Election Polling Errors across Time and Space

Will Jennings*, Department of Politics and International Relations, University of

Southampton, Southampton, SO17 1BJ, UK. E-mail: w.j.jennings@soton.ac.uk

Christopher Wlezien*, Department of Government, University of Texas at Austin, Austin,

TX 78712-1704, USA. E-mail: Wlezien@austin.utexas.edu.

* Corresponding authors

Are election polling misses becoming more prevalent? Are they more likely in some contexts than others? In this article we undertake an over-time and cross-national assessment of prediction errors in pre-election polls. Our analysis draws on more than 30,000 national polls from 351 general elections in 45 countries over the period between 1942 and 2017. We proceed in the following way. First, building on previous studies, we demonstrate how errors in national polls evolve in a structured way over the election timeline. Second, we examine errors in polls in the final week of the election campaign to examine performance across election years. Third, we undertake a pooled analysis of polling errors – controlling for a number of institutional and party features – that enables us to test whether poll errors have increased or decreased over time. We find that, contrary to conventional wisdom, recent performance of polls has not been outside the ordinary. The performance of polls does vary across political contexts, however, and in understandable ways.

In the wake of the 2015 UK general election, and the 2016 US presidential election, the performance of the polling industry has been under much scrutiny. Indeed, the performance of polls in the UK and US general elections prompted unusually comprehensive – and lengthy – reports into what went wrong [i,ii]. Those reports suggest that the performance of pre-election polls – at least the national surveys in the US and UK – was largely consistent with the historical norm in terms of the magnitude of polling error. This is not to say that the pre-election polls were without problems, particularly at the state level in the US, where errors reached historical highs [ii]. Further, the claim that polling is in crisis and that poll errors are increasing remains popular, especially with commentators [iii,iv] and even some scholars [v,vi,vii,viii]. But, are these claims true? Are polling misses becoming more common?

In this article we undertake an over-time and cross-national comparison of pre-election poll estimates and election outcomes. Our analysis draws on more than 30,000 national polls from 351 elections in 45 countries over the period between 1942 and 2017. To be clear, we focus only on national polls, and do not examine subnational polls, which fared poorly in the 2016 US presidential election [ii], or referendum polls, the object of scrutiny after the Brexit vote [ix]. We also focus on poll aggregations, not particular polls or survey organizations, that is, we examine the daily average from all available polls – the "poll-of-polls" – where each poll is dated by the mid-point of the fieldwork period.

We undertake a number of analyses. First, building on previous studies, we show how poll errors evolve in a structured way over the election cycle, and illustrate how this varies across election types. Second, we examine errors in polls in the final week of the campaign, to examine poll performance across election years, benchmarking recent polling "misses". Third, we undertake pooled analysis of polling errors – controlling for a number of institutional and party features – which enables us to explicitly test whether poll errors have increased or decreased over time. We find that, contrary to much conventional wisdom, recent performance of polls has not been outside the ordinary; if anything polling errors are getting smaller on average, not bigger.

Pollsters have sought to measure citizen's preferences for candidates or parties for almost three quarters of a century. While the wording of survey questions differs due to differences in context, most pre-election polls ask how citizens would vote "if the election were held today." We draw on what we believe is the most extensive cross-national dataset of polls of vote intentions for presidential and legislative elections [x], extended for this analysis to include elections up to 2017. This dataset consists of 30,916 polls spanning the period from 1942 to

2017. The data cover a total of 351 elections in 45 countries – presidential elections in 23

countries and legislative elections in 31 countries, summarized in Table 1. In most countries

the common practice is to report "headline" vote intention figures excluding don't knows and

refusals, which is what we use here. For our analysis, we therefore drop polls for legislative

elections in Japan because it is the only case where these are included in the headline figures.

Doing this reduces the number of polls by 199 across five elections. Not surprisingly, given

the large number of polls in the dataset, this has marginal consequences for the results.

On average, we have 687 separate polls for approximately eight elections per country, or

about 88 polls per election cycle. Since most polls are conducted over multiple days, we

"date" each poll by the middle day of the period that the survey is in the field. For days when

more than one poll result is recorded, we pool the results together into a single poll-of-polls,

taking the average poll estimate for each party or candidate. In the final week before the

election, we have 1,339 polls over 220 elections in 32 countries. The data used in our

analysis are available at the Harvard Dataverse (http://dx.doi.org/10.7910/DVN/8421DX).

-- Table 1 about here --

Recall that we are interested in the amount of prediction error in these pre-election polls and

how it varies across time and space. We thus need data on the vote shares that parties and

candidates received on Election Day to compare with poll results. For this, we rely on a wide

range of official sources and election data resources (for details, see Jennings and Wlezien [x]).

For our main analysis, the dependent variable is the simple absolute vote-poll error: the

absolute value of the difference between party or candidate share of the polls and the Election

Day vote share. For instance, polls in the final week of the 2015 UK general election

campaign showed the Conservative Party receiving 33.2 percent in the polls, substantially below their vote share of 37.8 percent, producing an absolute error of 4.6 percentage points. In addition to the absolute vote-poll error for each party or candidates, we also consider the log of the odds ratio of the poll to vote share and the absolute error on the margin between the leading two parties or candidates.

Our analysis considers three patterns. First, building on previous research, we examine how poll errors vary over the course of the election campaign. Second, we examine whether and how poll errors at the end of election cycles have varied over time, across election years, and particularly in recent years. Third, we examine whether and how electoral context matters for poll accuracy.

Although our primary interest is in the performance of vote intention polls just before Election Day, it useful to consider how they line up with the election result over the course of the election cycle [xi]. This is important because it reveals how aggregate electoral preferences evolve over the election "timeline" – whether the so-called fundamentals are in place early or come into focus late in the campaign. Fundamentals are those factors that matter on Election Day, and include variables that are "internal" to voters, like party identification, and those that are "external" and influence all voters, such as the performance of the sitting government or state of the economy [xi].

For this analysis, we focus on elections for which we have poll readings beginning 200 days before Election Day, that is, to avoid change in estimates due to the addition of cases over the timeline. This leaves us with 286 discrete election cycles and 216 parties, where independent candidates in presidential elections are treated as unique parties, and where we exclude those

parties whose vote share is less than 5 per cent. In the dataset, polls are missing on 86% of

days on average across parties, which implies that we typically have readings for around 131

parties on each day (with polls dated according to the mid-point of the fieldwork period).

Using these data, we can assess the degree to which the election results match poll estimates

on different days over the timeline.

-- Figure 1 about here --

Figure 1(a) provides a very general take. It plots the mean absolute error $|(Poll_i - VOTE_i)|$

for all parties using polls from each of the last 200 days of the cycle, pooling all 286 elections

for which we have poll data. In the figure, we can see that poll errors decline over the

election timeline. Using polls from 150-200 days before Election Day, the mean absolute

error is close to four percentage points; 50 days in in advance, it is approximately three

points; on the eve of elections, it is under two points. This is not surprising but is satisfying,

as it shows that polls become more reflective of the actual result, though *they remain*

imperfect even at the very end of the campaign. The declining error owes partly to the

increasing number of polls (and respondents) as the campaign unfolds [xii], which helps explain

the dampening oscillations we observe in the figure. Much of the early jaggedness is due to

the relatively sparse N of polls and the changing mix of elections (and parties) from day to

day, which stabilizes as the timeline unfolds and polling increases, thus reducing variation.

While the results in Figure 1(a) are informative about the global pattern in the poll-vote

match over time, there is reason to think that they conceal differences across political

institutions. Scholars have found that preferences evolve differently in different institutional

contexts, after all, which implies that the convergence of the polls on the vote comes into

focus differently. Of special importance is the difference between legislative and presidential

elections [x]. (There is no real difference between legislative elections in presidential and parliamentary systems.) Voters' preferences crystallize earlier in the electoral cycle in the former than the latter. The patterns in Figure 1(b) are consistent with these expectations. At the beginning of the timeline, 200 days out, polls are more informative about the vote in legislative elections, with an MAE of approximately 3.3 percentage points by comparison with 5.9 points for presidential elections. The gap narrows over time, especially during the last 50 days, but a slight difference remains the day before Election Day, as errors are still somewhat larger in presidential elections. By that point in time, preferences in both types of elections seemingly are fully formed.

We have seen that poll errors tend to decline over time within particular election cycles, and particularly in presidential elections. The timing of polls thus is an important factor in their predictive power and that context matters as well. This comes as little surprise but is reassuring. What we want to know now is whether and to what extent the predictive power of the polls has changed over time, across elections. Are polling errors more common today than in the past?

There are reasons to believe polling errors might have increased over time.

First, new, less expensive and easier polling methods – most notably online polling and interactive voice response (IVR) polls – have emerged. As such, pollsters now are using many different methods, the consequences of which are not fully understood. This can introduce error for each of the organizations employing such methods but also for the industry collectively, insofar as the errors of particular methods (including adjustment procedures) do not cancel out. Even to the extent the average performance remains the same,

the variation in poll accuracy may have changed.

Second, for more established methods such as face-to-face and telephone polling, response rates have declined. Twenty years ago, more than one-third of respondents contacted would take surveys; today, the number is less than 10% [xiii]. To be clear, their rate of contact is approximately 60% and the cooperation rate 15%, yielding a response rate of 9%. This potentially jeopardizes the representativeness of surveys, which has fairly obvious consequences for polling error, and has been implicated in recent polling misses [i,ii]. Non-response bias can be a particular problem when response rates are low, as was highlighted in analysis of the 2015 UK general election polls [xiv]. The trend in response rates thus could pose an increasing problem for pollsters in accurately representing the voting electorate.

Now, while the proliferation of approaches and declining response rates pose real challenges, we nevertheless have a lot more survey respondents, i.e., there are more polls often with larger sample sizes, and most survey organizations have incorporated weighting and other techniques designed to increase representativeness. It thus may be that we actually have a fairly reasonable portrait of electoral preferences, particularly when we combine polls from different survey organizations, where errors may to a large extent cancel out [xv]. Unfortunately, we do not have much information about the practices of different survey organizations, particularly going back in time, which precludes an analysis of the effects of many poll-level factors. Some recent studies of poll performance have examined survey mode, finding minimal effects [ii,xvi], while others find significant "house effects," which lead some pollsters to systematically under- or over-estimate support for particular parties or candidates [xv,xvii,xviii]. We still can examine the performance of poll aggregations.

Polling accuracy is not just about pollsters, of course, as the behaviour of voters matters as well. Of special importance for poll errors is the structure of the vote. We know that traditional cleavages have weakened over time in most countries [ix,xx], with the decline of class voting observed across many countries for several decades [xxi,xxii,xxiii]. As cleavages weaken, voter behaviour becomes less predictable and more susceptible to the influence of short-term factors [xxiv]. This may make polls – even close to Election Day – less predictive of the vote, at least where partisan and other cleavages have declined.

For our analysis of polling accuracy across election years, we focus on polls conducted during the last week of election campaigns in the 220 national elections in 32 countries between 1942 and 2017. (In the 12 other countries, we do not have polls in the final week of election campaigns.) For each party or candidate we calculate the absolute error of the average vote estimate of all polls conducted during the final week of the election campaign. Figure 2(a) plots these errors by year, with the error for each party/candidate indicated with a hollow grey circle, and the mean absolute error across all parties/candidates and elections in a given year indicated with a black circle.

-- Figure 2 about here --

From the figure, it is immediately evident that the number of elections for which we have poll data has increased over time. This partly reflects the growth in the number of democracies over the period, but it also reflects the growth in pre-election polling. While polling has increased over time, polling errors have not. Consider the annual averages of poll errors indicated by the bold circles in Figure 2(a). For the entire period, the mean absolute error was 2.1%. This seemingly is considerably greater than we would expect based on sampling error variance alone. Sampling theory indicates that the poll "margin of error" reflects the

size of the sample and the distribution of responses, specifically the proportion giving a particular survey response. The standard error of a proportion, p, can be calculated using the equation $\sqrt{\frac{p(1-p)}{n-1}}$, where n is the size of the sample. This means that the observed error variance increases as $p \sim 0.5$. The margin of error of a poll thus is equal to $\pm 1.96 \times \sqrt{\frac{p(1-p)}{n}}$. We have data on the sample size of polls for a subset of 16 countries, which enables us to determine that the absolute error expected based on sampling theory, assuming random sampling, is 0.7 percentage points, by comparison with the observed absolute error of 1.8 points for those cases. Given the widespread use of weighting protocols, our estimate may overstate the true amount of sampling error, though note that it also does not take into account other sources of survey error [xv].

As can be seen in Figure 2(a), the average absolute error has bounced around somewhat over the years but has not increased. The mean error was 2.1% during the 1940s and 1950s (in the early days of polling), 2.1% during the 1960's and 1970s, and has been 2.0% since 2000. The bivariate correlation between the polling year and the absolute error is -0.11 ($p = 0.40$, N = 61). Note that if the correlation is estimated for all polls (N = 763), not the annual averages, the correlation is -0.06 ($p = 0.08$). Poll performance just has not changed in a secular way over the last 60 years, certainly not for the worse.

If we restrict our analysis to those countries where we have regular polling over the *same* extended time period (Australia, Canada, Denmark, France, Germany, Ireland, the Netherlands, New Zealand, Norway, the U.K. and U.S.), we observe that poll errors actually have declined over time. This ensures that the result is not due to the changing mix of countries covered by the data. Taking poll readings for these 11 countries from 1977 on, the

negative correlation between polling year and error is even stronger, -0.15 ($p < 0.01$, N =
398); using annual averages, the correlation is a slightly larger in absolute terms, -0.28 ($p =$
0.09, N = 37).

So far, we have considered the trend in the absolute error between poll estimates and the vote
share, but there are other measures of polling accuracy. One possibility is the absolute error
on "the margin," i.e., the lead of the party receiving the highest vote share over the party
receiving the next highest vote share. For two parties, i and j, the absolute error on the
margin can be expressed in the form:

$$Absolute\ error\ on\ the\ margin = |(Poll_i - Poll_j) - (VOTE_i - VOTE_j)|.$$

This often can be more consequential in terms of predicting the election "winner," either in
terms of government formation or in the translation of votes into seats in single-member
district systems. Another possibility is the odds ratio of a poll compared to the actual
outcome [xxv,xxvi]. For party i, receiving p as a proportion of the poll share, and v as a
proportion of the vote share this measure takes the form:

$$A'_i = \ln\left(\frac{p_i}{1-p_i} \times \frac{1-v_i}{v_i}\right).$$

When pooling across parties and across elections, measures of the direction of errors are less
informative – since these tend to cancel out in the aggregate – so here we calculate the
absolute value of the log of the odds ratio.

Figures 2(b), 2(c) and 2(d) plot the LOWESS (locally weighted scatterplot smoothing)
curves with bandwidth of 0.3 for the absolute error, the absolute error on the margin, and the
absolute value of the log odds ratio of the poll and vote shares, using the annual average of
these measures). These reveal that, while polling seems to have become more inaccurate

during the 1960's, performance leveled off thereafter, though with some ebb and flow. In the case of the absolute error and the absolute error on the margin, accuracy actually appears to have increased in recent decades.

While the average errors may have levelled off and possibly declined over time, it could be the case that the spread has increased. To assess this possibility, we plot the LOWESS curve of the yearly variance of the absolute errors across elections in Figure 2(e). (Note that the calculation excludes outliers from 1985 and 2004, detected with the Grubbs-Stefansky test, that otherwise obscure the long-term trend in the variance of the absolute error.) The figure reveals a drop in the variance of absolute polling during the 1950s, where polls are relatively sparse, followed by a steady rise thereafter. This implies that the spread of poll errors has increased, approximately 0.02 of a point per year since 1965. Now, the trend could be due to the increasing diversity of polling methodologies, the greater difficulty of obtaining representative samples, or the nature of the electorates in countries that entered the poll dataset in later periods, among other possible explanations. If we limit analysis to the 11 countries where we have regular polling data from 1977 onwards, the trend in the variance of the absolute error is downwards rather than upwards ($b = -0.02$). This can be seen in Figure 2(f). The trend is flatter but still negative if we consider the annual average variance of absolute errors of the final polls within each election ($b = -0.01$). These results imply that the increase in the growing spread of poll errors is in large part driven by the growing mix of countries in which pre-election polls have been conducted over time.

This historical record provides us with a benchmark to consider the performance of polls in recent prominent elections, many of which have been the subject of scrutiny noted above. Figure 3 plots the absolute error of the polls in eleven elections that have occurred since summer 2015, specifically the U.K. in May 2015, Denmark in June 2015, Greece in

September 2015, Canada in October 2015, Ireland in February 2016, Spain in June 2016, Australia in July 2016, Iceland in October 2016, the U.S. in November 2016, France in March and April 2017, and the U.K. in June 2017. This analysis focuses on the absolute error of the polls for the two main parties or candidates, e.g., Labour and the Conservatives in the U.K., Le Pen and Macron in the 2017 French presidential election, since the error tends to be lower for smaller parties due to sampling theory. In some multi-party systems, the pair of parties receiving the highest vote share differs from the pair receiving the highest poll share (typically due to closeness of the election). In those cases (Denmark, Spain, Iceland), we consider the absolute error for the three largest parties – since our analysis might otherwise ignore an important part of polling misses.

Figure 3 includes a benchmark, denoted with a dashed black line, representing the mean absolute error for "large" parties, i.e., those parties receiving over 20% of the vote share, for the preceding 1942 to 2014 period (equal to 2.3 percentage points). Here we can see that, while for recent elections the poll errors vary from election to election, the average is 2.6 percentage points. This is just 0.3 points higher than the average for large parties; the difference is *exactly* the same if we restrict the comparison to those countries with regular polling between 1977 and 2014. Poll performance in recent elections are not out of line with what we have observed in the past.

-- Figure 3 about here –

Now, it is possible that the patterns we observe in polls over the post-war period reflect the kinds of political and electoral systems that have most frequently held elections. It may be, for instance, that the increasing number of countries using proportional representation (PR) has reduced one source of poll error – since party attachments matter more in those systems

and are more durable than candidate evaluations, which are more central in non-PR settings [x].

Table 2 summarizes the absolute vote-poll error (using the average poll estimate during the last week before the election) by election and party type. Since the numbers are weekly averages, they differ slightly from the daily eve-of-election polls depicted in Figure 1. The results in the table indicate that polls errors are higher in presidential elections (an average of 2.7 percentage points) compared to legislative elections (1.8 percentage points), higher in single-member district (SMD) systems (2.3 percentage points) compared to PR systems (1.6 percentage points) – with a similar difference between candidate- and party-centric systems respectively. Errors are moderately related to the effective number of parties, with an average error of 2.2 where there are fewer than 3.0 effective electoral parties, compared to an error of 1.7 percentage points for equal or more than 3.0 effective electoral parties. (Following Laakso and Taagepera [xxvii], the effective number of electoral parties (*ENP*) is calculated as the sum of the squared fraction of votes (*V*) for each party *i*, divided into one. That is, $= \frac{1}{\sum_{i=1}^{n} V_i^2}$.) Much the same is true for participation in government, as the error of poll estimates for parties in government (2.0 percentage points) is only slightly higher than for parties in opposition (1.7 points). While informative, basic descriptive analyses of poll errors may mislead. That is, some of the differences we observe for one factor may be due to the other factors. For example, opposition parties are more likely to be small parties so their errors may be less to do with their opposition status and more to do with the size of their vote share.

-- Table 2 about here --

A more general modelling strategy to address this would treat the absolute error as a dependent variable, enabling us to conduct simultaneous tests of party and system

characteristics and over-time trend. We could model the error as a function of various features of electoral systems and parties, along with the election year. The equation might take the form:

$$|VOTE - POLL| = a + b_1 Presidential + b_2 PR + b_3 Candidate + b_4 ENP +$$

$$b_5 Government + b_6 Size + b_7 Year,$$

where the absolute error for each party or candidate in each election is a function of some intercept (*a*) plus whether the election is a presidential race (*Presidential*), whether it takes place in a proportional representation system (*PR*), whether the system is candidate-centric (*Candidate*), whether the effective number of parties is greater than 3.0 (*ENP*), whether the party or candidate is in government (Government), whether they received more than 20% in the relevant election (*Size*), and the election year (*Year*). The latter enables us to determine whether polling errors have increased or decreased in magnitude, after accounting for the underlying features of elections.

Results of estimating this equation are reported in Table 3. To begin with, we estimate a bivariate regression relating the absolute error to year, that is, to document the trend in the absence of controls. The results shown in the first column of the table reveal that poll errors have not changed much over time, and that, if anything given the coefficient is less than 0, they may have declined slightly. As can be seen in the second column of Table 3, adding the various contextual covariates makes little difference to the over-time pattern. There is evidence from these results that poll errors are around 0.5 percentage points higher in presidential systems ($p < 0.01$), 0.8 points lower in PR systems ($p < 0.05$) and 0.9 points higher ($p < 0.001$) for large parties. (As noted earlier, the latter is exactly what we expect based on sampling theory.) There are not any significant differences in polling accuracy between party-centric and candidate-centric systems, systems with more or less effective

electoral parties, or government and opposition parties. Most notably for our analysis, the effect of the election year is a trivial -0.001 and does not even approach statistical significance ($p = 0.85$). There simply has been no discernible decline in the accuracy of polls over time.

-- Table 3 about here --

We observe a clearer pattern if we limit the analysis to elections in the eleven countries where we have regular poll data over a concurrent period, i.e., since 1977. This enables us to be sure that over-time trends in poll accuracy are not due to the changing mix of elections and democracies where polls were conducted (or where we have data). The results of estimating the same equations as above for the restricted set of cases are reported in Table 4. In some contrast with Table 3, the bivariate regression in the first column of Table 4 indicates that poll errors in this particular set of countries have not only not increased over time, but actually have decreased. The coefficient (-0.02) implies a fairly substantial decline of about 0.8 percentage points over the last 40 years, 40% of the mean absolute error (2.0 points) in our sample.

The pattern holds when the various covariates are included. As can be seen in the second column, errors tend to be around 0.9 points higher for large parties ($p < 0.001$), exactly as we saw in Table 3. The results in Table 4 also reveal that poll accuracy in the set of 11 countries still is significantly ($p < 0.01$) greater in PR systems, with errors approximately 0.8 percentage points lower. Most importantly, with the controls added, the effect (-0.01) of the election year still is negative and statistically significant ($p = 0.04$). In cases where there is regular polling for elections over a period of almost forty years, the evidence is that polls have become *more* accurate, not less.

16

If we instead use the absolute value of the log odds ratio of the poll and vote shares as the dependent variable, we find similar results – no evidence of an over-time trend in errors for all elections, and a negative and significant trend for countries with regular polling since 1977, again indicating a decline in polling errors. There likewise is no significant time trend when the absolute error on the margin or the absolute value of the log odds ratio of the poll over vote margin are taken as the dependent variable, either for all elections or those where we have regular polling since 1977. In combination, these results highlight that features of electoral systems and political parties are important factors in assessing polling accuracy, and that claims that polls have become increasingly unreliable are not supported by the evidence.

-- Table 4 about here --

Although claims about the demise of pre-election polling have become common in recent times, we find little basis in fact to support them. In fact, some of our findings point to the reverse. Relying on vote intention polls from more than 200 elections in 32 countries over a period of more than 70 years, there is no evidence that poll errors have increased over time. And the performance of polls in very recent elections is no exception. This implies that declining response rates and the growing variation in survey mode, sampling, and weighting protocols together have had little effect on the performance of pre-election polls, at least when taken together. Of course, those trends still may impact the performance of particular polls, which cancel out when different polls are combined. That may help explain why the variance in poll errors across elections has increased over time, though it is important to note that much of the apparent increase reflects the growing mix of countries in which pre-elections polls were conducted. This remains an important subject of future research.

What we do find is that a basic feature of political parties influences the accuracy of polls, namely, their size. This is exactly what we would expect based on sampling theory, and also makes more understandable the seeming surprises in recent elections. That is, while all polls contain error, it tends to be greatest for the largest parties (or candidates), which are the ones competing for power. Moreover, these errors are most consequential when elections are close, as they can be decisive for government control, as was the case in the 2015 and 2017 UK general elections and the 2016 US presidential election. The magnitude of the poll error in each of these elections was not especially unusual given electoral history. It did matter greatly for binary predictions of election winners and losers, however.

Characteristics of political systems also appear to influence the accuracy of polling. There is evidence that errors tend to be lower in PR systems, consistent with vote choices being based on partisan loyalties, which tend to be more structured. Errors also are lower for presidential elections, at least in the US and France, the only presidential systems for which we have data over long stretches of time.

We have not considered all possible explanations of polling error here. Some countries ban the publication of polling results for periods in the run-up to elections, which may condition our results. It is also possible that factors such as changes in electoral rules or registration, or volatility within the electorate itself (from one election to the next), might influence polling accuracy – either due to switching between parties or patterns of turnout that lead pollsters' likely voter models to become ineffective. Similarly, survey mode could give rise to systematic variations in poll accuracy that we were unable to test here. These are also important subjects for future research. And, to be clear, we are not claiming that traditional polling methods are working just as well (or better) now than in the past – or that polling has not changed over time. Rather, we infer based on our evidence that survey organizations

have adapted to the well-known challenges they face, described earlier in the paper. Most notable here may be the almost pervasive use of weighting protocols using relevant population parameters [xxviii,xxix].

Ultimately, while the polling industry faces a range of substantial challenges, we find no evidence to support the claims of a crisis *in the accuracy* of polling. Taking into account both space, i.e., countries, and time provides important perspective on polling accuracy. Periodically pollsters get it wrong, and are subject to a great deal of attention, particularly when this influences expectations of who will form the government. Indeed, this can lead to methodological reflection and innovation, and in turn improvements in polling.

DATA AVAILABILITY

The data used to produce these analyses are available at the Harvard Dataverse

(http://dx.doi.org/10.7910/DVN/8421DX).

CODE AVAILABILITY

Stata do-files to reproduce these analyses are also available at the Harvard Dataverse

(http://dx.doi.org/10.7910/DVN/8421DX).

REFERENCES

i. 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*. London: Market Research Society and British Polling Council.

ii. AAPOR. 2017. *An Evaluation of 2016 Election Polls in the US*. A Report by the American Association for Public Opinion Research Ad Hoc Committee on 2016 Election Polling.

iii. Silver, Nate. 2015. "The World May Have a Polling Problem." *FiveThirtyEight live blog*, 7 May 2015. http://fivethirtyeight.com/live-blog/uk-general-election-2015/

iv. Williams, Joseph P. 2015. 'The Problem with Polls.' *U.S. News and World Report*, 28 September 2015. https://www.usnews.com/news/the-report/articles/2015/09/28/why-public-opinion-polls-are-increasingly-inaccurate

v. Zukin, Cliff. 2015. "What's the Matter With Polling?" *New York Times*, 21 June 2015. https://www.nytimes.com/2015/06/21/opinion/sunday/whats-the-matter-with-polling.html?_r=0

vi. Cassino, Dan. 2016. "Why Pollsters Were Completely and Utterly Wrong." *Harvard Business Review*. https://hbr.org/2016/11/why-pollsters-were-completely-and-utterly-wrong

vii. Santos, Rob. 2016. "Why the polls get it wrong." *Los Angeles Times*, 27 March 2016. http://www.latimes.com/opinion/op-ed/la-oe-0327-santos-polling-problems-20160327-story.html

viii. Barfar, Arash and Balaji Padmanabhan. 2016. "Election watch." *Significance* 13(5): 30-33.

ix.   British Polling Council. 2016. *Performance of the polls in the EU referendum*. 24 June 2016. http://www.britishpollingcouncil.org/performance-of-the-polls-in-the-eu-referendum/

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

xi.   Erikson, Robert S. and Christopher Wlezien. 1999. "Presidential Polls as a Time Series: The Case of 1996." *Public Opinion Quarterly* 63(2): 163-177.

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

xiii.   Keeter, Scott, Nick Hatley, Courtney Kennedy and Arnold Lau. 2017. 'What Low Response Rates Mean for Telephone Surveys.' *Pew Research Center, 15 May 2017*. http://www.pewresearch.org/2017/05/15/what-low-response-rates-mean-for-telephone-surveys/

xiv.   Mellon, Jonathan, and Chris Prosser. 2017. "Missing Nonvoters and Misweighted Samples: Explaining the 2015 Great British Polling Miss." *Public Opinion Quarterly* 81(3): 661-687.

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

xvi.   AAPOR. 2009. *An Evaluation of the Methodology of the 2008 Pre-Election Primary Polls*. A Report by the American Association for Public Opinion Research Ad Hoc Committee on the 2008 Presidential Primary Polling.

xvii.   Pickup, Mark, and Richard Johnston. 2008. "Campaign trial heats as election forecasts: Measurement error and bias in 2004 presidential campaign polls." *International Journal of Forecasting* 24(2): 272-284.

xviii.   Ford, Rob, Will Jennings, Mark Pickup and Christopher Wlezien. 2016. "From Polls to Votes to Seats: Forecasting the 2015 British general election." *Electoral Studies* 41(1): 244-249.

xix.   Kriesi, Hanspeter. 1998. "The transformation of cleavage politics: the 1997 Stein Rokkan lecture." *European Journal of Political Research* 33(2): 165-185.

xx.   Mair, Peter. 2013. *Ruling The Void: The Hollowing of Western Democracy*. London : Verso.

xxi.   Franklin, Mark. 1985. *The Decline of Class Voting in Britain: Changes in the Basis of Electoral Choice, 1964-83*. Oxford: Clarendon Press.

xxii.   Knutsen, Oddbjørn. 2006. *Class Voting in Western Europe: A Comparative Longitudinal Study*. Lanham: Lexington Books.

xxiii.   Evans, Geoffrey, and James Tilley. 2017. *The New Politics of Class: The Political Exclusion of the British Working Class*. Oxford: Oxford University Press.

xxiv.   Kayser, Mark, and Christopher Wlezien. 2011. "Performance Pressure: Patterns of Partisanship and the Economic Vote." *European Journal of Political Research* 50(3): 365-394.

xxv.   Martin, Elizabeth A., Michael W. Traugott, and Courtney Kennedy. 2005. "A Review and Proposal for a New Measure of Poll Accuracy." *Public Opinion Quarterly* 69(3): 342-369.

xxvi.   Arzheimer, Kai, and Jocelyn Evans. 2014. "A New Multinomial Accuracy Measure for Polling Bias." *Political Analysis* 22(1): 31-44.

xxvii.   Laakso, Markku, and Rein Taagepera. 1979. ""Effective" Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12(1): 3-27.

xxviii.   Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua D. Clinton, Claire Durand, Charles Franklin and Kyley McGeeney. Forthcoming. "An Evaluation of 2016 Election Polls in the United States: AAPOR task force report." *Public Opinion Quarterly.*

xxix.   Pew Research Center. 2018. "Our survey methodology in detail." *Pew Research Center.* http://www.pewresearch.org/methodology/u-s-survey-research/our-survey-methodology-in-detail/

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

C.W. and W.J. developed the original study concept. W.J. and C.W. gathered and analyzed the data, and drafted and revised the manuscript. W.J wrote the computer code and generated figures and tables in Stata.

COMPETING INTERESTS

The authors declare no competing interests.

FIGURES

Figure 1. Poll Errors over the Election Timeline.

(a) Mean Absolute Vote-Poll Error over the Election Timeline for All Elections, Raw Polls

(b) Mean Absolute Vote-Poll Error over the Election Timeline for All Elections, Raw Polls, Presidential and Parliamentary Elections
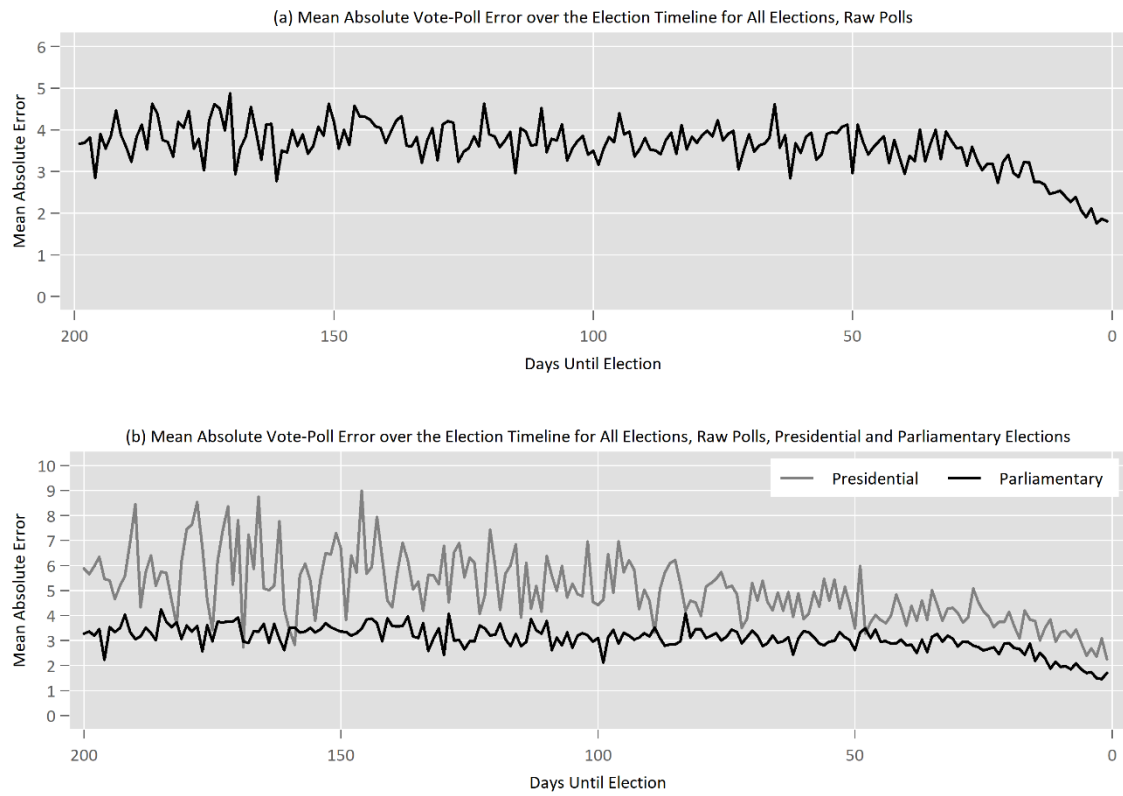
Figure 2. Poll Errors across Electoral History.
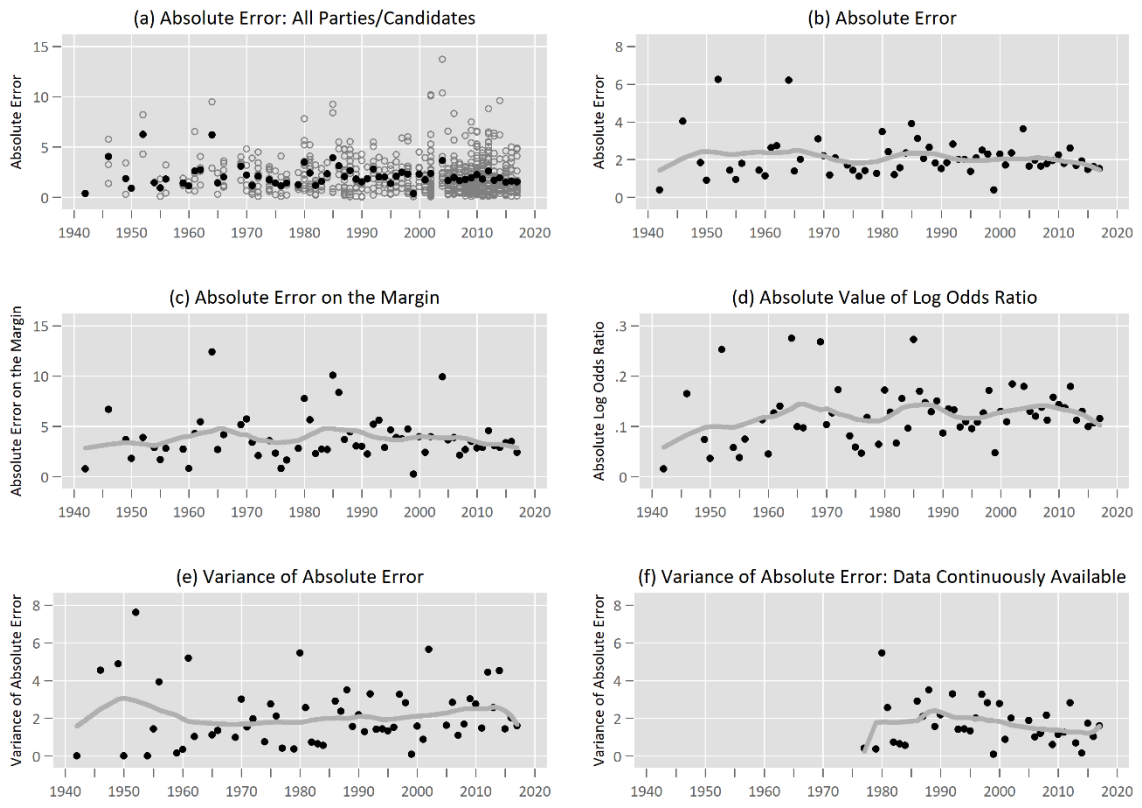
Figure 3. Poll Errors in Recent Elections, 2015-2017.
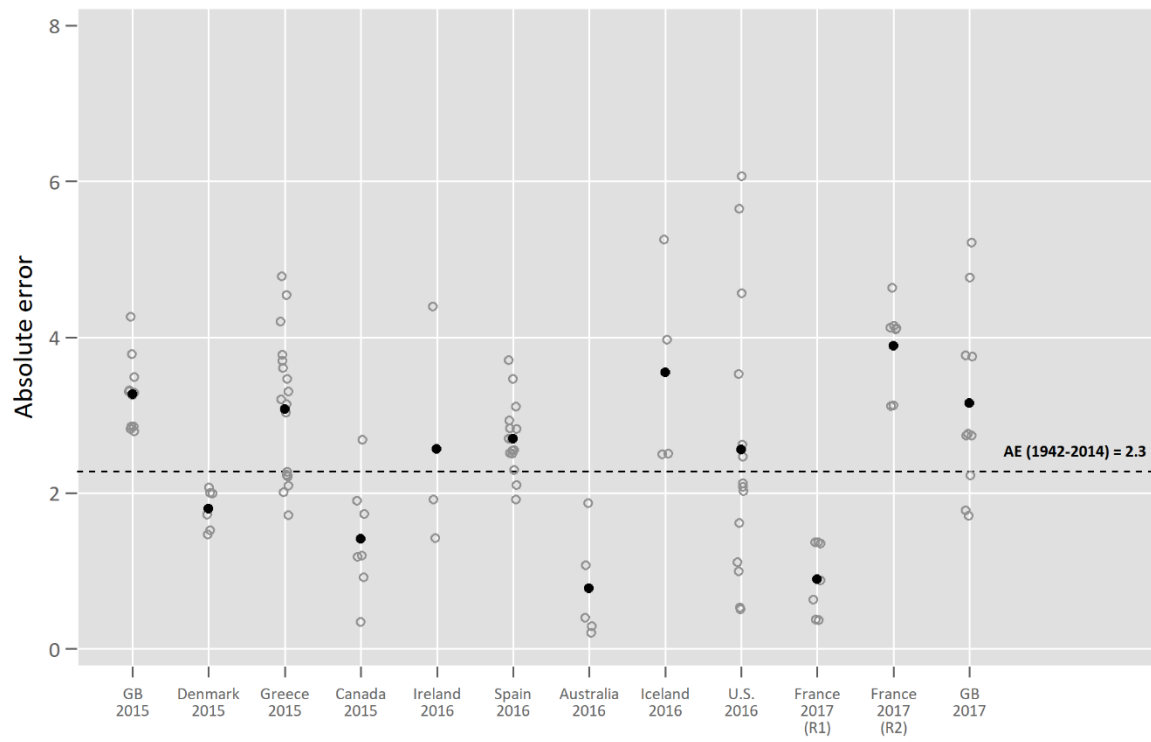
TABLES

Table 1. Poll Data in 45 Countries, 1942-2017

Country	Election	Rule	First poll	Last election	N of polls	N of elections
Parliamentary systems						
Australia	Legislative	SMDP	1943	2016	1,509	29
Belgium	Legislative	PR	2004	2010	32	2
Bulgaria	Legislative	PR	2009	2013	28	2
Canada	Legislative	SMDP	1942	2015	2,071	23
Croatia	Legislative	PR	2008	2011	74	1
	Presidential	Majority	2009	2010	39	2
Czech Republic	Presidential	Majority	2012	2013	9	2
Denmark	Legislative	PR	1960	2011	542	21
Finland	Legislative	PR	2010	2015	15	1
	Presidential	Majority	2006	2012	49	4
Germany	Legislative	PR	1961	2017	3,804	16
Greece	Legislative	PR	2007	2015	476	6
Hungary	Legislative	PR	2009	2014	97	2
Iceland	Legislative	PR	2009	2016	161	3
	Presidential	Plurality	2012	2012	9	1
Ireland	Legislative	PR	1974	2016	573	12
Italy	Legislative	PR	2012	2013	323	1
Japan	Legislative	PR	2000	2012	194	5
Malta	Legislative	SMDP	2012	2013	5	1
Netherlands	Legislative	PR	1964	2012	1,518	16
New Zealand	Legislative	SMDP/PR	1975	2017	816	15
Norway	Legislative	PR	1964	2017	1,571	14
Poland	Legislative	PR	2010	2011	139	1
	Presidential	Majority	2011	2011	34	2
Serbia	Legislative	PR	2008	2012	20	1
Slovakia	Legislative	PR	2010	2012	13	2
Slovenia	Presidential	Majority	2012	2012	19	2
Spain	Legislative	PR	1980	2016	1,434	11
Sweden	Legislative	PR	2000	2014	886	4
Switzerland	Legislative	PR	2010	2011	12	1
Turkey	Legislative	PR	2010	2011	40	1
U.K.	Legislative	SMDP	1943	2017	5,308	20

Presidential systems						
Argentina	Presidential	Majority	2006	2011	52	3
Brazil	Presidential	Majority	2002	2014	247	8
Chile	Presidential	Majority	2008	2010	90	2
Colombia	Presidential	Majority	2010	2010	41	2
Cyprus	Presidential	Majority	2007	2013	52	4
Ecuador	Presidential	Majority	2010	2013	44	1
Mexico	Presidential	Plurality	2005	2012	243	2
Paraguay	Presidential	Plurality	2013	2013	9	1
Peru	Presidential	Majority	2006	2011	189	4
Philippines	Presidential	Plurality	2010	2010	18	1
South Korea	Legislative	PR	2011	2012	15	1
	Presidential	Plurality	2012	2012	30	1
U.S.	Legislative	SMDP	1942	2016	2,336	38
	Presidential	Electoral College	1952	2016	3,778	17
Venezuela	Presidential	Plurality	2006	2013	92	3
Semi-presidential systems						
Austria	Legislative	PR	2006	2013	326	3
	Presidential	Majority	2010	2010	13	1
France	Presidential	Majority	1965	2017	955	20
Portugal	Legislative	PR	1985	2015	485	10
	Presidential	Majority	2010	2011	30	1
Romania	Legislative	PR	2008	2012	23	2
	Presidential	Majority	2009	2009	28	2

Table 2. Absolute Vote-Poll Error for the Last Week before Election Day,
    by Election and Party Types

Elections	*N*	Mean absolute error	Standard deviation
For all elections			
	763	2.03	1.75
By Election type			
Presidential	179	2.70	2.13
Legislative	584	1.83	1.56
For legislative elections			
Electoral system I			
PR	396	1.62	1.40
SMD	188	2.28	1.77
Electoral system II			
Party-centric	368	1.66	1.42
Candidate-centric	216	2.13	1.72
Effective Number of parties			
≤ 3	160	2.16	1.54
> 3	424	1.71	1.55
Party size			
Large (≥20%)	321	2.28	1.73
Small (<20%)	263	1.29	1.08
Incumbency			
Governing party	216	2.05	1.60
Opposition	368	1.71	1.52

Table 3. Regressions of Absolute Vote-Poll Error Using Polls from the Week before Election Day, All Elections with Polls

	Linear Trend Only	General Model With Linear Trend
Presidential elections		0.53
		(0.18)
Proportional representation		-0.83
		(0.33)
Candidate-centric		-0.46
		(0.30)
Effective number of parties > 3		0.17
		(0.16)
Large parties (>20% vote)		0.90
		(0.13)
Governing parties		0.02
		(0.14)
Linear trend	-0.01	-0.00
	(0.00)	(0.00)
Constant	15.70	3.52
	(7.69)	(8.29)
N	763	763
R-squared	0.00	0.13
Adjusted R-squared	0.00	0.12
RMSE	1.74	1.64

Table 4. Regressions of Mean Absolute Vote-Poll Error using
Polls from the Week before Election Day, Countries
Where Data Are Continuously Available since 1977

	Linear Trend Only	General Model With Linear Trend
Presidential elections		-0.26
		(0.22)
Proportional representation		-0.77
		(0.32)
Candidate-centric systems		-0.36
		(0.30)
Effective number of parties > 3		0.12
		(0.21)
Large parties (>20% vote)		0.92
		(0.17)
Governing parties		0.07
		(0.17)
Linear trend	-0.02	-0.01
	(0.01)	(0.01)
Constant	42.63	30.91
	(13.27)	(14.16)
N	398	398
R-squared	0.02	0.14
Adjusted R-squared	0.02	0.12
RMSE	1.58	1.50

Note: Countries are Australia, Canada, Denmark, France, Germany, Ireland,
Netherlands, New Zealand, Norway, the U.K. and the U.S.