
Peru	Presidential	Presidential	Majority	2006	2011
Philippines	Presidential	Presidential	Plurality	2010	2010
South Korea	Presidential	Legislative	PR	2011	2012
		Presidential	Plurality	2012	2012
U.S.	Presidential	Legislative	SMDP	1942	2014
		Presidential	Electoral College	1952	2014
Venezuela	Presidential	Presidential	Plurality	2006	2013
Austria	Semi-Presidential	Legislative	PR	2006	2013
		Presidential	Majority	2010	2010
France	Semi-Presidential	Presidential	Majority	1965	2012
Portugal	Semi-Presidential	Legislative	PR	1985	2011
		Presidential	Majority	2010	2011
Romania	Semi-Presidential	Legislative	PR	2008	2012
		Presidential	Majority	2009	2009

Table A2. Lead time of MAE predictions

	All elections	Presidential elections	Legislative elections	Legislative elections (parliamentary systems)	Legislative elections (presidential systems)
MAE at T = 1	**1.999**	**3.117**	**1.843**	**1.811**	**2.081**
Lower CI ≤ MAE	6 days	10 days	6 days	9 days	5 days
+ 0.5 points	27 days	18 days	39 days	52 days	16 days
+ 1.0 points	93 days	57 days	137 days	170 days	60 days
+ 1.5 points	>200 days	92 days	>200 days	>200 days	116 days
+ 2.0 points	>200 days	123 days	>200 days	>200 days	>200 days
+ 2.5 points	>200 days	145 days	>200 days	>200 days	>200 days