FORECAST ERROR: HOW TO PREDICT AN ELECTION: PART 1: POLLS
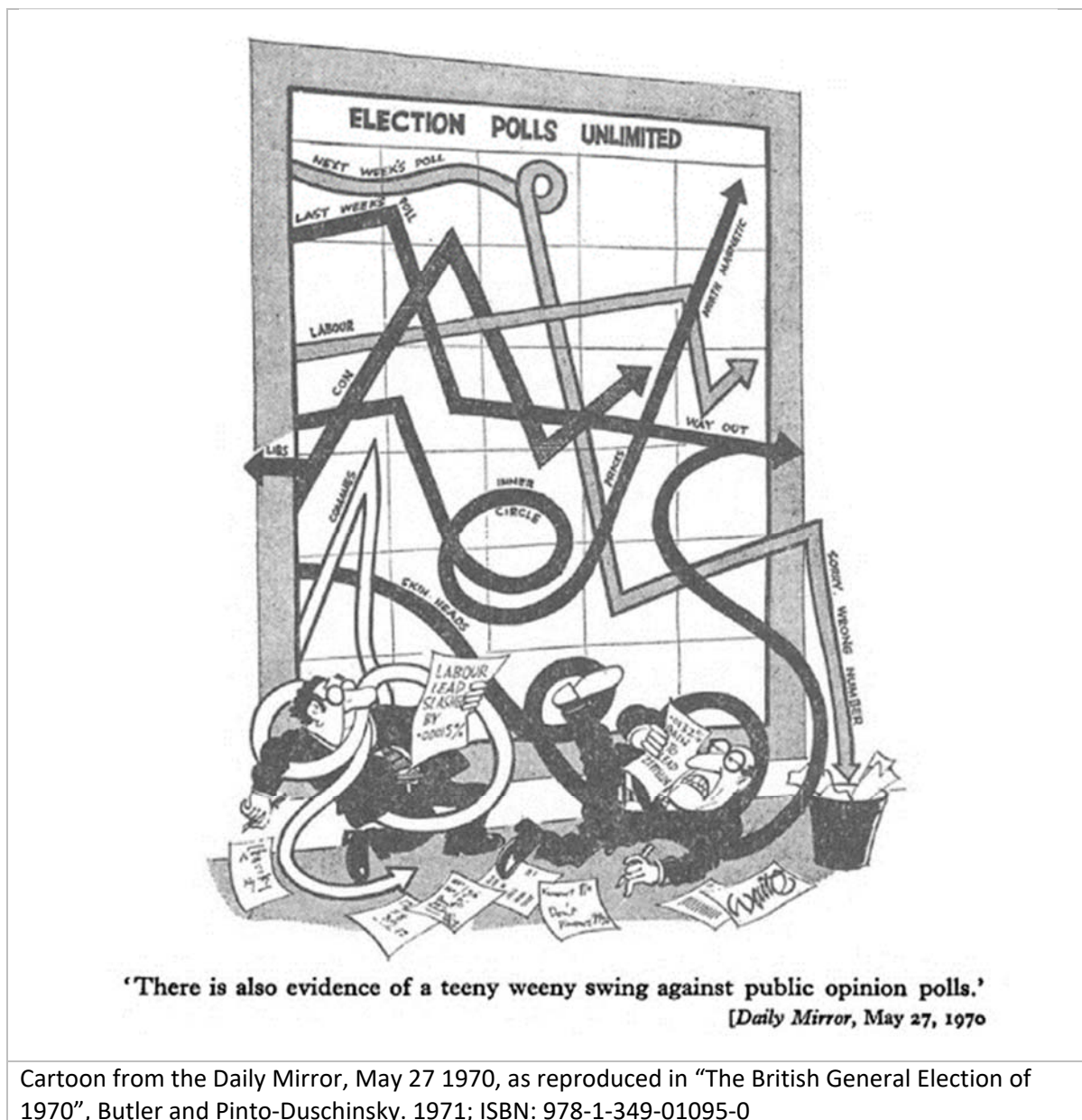
"Any attempt to predict the future depends on it resembling the past"[0423l]

'There is also evidence of a teeny weeny swing against public opinion polls.'
[*Daily Mirror*, May 27, 1970

Cartoon from the Daily Mirror, May 27 1970, as reproduced in "The British General Election of 1970", Butler and Pinto-Duschinsky. 1971; ISBN: 978-1-349-01095-0

1. INTRODUCTION

The "Forecast Error" series of articles started examining election predictors in 2015. Each article considered many predictors, but each article covered just one election. This article marks a new chapter in the "Forecast Error" series where we examine an individual class of predictor more closely across many elections. We begin with possibly the most prominent: opinion polls.

A political opinion poll is a method of assessing a population's opinion on the matters of the day by asking a sample of people. A subset is the voter intention poll, which asks each sampled person how

they intend to vote in an election. Pollsters then turn their intention into votes by applying certain assumptions. Those assumptions may not be valid over the long term, or even from one election to the next.

The problem faced when writing about opinion polls is not how to start writing, it's how to stop. It is entirely possible to write a full article about any given facet of opinion polling, and examples immediately spring to mind: whether one should still use "margin of error" for online panel polling, Nate Silver's insight regarding the borrowing of strength from polls in similar states, is it meaningful to speak of a polling threshold, and so on. To prevent this article becoming infinitely large, it will concentrate on only two things: how well did opinion polls perform as predictors, and how do you measure that performance?

In fairness we must acknowledge that their creators state that political opinion polls should not be used as predictors. In realism we must acknowledge that they are.

2. SCOPE

We cannot cover the whole of space and time. There are so many examples worldwide but this article is finite and we have to limit our scope. Consequently, for a global view we will refer to the work of others, specifically "Election polling errors across time and space" (2018) by Jennings and Wlezien, see [0421f]. For the British view we will concentrate on British General Election opinion polling since 1945 & include other countries/elections when they inform the British experience. We will restrict our use of the term "opinion polls" to refer to voter intention polling only. Exit polls, leader ratings, models and seat projections will be covered in a later article.

3. HOW TO ASSESS ACCURACY

If we are going to judge how predictive polls are, we need to decide on a metric for predictiveness and its close relative, accuracy. In previous articles we discussed metrics and preparation for this article we spoke to notable academics and pollsters to see how they measured it

For the academics we spoke to the Professors Patrick Sturgis, Will Jennings and Stephen Fisher. They used arithmetic metrics to judge the distance between the poll and the outcome to assess accuracy. They used metrics such as root mean square error[0421d][0421e] or log of the odds ratio[0421f] or the net error of the Con/Lab lead[IF1] but the most popular metric was mean absolute error, and it's pretty much a commonplace. The advantages of MAE are too good to overlook: you can use it to enable comparisons with other reports past and present, and it is relatively easy to interpret and understandable to the layman.[IS1] MAE or a variant is the sensible shoe of polling; everybody uses them when they go to work.

For the pollsters we spoke to the pollsters Luke Taylor of Kantar and Anthony Wells of YouGov. They used MAE, but to be precise they looked at each error individually rather than all errors collectively, and tried to reduce them all. That would be closer to maximum error or even just simple error.

For the general public, it's a different matter. In a lecture to the RSS I noted that a gentleman in the audience was getting agitated. He left before I could speak to him but I spoke to some of the other attendees and they explained that the direction of the error should be considered: if the prediction implies a win for one party and the outcome is a loss, then that is important. The public assess polls on who wins,[0430a][IW1][0501a] not the size of the lead, and inquiries are not launched when the polls over-estimate a landslide.[0430a]

This observation about the importance of predicting the winner had been echoed some years prior by Brian Tarran (the editor of Significance) who encouraged the use of a "WIN" indicator. This indicator is set to 1 if the prediction correctly predicted the winner, 0 if not. This indicator is a variant of the Brier score with the f-values replaced by 1 or 0, and the Brier score is recommended by the Good Judgement project[0421a][0421b] and was mentioned by Fisher[IF1]. Other indicators that take the winner into account include the net error of the Con/Lab lead,[0430a][IF1]

So which metric shall we choose? In order to retain compatibility with previous articles, we shall use the WIN on the popular vote and MAE on the popular vote.

4. PREDICTIVENESS VERSUS ACCURACY

At this point we have to discuss the difference between predictiveness and accuracy. To measure predictiveness as distinct from accuracy you must define your metric without knowing the election results beforehand. This is a problem because in the absence of a handy time machine, polls are assessed retrospectively and when a surprise happens (the Canadian Conservatives losing all seats except two in 1993, the UK Liberal Democrats reduced to eight seats in 2015, the Scottish SNP gaining 56 out of 59 seats in Scotland in 2015) the natural human temptation is to measure against that.

Unfortunately this leads to comparison problems: you end up tuning your metric against the SNP and UKIP in 2015, the SDP in 1983…and suddenly you don't have one consistent metric across many elections, you have many different metrics each predicting a different election. To avoid this we need a robust metric that will cope with many elections consistently.

To that end we shall use CON/LAB/OTHER - "three-party-forced" - to judge accuracy for our MAEs. What we lose in detail we gain in robustness as this combination will cope with all UK General Elections since 1945. We did consider even lower numbers than three - Sabato[0421c] just takes the top two and discards all the others for US Presidential elections - but we think that goes too far. Using three-party-forced also gets us out of a rather thorny problem with the Liberal party in the 1980's – namely, should we assess against the Liberal share or the share of their electoral alliance.

5. THRESHOLD

Here's a question: how good is good enough? How bad does a poll have to get before we say it's too bad? What is our threshold?

An academic will assess a poll to see if it is better or worse than another poll. To do this they will compare polls across many years and different countries. These are *relative* comparisons, and a threshold is not necessarily useful for relative comparisons

A pollster will assess a poll to try and make it better. To do this they will compare the poll to their competitors in the present, and to the last election and possibly the one before in the past. Unlike academics a pollster will not assess their poll against the far past, because the salience and comparability of past elections grows less as time passes[IT1]. But again, these comparisons are relative comparisons and a rigid threshold is not useful.

The time when thresholds really come into their own is during a polling failure. The polls are deemed to have failed if the majority of polls predicted the wrong winner, so that is the threshold we will use for WIN. During a polling failure the threshold for an acceptable MAE on a four-party-forced basis - CON/LAB/LIB/OTH - is 2% or less, which matches comments by academics[IJ1] and others[0510e].

Problem is, as we have noted in previous articles, MAEs vary as the number of parties change, so that threshold won't work for three-party-forced. After some consideration, we settled on 2.7% or less as our threshold for an acceptable three-party-forced poll. This rather awkward number is a result of the fact that in past years poll shares were published to one decimal point – see Appendix 2 for a discussion of this.

So we will use 2.7% or less as our threshold for MAE three-party-forced. For WIN, it will be simpler: did a majority of the polls predict the correct winner?

6. ELECTION RESULTS

A question that is not often considered is the election results. There are two problems here. Firstly, for reasons of history and differences in polities, polls are usually for Great Britain only and are correctly judged against the results for Great Britain, ignoring Northern Ireland. But the general public may and do[0512a] judge against the results for the United Kingdom, including Northern Ireland.

Secondly, the results themselves. Results on the night are stated by the Returning Officer for each constituency and various organisations keep a running total of votes. Over the next day small piles of votes may be discovered or totals change, and the official total for a given constituency may change very slightly. A constituency election may be delayed by a few days due to the death of a candidate[0512b] or later set aside[0512c]. After some weeks or months academics will collate totals and the Electoral Commission and the House of Commons Library will publish the final results. All these factors mean that sources may disagree slightly on what the final results were: for example, was the Conservative GB vote in 1997 31.4%[0512d] or 31.5%[0512e]? Since our mission is to assess predictiveness, should we judge against the results on the night or the final results, given that these will be very slightly different?

For consistency's sake we will measure against GB results. For logistical reasons we will use the House of Commons briefing paper CBP-7529[0512e] to 1 decimal place (eg 37.7%) as the final results and try to note where sources differ.

7. GLOBAL ACCURACY

Let's start big. All of polls, everywhere and everywhen. How are they doing?

As mentioned above, this article cannot cover the whole of polling across time and space. Fortunately, we knew some people who could. On 12 March 2018 Nature published online the paper "Election polling errors across time and space"[0421f] by Jennings and Wlezien, the latest in a series of articles by the same authors and others[0421d][0421e].

The paper looked at 30,916 polls from 1942 to 2017, covering 351 elections in 45 countries. They examined annual MAE and log-odds ratio between 1942 and 2017, MAE per day from t-minus 200 days to t-minus zero, MAE for large parties in recent elections 2015–2017, and modelled MAE over the last week by election type (presidential/legislative), electoral system, effective number of parties, party size and incumbency. The paper selected the two largest parties for MAE post-facto which affected predictiveness, but the authors compensated by selecting the three largest parties when the pair with the highest votes wasn't the pair with the highest poll share.

The paper found that on a global basis:

Figure 1: copy of figure 2b from "Election polling errors across time and space" (2018)
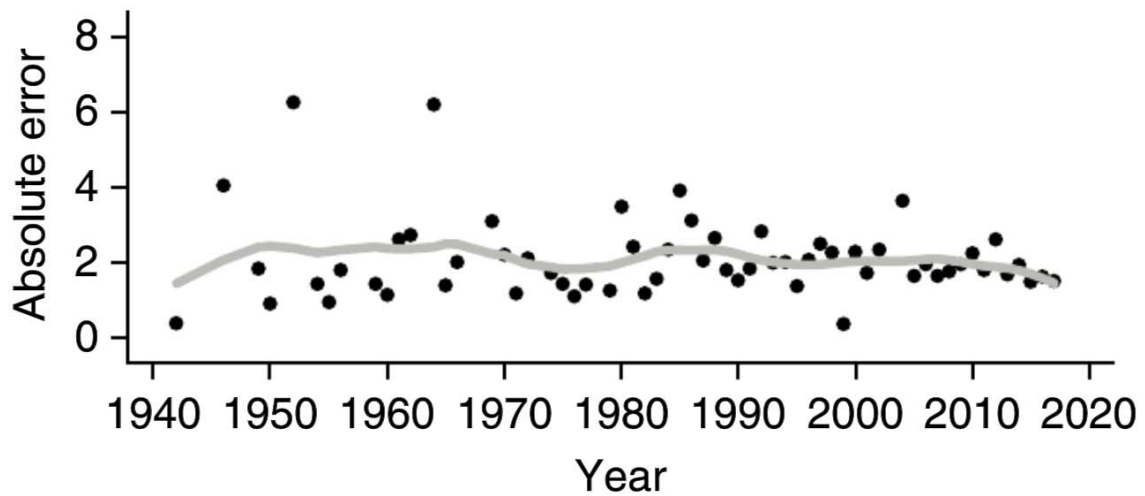
For original, see doi:10.1038/s41562-018-0315-6. The mean absolute error across all parties/candidates and elections in a given year is indicated with a black circle

Figure 2: copy of figure 3 from "Election polling errors across time and space" (2018)

MAE large parties (1942–2014) = 2.3
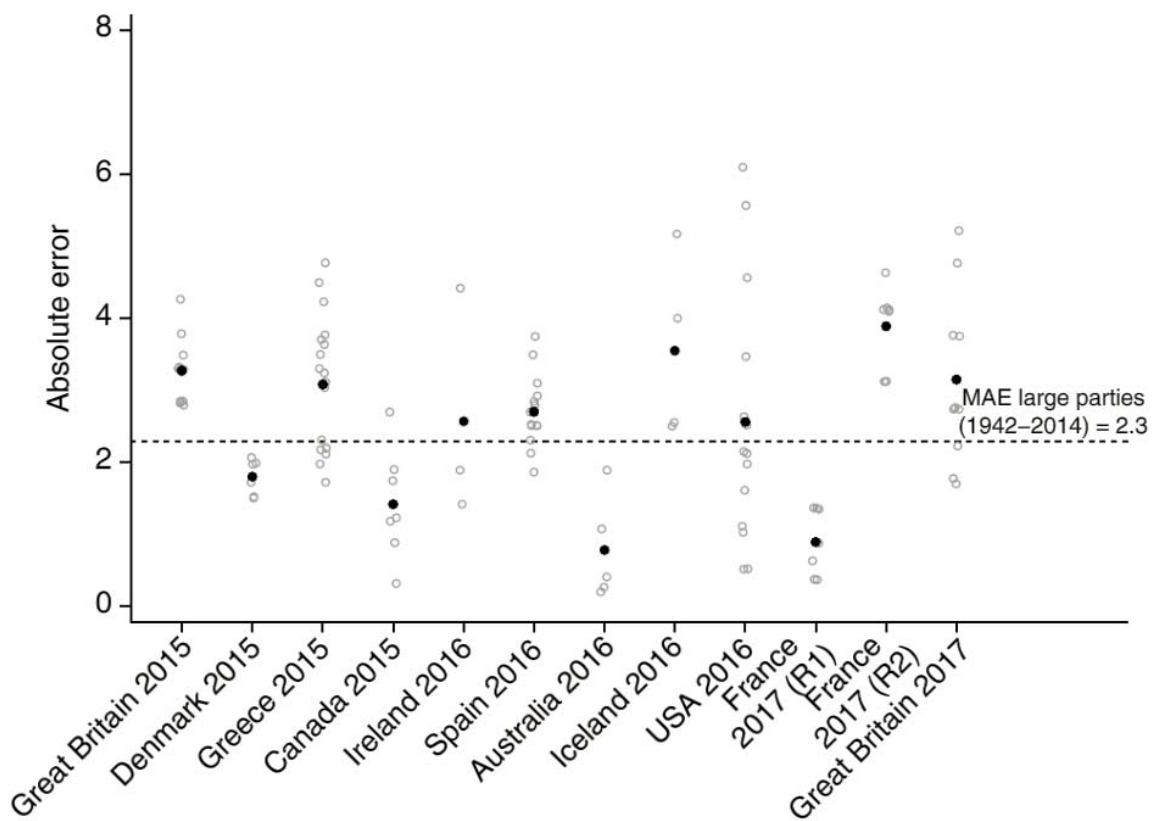
For original, see doi:10.1038/s41562-018-0315-6. The dashed black line is the mean absolute error (MAE) for 'large' parties over the 1942–2014 period. The black circles indicate the mean absolute error across polls for each election. The grey unfilled circles indicate the absolute error of individual pollsters

* There is no evidence that poll errors have increased over time, and the performance of polls in very recent elections is not exceptional.
* Larger parties have larger polling errors than smaller parties, and this is important when elections are close as it affects government formation
* Errors are lower in PR systems
* Errors are also lower for presidential elections (at least in the United States and France) compared to legislative elections

So, in terms of global accuracy, polls are accurate and aren't getting worse.

8. UK ACCURACY BY ELECTION

But macroaccuracy cannot be the whole picture: on average the world is flat and life expectancy is rising globally, but mountains and valleys still exist and some people die young. So, we must again restrict our scope and focus on the United Kingdom: how are polls doing in the UK?

Again, we knew some people who could answer this. On 31 March 2016 the British Polling Council and the Market Research Society published the NCRM report of the inquiry into the 2015 British General Election polls. That inquiry contained a graph of the average MAE for the final election polls for each year from 1945 to 2015. We created our own version from 1945 to 2017 and joined the dots. The resultant graphs are given below:

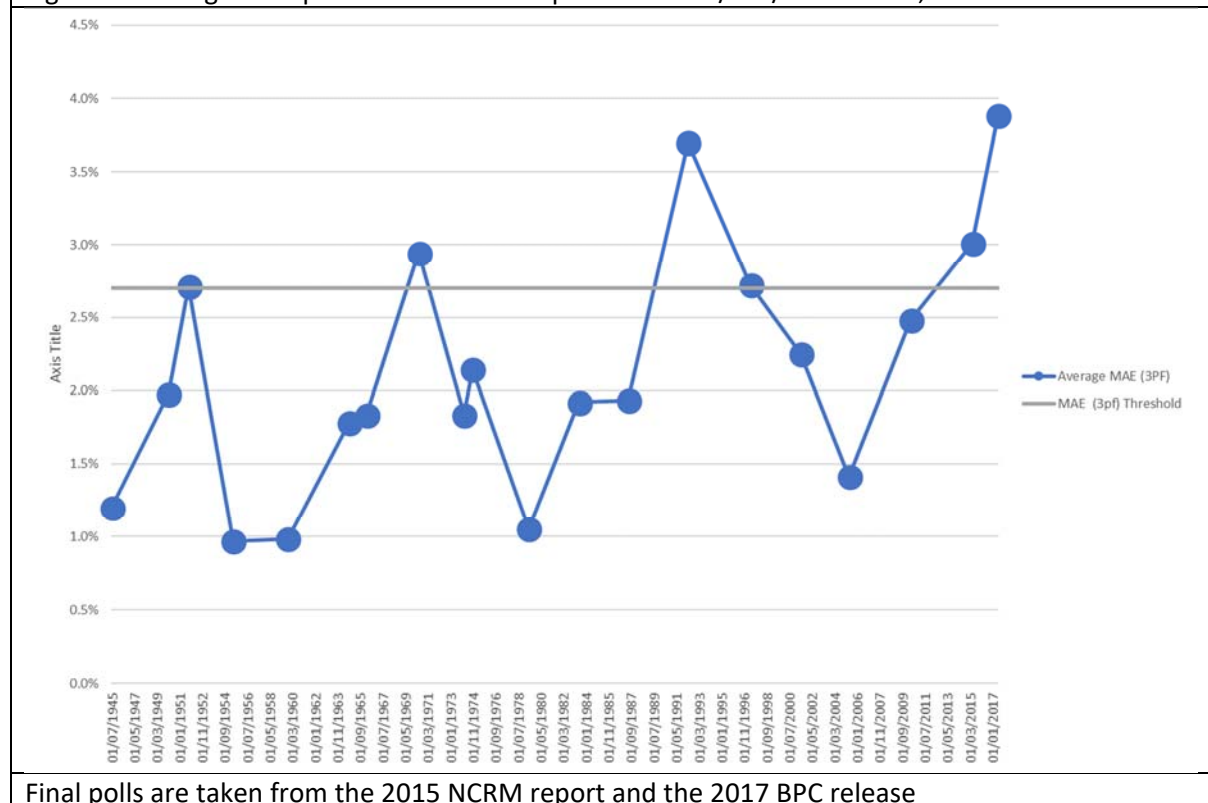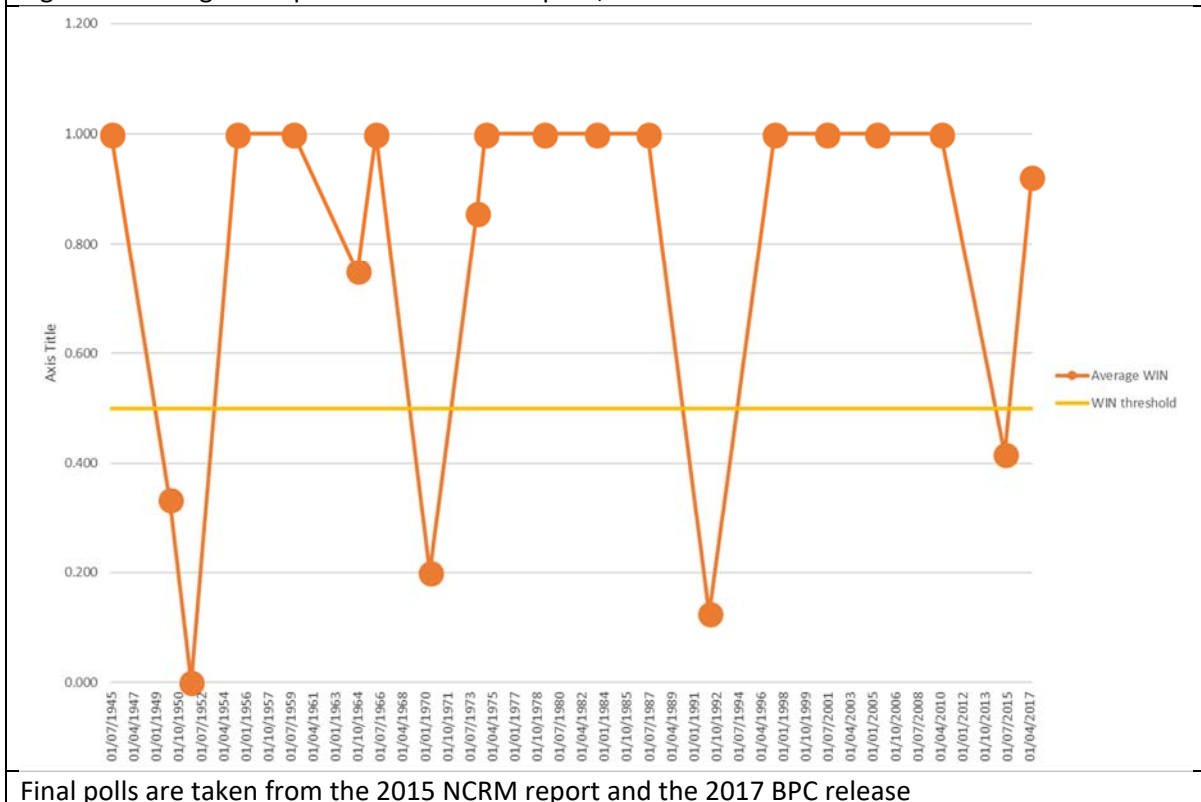Figure 3: Average MAE per election for final polls on a Con/Lab/Other basis, 2.7% threshold

Final polls are taken from the 2015 NCRM report and the 2017 BPC release

Figure 4: Average WIN per election for final polls, 0.5% threshold

Final polls are taken from the 2015 NCRM report and the 2017 BPC release

The data appear cyclical, with the cycles being 37-51, 52-70, 71-92, 93-17. But is this a real phenomenon? To examine this, we looked at the history of political opinion polls in the UK below.

9. UK POLLING HISTORY

1937-1951

On 1 January 1937 the UK spinoff of Gallup's American Institute of Public Opinion, inevitably entitled the British Institute of Public Opinion, was opened by Dr Henry Dunant. That autumn it asked its first voting intention question "For whom would you vote today", and the results were reported in the October 1937 edition of the weekly news magazine "Cavalcade" thus: "For the government, 68 per cent; for the Opposition, 32 per cent".[0422c] It polled on the Fulham West by-election of 6 April 1938[0422b] and switched from "Cavalcade" to the daily "News Chronicle" in October 1938[0422b], where it stayed until 1960[0422c].

Gallup/BIPO continued to poll intermittently thru the war years, and in 1946 the Market Research Society was founded in the offices of the London Press Exchange[0424a]. Then in 1948 the US presidential polls failed: a predicted win for the Republican candidate Dewey became an actual win for Democratic candidate President Truman. Following this the US Social Science Research Council organized a "Committee on Polls and Election Forecasts" and a conference entitled "Polls and Public Opinion" was held to discuss it in February 1949 at the University of Iowa.

But any schadenfreude the British pollsters might have felt at the US failure vanished in 1951 when the first notable British polling failure occurred. The lead between the top two parties according to the polls was wrong by over 5%[0423a]. Three entities polled for that election, Dunant's Gallup/BIPO for the News Chronicle, Research Services Ltd under Dr Mark Abrams for the Daily Graphic, and the Daily Express.

Final polls 1951	Con	Lab	Other	Source
RSL	50	43	7	[0423b page 16][0430a]
Gallup	49.5	47	3.5	[0423b page 16][0430a]
Daily Express	50	46`	4	[0423b page 16][0430a]
Result GB	47.8	49.4	2.8	[0423c][0512e]
Result UK	48	48.8	3.2	[0423b, table 3][0423d][0512e]

These were thought to be poor predictors of the result, with RSL's predicted 7% Conservative lead a particularly egregious prediction of the small Labour lead.[0423b, page 16]. Reasons proposed for this were not polling to the last day - the first occurrence of the perennial "late swing" - and failing to consider that the Liberals weren't standing in many seats.[0423b, page 17]

1952-1970

The years wore on. BIPO changed its name to "Social Surveys (Gallup Poll) Ltd" in 1952. A new crop of pollsters started to join in and by the late 60's the polling landscape was recognisable to older statisticians: Gallup, Harris, Marplan under Robert Worcester, NOP chaired by Mick Shields, ORC.[0423e] Then in the 1970 the opinion polls failed again: the lead was wrong by over 6%[0423a].

Final polls 1970	Con	Lab	Other	Source
Marplan	41.5	50.2	8.3	[0423e][but note 0423n][0430a]
NOP	44.1	48.1	7.7	[0430a]
ORC	46.5	45.5	8	[0423e][but note 0423n][0430a]
Gallup	42	49	9	[0423e][but note 0423n][0430a]
Harris	46	48	6	[0423e][but note 0423n][0430a]
Result GB	46.2	43.8	10	[0423e, page 179][0423c]
Result UK	46.4	43	11	[0512e]

Suspects this time included differential turnout[0423e] and - again - a late swing[0423e][0423b]

1971-1992

The Seventies turned into the Eighties and then the early Nineties. Things seemed to have improved: in all five elections since the 1970 debacle the average error on each party share across all the final polls were well below 2%[0432b, page 106]. The old pollsters were still there but new pretenders had popped up: MORI in the 70's, ASL and G9000 in the 80's, Neilsen in the 90's[0423c] Phone polling had begun to appear. And in 1992 the nation sat to watch the coverage of the 1992 General Election. On the morning of election day four companies had published polls which, on average, put Labour one percentage point ahead[0423g][0423h]. Would Labour be in the lead, like the polls said on Red Wednesday[0423f]?

No. No, it wouldn't.

By the end of the night it became apparent that there had been another polling failure[0430b]: the poll lead was wrong by over 9%[0423a] and the Conservatives had won by eight percentage points.[0423g][0423h]

Final polls 1992	Con	Lab	Other	Source
ASL	35	38	23	[0423c][0423b, page 106]
Harris	38	40	18	[0423c][0423b, page 106]
MORI	38	39	20	[0423c][0423b, page 106]
NOP	39	42	17	[0423c][0423b, page 106]
Gallup	38.5	38	20	[0423c][0423b, page 106]
ICM	38	38	20	[0423c][0423b, page 106]
Result GB	42.8	35.2	18.3	[0423b, page 106][0423i][0423c]
Result UK	41.9	34.4	17.8	[0423j][0512e]

The Market Research Society conducted a review which blamed differential non-response (aka "shy Tories")[0423k][0423l][0423a][0423m], unrepresentative quota samples[0423l] and – yes, yet again - late swing[0423k][0423l][0423a][0430b].

1993-2015
After 1992 pollsters reacted by improving quotas and weightings[0423a] to make the sample more representative[0423g], weighting by past vote[0423a], changing the way they reallocated don't knows[0423n], and in some cases moving to telephone polling[0423g][0423m], following the urging of (amongst others) the younger John Curtice.

ASL stopped polling voting intention soon after, and G9000 in May 1998, but pollsters came in to plug the gaps and more: Angus Reid, ASL, BMG, BPIX, CommR, ComRes, G9000, Gallup, Harris, ICM, Live St, Lord Ashcroft, MFS, Markting Sciences, MORI, NOP, Onepoll, Opinium, ORB, Panelbase, Populus, Rasmussen, RNB, Survation, SurveyMonkey, TNS-BMRB, YouGov, the list kept growing. Some of them stayed only briefly, some stayed for the long term, and in 2004 some of them formed the British Polling Council.[0424b] As the 2015 General Election rolled round, the polling landscape seemed settled: so many pollsters, so much theory, so much skill. The polls couldn't go wrong again, could they?

Yes. Yes they could.

Final polls 2015	Con	Lab	Other	Source
Ashcroft*4	33	33	34	[0423c][0430a]
BMG*4	34	34	32	[0423c][0430a]
ComRes	35	34	31	[0423c][0424f][0430a]
ICM	34	35	31	[0423c][0424f] [0430a]
Ipsos MORI	36	35	29	[0423c][0424f] [0430a]
Opinium	35	34	31	[0423c][0424f] [0430a]
Panelbase	31	33	36	[0423c][0424f] [0430a]
Populus	34	34	32	[0430a]*3
Survation	31	31	38	[0506d][0424f][0430a]*1
SurveyMonkey*4	34	28	38	[0430a]
TNS UK	33	32	35	[0430a]
YouGov	34	34	32	[0423c][0424f][0430a]
Result GB	37.7	31.2	31.1	[0423c][0512e]*2
Result UK	36.8	30.4	32.8	[0512e]

*1 [0423c] has different figures for this
*2 [0424f] has 37.8 for the Conservatives
*3 [0423c] and [0424f] have different figures for this
*4 Non-BPC. BMG joined BPC in February 2016

The public reaction to the failure of the polls was derisory and the Market Research Society and the British Polling Council announced a review, usually known as the "Sturgis inquiry". The report indicted unrepresentative samples as the prime cause of the failure caused by too many young people and people under the age of 70, and not enough aged 75 and older. It recommended that the pollsters make their samples more representative and suggested some procedural improvements. [0424c]

2015-2017

The period was marked by a flurry of activity: the 2015 UK General Election was followed by the 2016 UK EU referendum and the 2016 US POTUS election, and the polls for each were thought to be inadequate in some way. There was another review, this time the Kennedy review in the USA. The pollsters tried different techniques: some concentrated on making their samples more representative, others tried compensating with differential turnout models. The Referendum pitted online panels against telephone polling, and the latter lost. The pollsters took note and by 2017 had mostly moved to online panels. And then the Prime Minister announced a surprise General Election.[0424c] So on 8 June 2017 we voted again.

Final polls 2017	CON	LAB	Other	Source
BMG	46	33	21	[0424d]
ComRes	44	34	22	[0424d]
ICM	46	34	20	[0424d]
Ipsos MORI	44	36	20	[0424d]
Kantar Public	43	38	19	[0424d]
Opinium	43	36	21	[0424d]
Panelbase	44	36	20	[0424d]
Survation	41	40	19	[0424d]
YouGov	42	35	23	[0424d]
Norstat*4	39	35	26	[0423c]
ORB*5	45	36	19	[0423c]
Qriously*4	39	41	20	[0423c]
SurveyMonkey*4	42	38	20	[0423c]
Result GB	43.4	41	15.6	[0424e][0512e]*1
Result UK	42.3	40	17.7	[0424e][0512e]*2

* 1 [0424d] gives 43.5 and 41 for GB
* 2 [0423c] gives 42.4 and 40 for UK
* 4 Non-BPC. BMG joined BPC in February 2016
* 5 ORB is a BPC member but BPC did not include it as a final poll, see [0424d]

The predictiveness was good for the Conservatives, not so good for Labour. As the post-2015 improvements were still being implemented, the BPC declined to call another inquiry[0424e]. But the House of Lords did hold an enquiry of their own, which we will discuss later.[0424h]

SUMMARY

The history of political opinion polls in the UK upholds the NCRM graph and our supposition that UK poll accuracy over elections is cyclical. Periodically since 1945 there have been UK polling failures in 1951, 1970, 1992, 2015. Each time the pollsters have generated new techniques or revised their assumptions. The MAE for succeeding elections are lower, but eventually begin to rise again until they become unacceptable, and the cycle begins again.

10. CURRENT UK ACCURACY

So we have considered papers about global poll accuracy and looked at the cyclical nature of British polling accuracy over the years. So now we need to ask: how are polls doing in the UK right now? This article was written in 2018, so we will define "current accuracy" as the accuracy of the 2015 and 2017 elections

2015

Firstly, let's have a look at 2015, the final polls of which are listed below. Unpublished polls are unpredictive so the unpublished Survation poll of May 8 2015[0505c] is not included here and the Survation poll published on May 7 2015[0505d] is used instead. ICM released an early[0505b] "interim" version of its final poll: we have ignored it in favour of its actual final poll. The non-BPC Survey Monkey, BMG and Ashcroft polls are included for completeness.

Final polls 2015	Published	CON	LAB	OTH	Source	MAE GB	MAE UK	WIN
ComRes	07/05/2015	0.35	0.34	0.31	[0423c][0424f][0430a]	0.0187	0.0240	1
ICM (final)	07/05/2015	0.34	0.35	0.31	[0423c][0424f] [0430a]	0.0253	0.0307	0
Ipsos-Mori	07/05/2015	0.36	0.35	0.29	[0423c][0424f] [0430a]	0.0253	0.0307	1
Opinium	06/05/2015	0.35	0.34	0.31	[0423c][0424f] [0430a]	0.0187	0.0240	1
Panelbase	06/05/2015	0.31	0.33	0.36	[0423c][0424f] [0430a]	0.0447	0.0387	0
Populus	07/05/2015	0.34	0.34	0.32	[0430a]*3	0.0247	0.0240	0
Survation	07/05/2015	0.31	0.31	0.38	[0506d][0424f][0430a]*1	0.0460	0.0387	0
TNS_BMRB	06/05/2015	0.33	0.32	0.35	[0430a]	0.0313	0.0253	1
YouGov	07/05/2015	0.34	0.34	0.32	[0423c][0424f][0430a]	0.0247	0.0240	0
Ashcroft*4	07/05/2015	0.33	0.33	0.34	[0423c][0430a]	0.0313	0.0253	0
BMG*4	05/05/2015	0.34	0.34	0.32	[0423c][0430a]	0.0247	0.0240	0
SurveyMonkey*4	06/05/2015	0.34	0.28	0.38	[0430a]	0.0460	0.0347	1
GB result	07/05/2015	0.377	0.312	0.311	[0423c][0512e]*2	**Avg GB MAE**	**Avg UK MAE**	**Avg WIN**
UK result	07/05/2015	0.368	0.304	0.328	[0512e]			
					BPC members	**0.0288**	**0.0289**	**0.4444**
					Non-BPC	**0.0340**	**0.028**	**0.3333**
					All	**0.0301**	**0.0287**	**0.4167**

*1 [0423c] has different figures for this
*2 [0424f] has 37.8 for the Conservatives
*3 [0423c] and [0424f] have different figures for this
*4 Non-BPC. BMG joined BPC in February 2016

2017

Now let's look at 2017. The non-BPC Norstat, Qriously and SurveyMonkey polls are included for completeness. The BPC omitted the ORB poll of 2017-06-04 from their analysis but we have included it.

Final polls 2017	Published	CON	LAB	OTH	Source	MAE GB	MAE UK	WIN
BMG	08/06/2017	0.46	0.33	0.21	[0424d]	0.0533	0.0467	1
ComRes	07/06/2017	0.44	0.34	0.22	[0424d]	0.0467	0.0400	1
ICM	07/06/2017	0.46	0.34	0.2	[0424d]	0.0467	0.0400	1
Ipsos MORI	08/06/2017	0.44	0.36	0.2	[0424d]	0.0333	0.0267	1
Kantar Public	07/06/2017	0.43	0.38	0.19	[0424d]	0.0227	0.0133	1
Opinium	06/06/2017	0.43	0.36	0.21	[0424d]	0.0360	0.0267	1
Panelbase	07/06/2017	0.44	0.36	0.2	[0424d]	0.0333	0.0267	1
Survation	07/06/2017	0.41	0.4	0.19	[0424d]	0.0227	0.0087	1
YouGov	07/06/2017	0.42	0.35	0.23	[0424d]	0.0493	0.0353	1
Norstat*4	04/06/2017	0.39	0.35	0.26	[0423c]	0.0693	0.0553	1
ORB*5	04/06/2017	0.45	0.36	0.19	[0423c]	0.0333	0.0267	1
Qriously*4	07/06/2017	0.39	0.41	0.2	[0423c]	0.0293	0.0220	0
SurveyMonkey*4	07/06/2017	0.42	0.38	0.2	[0423c]	0.0293	0.0153	1
GB result	08/06/2017	0.434	0.41	0.156	[0424e][0512e]*1	**Avg GB**	**Avg UK**	**Avg**
UK result	08/06/2017	0.423	0.4	0.177	[0424e][0512e]*2	**MAE**	**MAE**	**WIN**
					BPC members	**0.0382**	**0.0293**	**1.0000**
					ORB*5 & Non-BPC	**0.0403**	**0.0298**	**0.7500**
					All	**0.0389**	**0.0295**	**0.9231**

* 1 [0424d] gives 43.5 and 41 for GB
* 2 [0423c] gives 42.4 and 40 for UK
* 4 Non-BPC. BMG joined BPC in February 2016
* 5 ORB is a BPC member but BPC did not include it as a final poll, see [0424d]

These are not good numbers. and the MAE is over our threshold in both cases. However note the WIN indicator: in 2015 a majority of polls did not predict the winner, but in 2017 they did.

PUBLIC PERCEPTION

Now we need to ask: how were these polls received by the general public? Following the 2015 election, general opinion was that the polls had failed. This opinion was not restricted to academics[0506d][0506e][0506f] and pollsters,[0506i] it was shared by newspapers,[0506k] the media,[0506c][0506l] data scientists,[0506g] the commentariat,[0506a][0506j] election strategists[0506h] and celebrities.[0506g] Following the 2017 election reactions were more diffuse, with YouGov[0506m][0506n] and Survation[0506m] being thought of as good, but nobody thought the 2017 polls were a roaring success.[0506q][0609a]

But by then pressure had built up. Even before the 2015 election regulation of polling had been proposed[0506b] and following 2017 it was proposed again.[0506r] At that point the House of Lords had had enough and announced a committee to investigate.[0506o]. That committee reported on 17 April 2018.[0506q] It declined regulation and it agreed with Jennings that polls had not become more inaccurate over time, but nevertheless it thought that for various reasons polls may have recently taken a turn for the worse and may continue to deteriorate.

So the perception of British opinion polls is that they are not accurate and may get worse.

11.   UK ACCURACY BY TIME TO POLL

So we've looked at the final election polls, but they are still just the final election polls. How about predicting twenty days before? A hundred days? Two hundred days? A year? The Jennings and Wleizen paper had done this globally, but what about UK-only?

To examine this, we used the Mark Pack PollBase. Mark Pack has a spreadsheet of British General Election opinion polls since 1945. It's rather comprehensive but it has one interesting flaw. If we want to measure predictiveness instead of accuracy, we must only use information available to the observer at the time. A layman only becomes aware of a poll when it is published...but Pack's spreadsheet didn't have publishing dates for all the polls. So, we will use publishing dates when available, and the fieldwork or month date when not.

So let's look at the graphs. We'll look at the cycles - 37-51, 52-70, 71-92, 93-17 - and because of recent concerns we'll split the latter cycle into 93-15 and 15-17. The results looked like this:

Figure 5: Average MAE by time to election on a Con/Lab/Other basis, 2.75% threshold or less
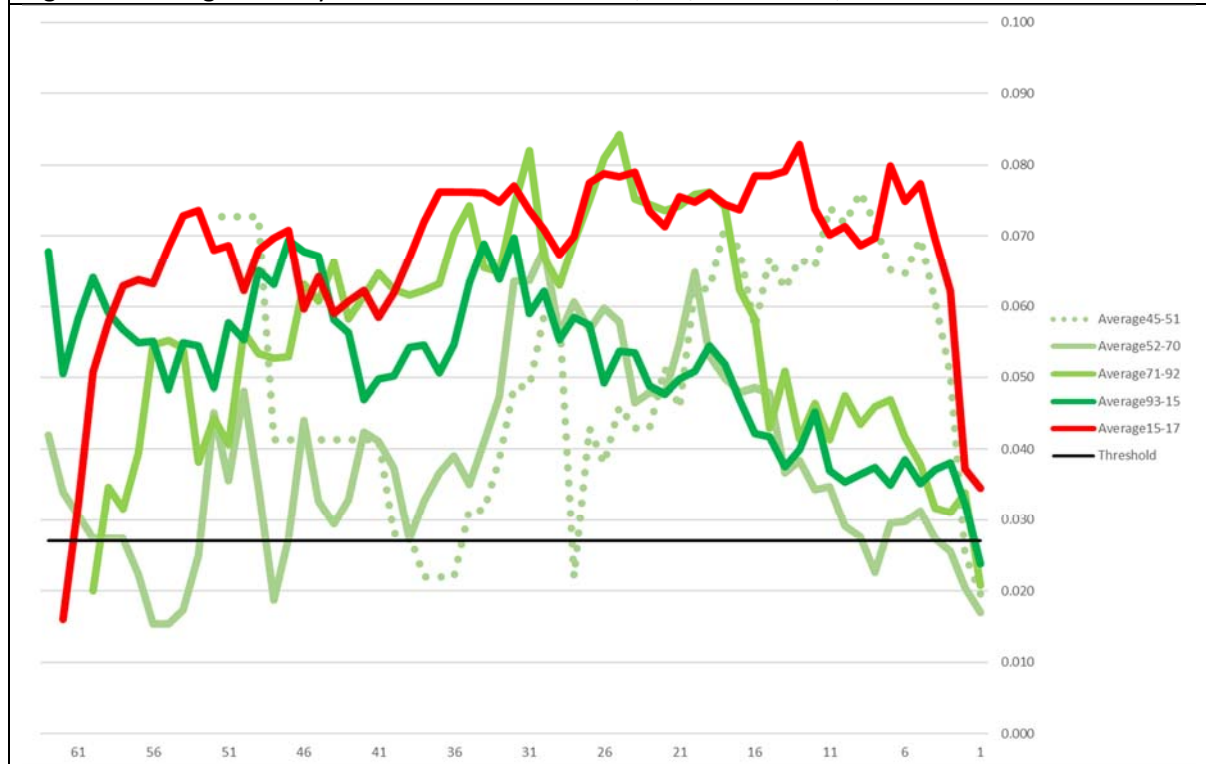
The time axis is 30-day chunks, 2015-2017 in red, the threshold is in black

Figure 6: Average WIN by time to election on a Con/Lab/Other basis, 0.5% threshold or more

The time axis is 30-day chunks, 2015-2017 in red, the threshold is in black
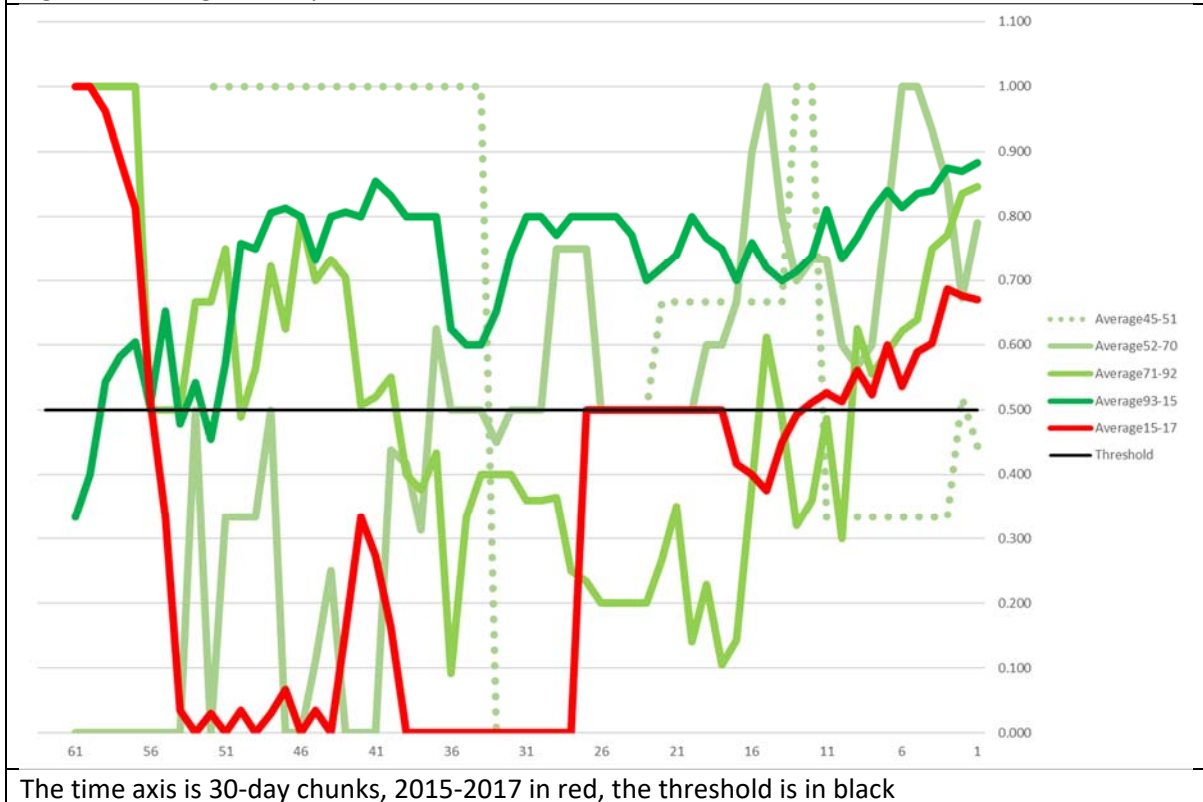
On average the polls start to stabilize about a year out, then become acceptable 30-60 days before the election. The exception is 2017, which had a high error until the campaign started then the error decreased rapidly. This is presumably because 2017 was a surprise snap election and were thought to be mid-term polls until the election was called. But even under those conditions the polls still passed the threshold for predicting the winner.

12.   INVESTIGATIVE TOOLS

Looking at opinion poll data is simple, but are there any investigative tools we can apply to improve predictiveness?

Firstly, house or mode fixed effects. There is always the urge to believe in a "gold standard", the fabled indicator that is better at predicting an election than the other indicators and does so consistently across elections. Unfortunately there just isn't one. Academics[IJ1] and pollsters insist that there is no such thing as a consistent fixed effect. Poll modes are dictated by economics as well as accuracy and mode effects may simply not be consistent from election to election. Pollster rankings are not constant as old pollsters leave, new pollsters join, and new techniques are adopted. ICM were one of the best in 2010[0501c] and one of the worst in 2017[0501d].

So there's no such thing as a consistent fixed effect. But how about an *inconsistent* fixed effect? During an election, a fixed effect may bubble to the surface, as ephemeral as a mayfly and as delicate as a soap bubble, that wasn't there in the last election but is undeniably here now. What can be done with such a transitory fixed effect?

This isn't a hypothetical point, it really happens and here's an example. During the 2016 EU Referendum it became apparent that a mode effect was present: the polls based on online panels said the race was close, but the polls based on telephone calling showed a consistent REMAIN lead.

One of them was presumably wrong, but people could not agree on *which* one and some just plumped for half-way between the two. As it turned out the polls based on online panels were closer to the truth and if the polls based on telephone polling had been discarded then the predictiveness would have been improved.

That example is famous but there are others and a similar situation arose during the 2017 General Election. As the campaign progressed it was hypothesised by Sturgis and Jennings (amongst others)[0606a][0606b] that there was a mode effect: polls using turnout models had higher Conservative leads than those using self-reported. If those polls that used turnout models had been discarded, then predictiveness would have improved. The graph is given below, using The Polling Observatory's categories

Figure 7: Graph depicting average poll lead from 1/12/2016 to 8/6/2017.
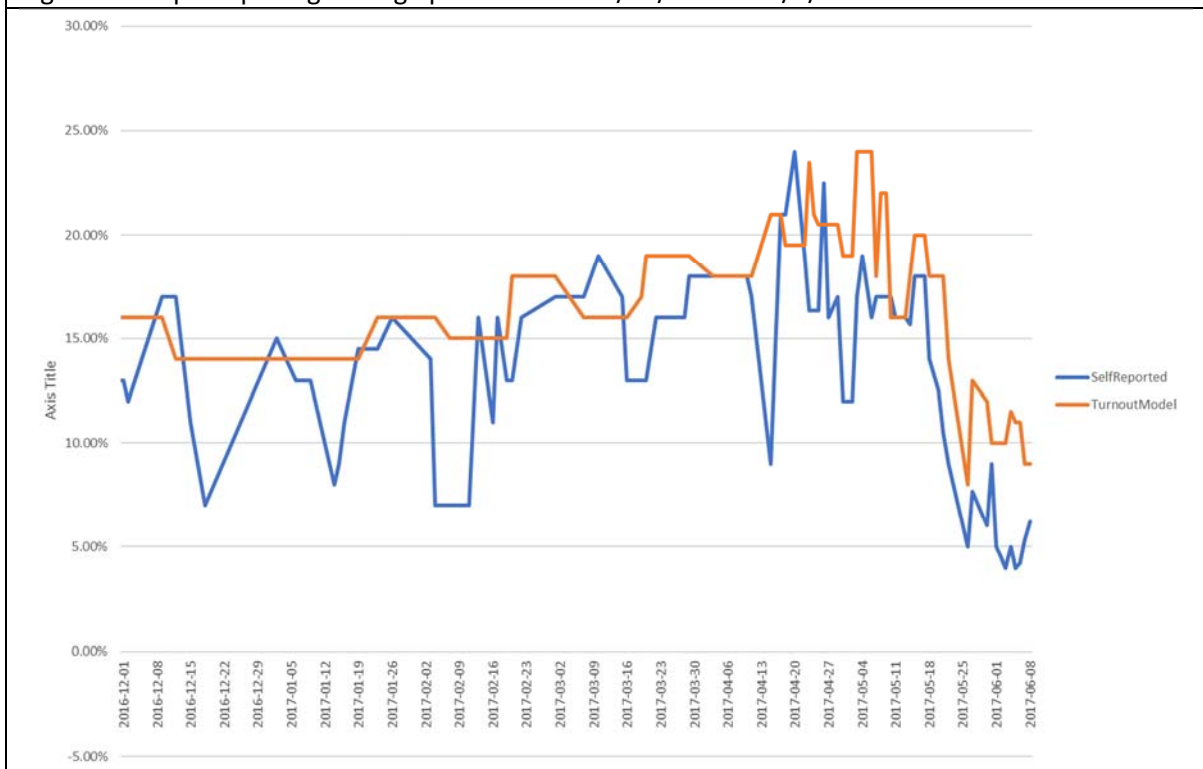
Days with no polls inherit value of latest prior. Blue=Self reported, orange=Turnout Model

This is all well and good, but can we do anything useful with this? Possibly, but it's a risk. If a transitory fixed effect is found to exist, there is no obvious way of telling which mode is correct. Even worse, it might simply be an artefact: the more you look for a pattern, the more likely you are to find one, so this might be spurious instead of a real phenomenon. So it's a triple-edged sword. Jennings urges caution and notes that they are only useful insofar as they inform the methodology.[IJ1]

Another possibility is leader ratings. Some modellers use leader ratings instead of voter intention polls to predict election outcomes, so the question arises whether they can be used to predict a polling failure. This will be the subject of later work.

13. FUTURE DEVELOPMENT

Currently the polling industry is evolving in a Big Data direction. The increased availability of datasets and the tools to link them together make it inevitable that someone will do so. New techniques can

be discovered via the internet, rapidly applied via easily-downloadable software, and used with minimal delay and cost.

The increased use of online panel polling, the popularity of the open-source R language and the availability of open-source R-compatible modelling functions makes the increased use of models irresistible. One such technique is MRP (Multi-level Regression and Poststratification). This takes the results of an online panel, infers from them the voting behaviour of specific characteristics such as age, gender and area type, and uses that to predict how each individual constituency will vote. This technique was applied to good effect by YouGov in the 2017 General Election, though less well by Lord Ashcroft at the same time.

One problem with innovation is that it tends to come from outside politics. The driver for innovation comes from the interaction between the research organisation and the clients, as the clients will seek out new techniques that they believe to be good and the research organisation will generate and offer techniques in the hope of attracting them.[IW1] Problem is, political pollsters in the UK tend to be smaller subdivisions of larger market research organizations (MRO) such as YouGov or Kantar, or entities with specialised funding. Those organisations will focus on their higher-value clients and in the UK such high-value clients will be commercial clients: the clients who want to discuss political opinion will be the media, academics and political parties, and they tend to have less money than the commercial clients or prefer simpler narratives to complex techniques.

Consequently, political pollsters in larger firms may be constrained by the techniques and concerns of the commercial parent, instead of generating innovations of their own. Since the commercial pollster and the political pollster may be the same people in different rooms, this is not necessarily a problem (and there is a tendency for the firms to use political polling as a loss-leader or "shop window"), but a smaller or hungrier firm may be more flexible. A recent example of such flexibility is Matt Singh's use of river sampling[0430d], a sample where respondents are invited via the placement of ads, offers or invites online.[0430c]

An ongoing development is the move towards expressing polls as ranges instead of single datapoints. On 1 May 2018[0506t] the BPC announced that it now requires its members to include this statement in its polls: "…*All polls are subject to a wide range of potential sources of error. On the basis of the historical record of the polls at recent general elections, there is a 9 in 10 chance that the true value of a party's support lies within 4 points of the estimates provided by this poll, and a 2 in 3 chance that they lie within 2 points…*" This sentence was derived from a study of the final polls for each General Election since 2001 and is a practical solution to the problem of calculating confidence intervals for nonprobability sampling.

The UK has been lagging behind the US in addressing this problem, as indicated by this sentence from a US poll in November 2016: "*...Because the sample is based on those who initially self-selected for participation rather than a probability sample, no estimates of sampling error can be calculated. All surveys may be subject to multiple sources of error, including, but not limited to sampling error, coverage error, and measurement error…* ".[0506s] So this change is a welcome uplift of UK polling to US standards, although it is not as sophisticated in that regard.

14.   CONCLUSIONS

When they discussed the polls the Lords were much taken with recentism: they acknowledged Jennings's thesis that polls were not getting worse globally but were still convinced that things had gone wrong recently and that things might get worse. As Sturgis points out, they may be correct: polls are currently not good and there is no *a priori* reason why they should get better.

But…we have been here before, several times. The current situation of inaccurate polls, soul-searching by the industry, calls for regulation and so on is not unique but has happened several times since 1945. There are *a posteriori* reasons for supposing polls will get better because the UK poll accuracy is cyclical.

Polling is subject to periodic failures and those failures are followed by experiment, the development of new techniques and updating of assumptions. Those adjustments cause the errors to decrease and the problem is thought to be solved. And due to human inertia that is true, at least for a while. But invariably assumptions will eventually date, the errors will creep back up and the cycle begins again. The cycle length is around twenty years from peak to peak and trough to trough: 51-70, 70-92, 92-2017.

We are currently in another trough. We may not get out of it but historically we have, and the signs of adjustment are present: the industry is experimenting, generating techniques and questioning assumptions, and hopefully this will lead to another upswing.

15. VERDICT

Polls occupy a paradoxical position. Strictly speaking they assess voter intention but are routinely judged by second-order effects like how many seats a party gets, or who forms the government. It is important to resist that temptation and judge polls purely against popular vote. With that caveat in place, the verdict looks like this.

In the twenty elections from 1945 to 2017, the polls were unacceptable for vote share on six occasions: 1951, 1970, 1992, 1997, 2015 and 2017. The threshold was a mean absolute error MAE of 2.7% or less on a three-party-forced basis: Con/Lab/Other.

In the twenty elections from 1945 to 2017, the polls were unacceptable for the winner of the popular vote on five occasions: 1950, 1951, 1970, 1992 and 2015. The threshold was at least fifty percent of the final polls predicting the winner of the majority vote.

Polls start to predict the winner of the popular vote about a year out, but do not meet our threshold for vote share until about thirty days out. Predictiveness is cyclical and we are currently in a trough from which we may or may not recover, but even in their present state polls still predict the winner of the popular vote.

16. ACKNOWLEDGEMENTS

17. REFERENCES

[0421]

- [0421a] "Good Judgement Open: Frequently Asked Questions (FAQ)", see https://www.gjopen.com/faq
- [0421b] "Superforecasting: The Art and Science of Prediction" by Philip Tetlock, Dan Gardner, ISBN-10: 1847947158, ISBN-13: 978-1847947154
- [0421c] "How Accurate Were the Political Science Forecasts of the 2016 Presidential Election?" James E. Campbell, November 17th, 2016, Sabato's Crystal Ball, University of Virginia Center for Politics, see http://www.centerforpolitics.org/crystalball/articles/how-accurate-were-the-political-science-forecasts-of-the-2016-presidential-election/
- [0421d] "The Timeline of Elections: A Comparative Perspective" Will Jennings, Christopher Wlezien: First published: 22 June 2015 https://doi.org/10.1111/ajps.12189 https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12189
- [0421e] "Polls and the Vote in Britain" Wlezien, Jennings, Fisher, Ford, Pickup: http://users.ox.ac.uk/~nuff0084/PollsandtheVoteinBritain.pdf doi: 10.1111/1467-9248.12008
- [0421f] "Election polling errors across time and space": Will Jennings & Christopher Wlezien: Nature Human Behaviour volume 2, pages 276–283 (2018), see doi:10.1038/s41562-018-0315-6: https://www.nature.com/articles/s41562-018-0315-6.epdf https://www.nature.com/articles/s41562-018-0315-6.epdf?author_access_token=Dv9fzB5mxgATeTZwsxLioNRgN0jAjWel9jnR3ZoTv0PalVG3evbr7Ah AkXUDTrGPuxQLUa5my1W6hESsmpNFdYsKXZoBT0m-zNSyqIupcNM3PbrEMvayUOKxPYc_acJq1F0WQDhCRL87YXwQ3q4FRg%3D%3D

[0422]

- [0422b] "Britain Speaks Out, 1937-87" by Robert J. Wybrow, ISBN-10: 0333396642 ISBN-13: 978-0333396643
- [0422c] "British Political Opinion, 1937-2000: The Gallup Polls" by Anthony King (Editor), Robert J. Wybrow (Editor), ISBN-10: 1902301889, ISBN-13: 978-1902301884

[0423]

- [0423a] "Getting Opinion Polls 'Right'" Postnote 96, March 1997, Parliamentary Office Of Science And Technology, see http://researchbriefings.files.parliament.uk/documents/POST-PN-96/POST-PN-96.pdf
- [0423b] "Opinion Polls: History, Theory and Practice" by Nick Moon, Manchester University Press, 1999, ISBN-10: 0719042240, ISBN-13: 978-0719042249
- [0423c] Mark Pack's opinion poll database, "PollBase-Q4-2017.xls" from https://www.markpack.org.uk/opinion-polls/
- [0423d] BBC Politics 97 website, past elections section, "25 October 1951", see http://www.bbc.co.uk/news/special/politics97/background/pastelec/ge51.shtml
- [0423e] The British General Election of 1970 By David Butler, Michael Pinto-Duschinsky, ISBN-10: 0333121422, ISBN-13: 978-0333121429
- [0423f] BBC Politics 97 website, past elections section, "Key Issues in the 1992 Campaign", http://www.bbc.co.uk/news/special/politics97/background/pastelec/92keyiss.shtml
- [0423g] "Why the polls got it wrong last time", The Economist, Mar 20th 1997, see https://www.economist.com/node/145916
- [0423h] "The Opinion Polls: Confidence Restored?", by Ivor Crewe, Parliamentary Affairs, Volume 50, Issue 4, 1 October 1997, Pages 569–585, see https://doi.org/10.1093/oxfordjournals.pa.a028756 or https://academic.oup.com/pa/article-abstract/50/4/569/1475511?redirectedFrom=PDF

- [0423i] "GENERAL ELECTION, 9th APRIL 1992", see
  http://www.election.demon.co.uk/ge1992.html
- [0423j] Commons Information Office, No 61, GENERAL ELECTION RESULTS, 9 APRIL 1992, see
  https://www.parliament.uk/documents/commons-information-office/m13.pdf
- [0423k] "Are the opinion polls ready for 1997", John Curtice, March 1, 1997, International
  Journal of Market Research, see
  http://journals.sagepub.com/doi/abs/10.1177/147078539703900206
- [0423l] "The Opinion Polls and the 1992 General Election: a report to the Market Research
  Society", July 1994, see
  https://www.ncrm.ac.uk/polling/documents/The%20Opinion%20Polls%20and%20the%201992%
  20General%20Election.pdf
- [0423m] "Editorial: The 1997 UK Pre-election Polls", Colm O'Muircheartaigh Peter Lynn, First
  published: 23 July 2007, RSS Series A, Volume160, Issue3, September 1997, Pages 381-385, see
  https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-985X.1997.00071.x
- [0423n] Note that Nick Moon gives the 1970 results as Marplan (42/50/7/1), NOP (44/48/6/1),
  ORC (46.5/45.5/6.5/1.5), Gallup (42.0/49.0/7.5/1.5) and Harris (46.0/48.0/5/1)

[0424]
- [0424a] "MRS: Our History", see https://www.mrs.org.uk/about/history
- [0424b] "On November 15th the British Polling Council (BPC) will be launched", 1 November
  2004, press release by BPC, see http://www.britishpollingcouncil.org/on-november-15th-the-
  british-polling-council-bpc-will-be-launched/ and http://www.britishpollingcouncil.org/british-
  polling-council-launched/
- [0424c] "What's happened to the polls since the 2015 UK election?" Timothy Martyn Hill,
  Significance, 15 May 2017, see https://www.significancemagazine.com/politics/549-what-s-
  happened-to-the-polls-since-the-2015-uk-election, or see
  https://www.significancemagazine.com/files/What-has-been-done-to--fix-the-polls---with-
  citations.pdf
- [0424d] "General Election: 8 June 2017", BPC, 9 June 2017, see
  http://www.britishpollingcouncil.org/general-election-8-june-2017/
- [0424e] "House of Commons Library BRIEFING PAPER Number CBP 7979, General Election 2017:
  results and analysis Second Edition", updated 3 Apr 2018, see
  http://researchbriefings.files.parliament.uk/documents/CBP-7979/CBP-7979.pdf
- [0424f] "General Election: 7 May 2015", BPC, 8 May 2015, see
  http://www.britishpollingcouncil.org/general-election-7-may-2015/
- [0424g] BBC Election 2015, Results webpage, see
  http://www.bbc.co.uk/news/election/2015/results
- [0424h] https://publications.parliament.uk/pa/ld201719/ldselect/ldppdm/106/10602.htm
- [0424i] Select Committee on Political Polling and Digital Media, "The politics of polling" Report of
  Session 2017-19 - published 17 April 2018 - HL Paper 106, see
  http://www.britishpollingcouncil.org/british-polling-council-welcomes-lords-committee-report-
  on-polling/

[0430]
- [0430a] Sturgis, P. Baker, N. Callegaro, M. Fisher, S. Green, J. Jennings, W. Kuha, J. Lauderdale, B.
  and Smith, P. (2016) "Report of the Inquiry into the 2015 British general election opinion polls",
  London: Market Research Society and British Polling Council, see
  http://eprints.ncrm.ac.uk/3789/1/Report_final_revised.pdf
- [0430b] "1970: Shock election win for Heath "
  http://news.bbc.co.uk/onthisday/hi/dates/stories/june/19/newsid_3829000/3829819.stm , see

also "Patrick Cosgrave on the election", see http://archive.spectator.co.uk/article/1st-may-1971/15/patrick-cosgrave-on-the-election
- [0430c] "ESOMAR/GRBN Guideline For Online Sample Quality Consultation", Draft April 2014, ESOMAR and GRBN. See http://www.grbn.org/userfiles/GRBN-ESOMAR_Online_Sample_Quality_Guideline_for_consultation_final_April_2014.pdf
- [0430d] https://www.ncpolitics.uk/2018/04/we-did-a-poll-heres-why-and-what-we-found.html/

[0501]
- [0501a] "Projecting confidence: How the probabilistic horse race confuses and demobilizes the public", Sean J. Westwood, Solomon Messing, and Yphtach Lelkes, Pew Research Center, Annenberg School of Communication, University of Pennsylvania February 6, 2018, SEE ALSO https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3117054 AND http://www.pewresearch.org/fact-tank/2018/02/06/use-of-election-forecasts-in-campaign-coverage-can-confuse-voters-and-may-lower-turnout/ AND https://www.nytimes.com/2017/12/24/opinion/2017-wrong-numbers.html?_r=0
- [0501b] "British Political Facts" By D. Butler, ISBN-10: 023025229X; ISBN-13: 978-0230252295, see https://books.google.co.uk/books?id=0l9BDAAAQBAJ&pg=PA297&lpg=PA297
- [0501c] "Predicting elections: who's the most accurate?", Timothy Martyn Hill, Significance, 09 April 2015, see https://www.statslife.org.uk/politics/2186-predicting-elections-who-s-the-most-accurate , see also https://www.statslife.org.uk/files/Predicting-Elections-Article-and-Appendices.pdf
- [0501d] "Forecast error: Predictors of the 2017 UK general election", Timothy Martyn Hill, Significance, 05 July 2017, see https://www.significancemagazine.com/politics/555-forecast-error-predictors-of-the-2017-uk-general-election , see also https://www.significancemagazine.com/files/GE2017-final-with-appendices.pdf

[0505]
- [0505a] "Conservatives heading for election win", LUKE TAYLOR, Head of Social and Political Attitudes, Kantar, 07.06.2017 / 21:00, see https://uk.kantar.com/ge2017/2017/conservatives-heading-for-election-win/
- [0505b] "The final four polls", 7 MAY 2015, Anthony Wells, "UK Polling Report", see http://ukpollingreport.co.uk/blog/archives/9410
- [0505c] The unpublished Survation poll of May 8 2015 can be found on http://survation.com/snatching-defeat-from-the-jaws-of-victory/ and http://survation.com/wp-content/uploads/2015/05/Final-Phone-Poll-2.pdf
- [0505d] The Survation poll published on May 7 2015 can be found on https://www.mirror.co.uk/news/uk-news/general-election-2015-britain-goes-5649629 and http://survation.com/wp-content/uploads/2015/05/Final-GE2015-Mirror-Poll-4c0d8h8.pdf

[0506]
- [0506a] "Why did the polls get it wrong at the general election? Because they lied", by Dan Hodges https://www.telegraph.co.uk/news/general-election-2015/politics-blog/11695816/Why-did-the-polls-get-it-wrong-at-the-general-election-Because-they-lied.html
- [0506b] "Pollsters would be curbed by peer's new regulation Bill", Mark Leftly (@MLeftly), The Independent, Sunday 11 January 2015 01:00 , see https://www.independent.co.uk/news/uk/politics/pollsters-would-be-curbed-by-peers-new-regulation-bill-9970427.html. See also https://www.huffingtonpost.co.uk/2015/02/05/ban-opinion-polls-during-election-campaign-says-labour-peer_n_6622300.html
- [0506c] "Election polls 2015: what went wrong?", unnamed author, "The Week", Jan 19, 2016, see http://www.theweek.co.uk/election-2015/62834/election-polls-2015-what-went-wrong

- [0506d] "Comment: The dark night of the polls: How did the polls get the 2015 election so wrong?" 22 May 2015, by Professor Charles Pattie and Dr Todd Hartman from the Sheffield Methods Institute at the University of Sheffield, see https://www.sheffield.ac.uk/news/nr/general-election-2015-polls-charles-pattie-todd-hartman-1.466666
- [0506e] "Representative samples are an issue for the pollsters – but so are respondents who lie", January 25 2016, Paul Whiteley (Professor in the Department of Government, University of Essex) and Harold D. Clarke (Professor in the School of Economic, Political and Policy Sciences, University of Texas), writing in the "British Politics and Policy" LSE blog, see http://blogs.lse.ac.uk/politicsandpolicy/the-2015-polling-debacle-the-british-polling-council-is-right-to-focus-on-problematic-samples-but-what-about-respondents-who-dont-tell-the-truth/
- [0506f] "Failed political predictions and the future of opinion polls", Thomas Loughran, Lecturer in Electoral Politics at the University of Manchester, December 5, 2016, see http://blog.policy.manchester.ac.uk/posts/2016/12/failed-political-predictions-and-the-future-of-opinion-polls/
- [0506g] "Election results 2015: How did the polls get it so wrong?", Caroline Preece, IT Pro, 8 May, 2015, see http://www.itpro.co.uk/public-sector/24580/election-results-2015-how-did-the-polls-get-it-so-wrong
- [0506h] "Lynton Crosby: Betrayal of British voters" Peter Dominiczak, Political Editor, The Telegraph, 10:00PM BST 15 May 2015, see https://www.telegraph.co.uk/news/general-election-2015/11608648/Lynton-Crosby-Betrayal-of-British-voters.html
- [0506i] "Why the polls were wrong about the Conservative campaign", By Rob Hayward, The Telegraph, 7:00PM BST 08 May 2015, see https://www.telegraph.co.uk/news/general-election-2015/11591779/Why-the-polls-were-wrong-about-the-Conservative-campaign.html
- [0506j] "Lord Ashcroft's polling overlooked many of the real election battlegrounds", Robert Smith, The Spectator, 15 May 2015 5:50 PM, see https://blogs.spectator.co.uk/2015/05/lord-ashcrofts-polling-overlooked-many-of-the-real-election-battlegrounds/
- [0506k] "General Election 2015: How the newspapers responded to Cameron's victory: The Conservatives won 33.8per cent of the vote", Kashmira Gander (@kashmiragander), The Independent, Friday 8 May 2015 23:09, see https://www.independent.co.uk/news/uk/politics/generalelection/general-election-2015-how-the-newspapers-responded-to-camerons-victory-10237949.html
- [0506l] "Election 2015: How the opinion polls got it wrong", By David Cowling, Editor, BBC Political Research Unit, 17 May 2015, see http://www.bbc.co.uk/news/uk-politics-32751993
- [0506m] "General election 2017 poll tracker: How the parties rated in the polls", by Mikey Smith Political Reporter, Joshua Taylor Assistant News Editor, Dan Bloom, The Mirror 13:25, 11 JUN 2017, UPDATED 14:31, 25 JAN 2018, see https://www.mirror.co.uk/news/politics/uk-politics-poll-tracker-2017-10266121
- [0506n] "Why did journalists get it so wrong on the election? Even when they spoke to real people, they didn't believe them", by Will Gore, (@willjgore), The Independent, Monday 12 June 2017 14:45, see https://www.independent.co.uk/voices/general-election-result-jeremy-corbyn-labour-journalists-mainstream-media-real-people-believe-a7786256.html
- [0506o] "House of Lords Committee on Political Polling and Digital Media: BPC Statement 6 July 2017", BPC, see http://www.britishpollingcouncil.org/house-of-lords-committee-on-political-polling-and-digital-media-bpc-statement/
- [0506p] "Political Polling and Digital Media Committee – timeline", see https://www.parliament.uk/business/committees/committees-a-z/lords-select/political-polling-digital-media/timeline/
- [0506q] "Select Committee on Political Polling and Digital Media: The politics of polling. Report of Session 2017-19 - published 17 April 2018 - HL Paper 106", see https://publications.parliament.uk/pa/ld201719/ldselect/ldppdm/106/10602.htm

- [0506r] Letter to the Editor, "Polls are powerful – and out of control. We must rein them in", from George Foulkes (Labour, House of Lords), first published Sun 2 Jul 2017 20.04 BST, see
- https://www.theguardian.com/politics/2017/jul/02/polls-are-powerful-and-out-of-control-we-must-rein-them-in
- [0506s] "NBC News SurveyMonkey Toplines and Methodology 1031 11 6", Monday, November 7, 2016 at 6:00 AM ET, NBC/Survey Monkey, see https://www.scribd.com/document/330243723/NBC-News-SurveyMonkey-Toplines-and-Methodology-1031-11-6
- [0506t] "British Polling Council Introduces New Rule on Uncertainty Attached to Polls", 1 May 2018, BPC Press Release, see http://www.britishpollingcouncil.org/british-polling-council-introduces-new-rule-on-uncertainty-attached-to-polls/

[0510]

- [0510e] "With the votes now counted, we can now see that they under-estimated the Conservatives by 4%, overestimated Labour by 3%, overstated the LibDems by 1% and understated others by 1%. The error in the estimation of the lead was 7%, the average error in the four estimates (Con, Lab, LibDem and Others as a group) of 2.25%. It is all eerily similar to 1992. In what has previously been labelled the great polling debacle, the pollsters, on average underestimated the Conservative share by 4% and overestimated Labour by 4%, an error in the lead estimate of 8% and an average error only marginally higher than 2015 at 2.75%." Nick Sparrow, former ICM, May 8th 2015, see http://www2.politicalbetting.com/index.php/archives/2015/05/08/ex-icm-boss-political-polling-pioneer-nick-sparrow-on-the-2015-polling-debacle/

[0512]

- [0512a] "Lies, damn lies and election polls: Why GE2015 pundits fluffed the numbers so badly. The lessons of shaping a mathematical 'reality'", by Jane Fae, The Register 11 May 2015 at 14:59, see https://www.theregister.co.uk/2015/05/11/why_the_election_pollsters_got_it_wrong/
- [0512b] "Thirsk and Malton election to put coalition government to test", Martin Wainwright, the Guardian, Wed 12 May 2010 19.12 BST, https://www.theguardian.com/politics/2010/may/12/thirsk-malton-election-coalition-government
- [0512c] "Shadow minister Woolas vows to fight election decision", BBC, 5 November 2010, see http://www.bbc.co.uk/news/uk-politics-11699888, see also http://www.bailii.org/ew/cases/EWHC/QB/2010/2702.html
- [0512d] "The Election. The Statistics. How the UK voted on May 1st.", BBC Politics 97 website, see http://www.bbc.co.uk/news/special/politics97/news/05/0505/stats.shtml
- [0512e] "House of Commons Library: UK Election Statistics: 1918-2017", published Wednesday August 23, 2017, see http://researchbriefings.files.parliament.uk/documents/CBP-7529/CBP-7529-UK-election-stats-23.08.2017.DOWNLOAD.xls , see also https://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7529

[0602]

- [0602a] "Good Judgement Open: Frequently Asked Questions (FAQ)", see https://www.gjopen.com/faq
- [0602b] "Verification Of Forecasts Expressed In Terms Of Probability" Glenn W. Brier, U. S. Weather Bureau, in the "Monthly Weather Review", Volume 78 Number 1, January 1950, see https://Web.Archive.Org/Web/20171023012737/Https://Docs.Lib.Noaa.Gov/Rescue/Mwr/078/Mwr-078-01-0001.Pdf

[0606]
- [0606a] "Will turnout weighting prove to be the pollsters' Achilles heel in #GE2017?" Patrick Sturgis and Will Jennings, University of Southampton. June 4 2017. See https://sotonpolitics.org/2017/06/04/will-turnout-weighting-prove-to-be-the-pollsters-achilles-heel-in-ge2017/
- [0606b] "Polling Observatory campaign report #3: All changed, changed utterly", May 31, 2017 The Polling Observatory (Robert Ford, Will Jennings, Mark Pickup and Christopher Wlezien). See https://sotonpolitics.org/2017/05/31/polling-observatory-campaign-report-3-all-changed-changed-utterly/

[0609]
- [0609a] "Why the polls were wrong in 2017", Anthony Wells, 7 June 2018, see http://ukpollingreport.co.uk/blog/archives/10002

18.   INTERVIEWS

[IS1] Interview with Professor Patrick Sturgis
[IJ1] Interview with Professor Will Jennings
[IT1] Interview with Luke Taylor of Kantar
[IW1] Interview with Anthony Wells of YouGov
[IF1] Interview with Stephen Fisher

APPENDIX 1: VERDICT TABLE

The table used for the verdict section is given below:

ElectionDate	Average MAE (3PF)	Average WIN
05/07/1945	0.012	1.000
23/02/1950	0.020	0.333
25/10/1951	0.027	0.000
26/05/1955	0.010	1.000
08/10/1959	0.010	1.000
15/10/1964	0.018	0.750
31/03/1966	0.018	1.000
18/06/1970	0.029	0.200
28/02/1974	0.018	0.857
10/10/1974	0.021	1.000
03/05/1979	0.011	1.000
09/06/1983	0.019	1.000
11/06/1987	0.019	1.000
09/04/1992	0.037	0.125
01/05/1997	0.027	1.000
07/06/2001	0.023	1.000
05/05/2005	0.014	1.000
06/05/2010	0.025	1.000
07/05/2015	0.030	0.417
08/06/2017	0.039	0.923

APPENDIX 2: THRESHOLD FOR A THREE-PARTY-FORCED MAE

To work out what the MAE threshold is for acceptable elections, we need to work backwards by looking at unacceptable elections and working out what the MAE was for them. Fortunately we have one close at hand: the 2015 election. In the immediate post-mortem for 2015 the pollster Nick Sparrow, formerly of ICM, made these remarks:

"With the votes now counted, we can now see that they under-estimated the Conservatives by 4%, overestimated Labour by 3%, overstated the LibDems by 1% and understated others by 1%. The error in the estimation of the lead was 7%, the average error in the four estimates (Con, Lab, LibDem and Others as a group) of 2.25%. It is all eerily similar to 1992. In what has previously been labelled the great polling debacle, the pollsters, on average underestimated the Conservative share by 4% and overestimated Labour by 4%, an error in the lead estimate of 8% and an average error only marginally higher than 2015 at 2.75%." [0510e]

From Sparrow's remarks we can tell that the threshold for an election of a four-party-forced basis (Con/Lab/Lib/Oth) is 2% or less: if the mean absolute error of the Conservative, Labour, Liberal and Other vote share is 2% or less then the poll is acceptable.

But what is the threshold of a three-part-forced poll? If we force the poll into Con/Lab/Oth format, then how close do the Conservative, Labour and Other vote shares have to be? The MAE's for 1992 and 2015 on a three-party-forced basis are as follows:

2015-05-07 election. Source: [0430a], Table 1

Pollster	Fieldwork	Con	Lab	Other 3pf	Con Actual	Lab Actual	Oth3PF Actual	MAE (3pf)	WIN
Populus	5–6_May	0.340	0.340	0.320	0.377	0.312	0.311	0.0247	0
Ipsos-MORI	5–6_May	0.360	0.350	0.290	0.377	0.312	0.311	0.0253	1
YouGov	4–6_May	0.340	0.340	0.320	0.377	0.312	0.311	0.0247	0
ComRes	5–6_May	0.350	0.340	0.310	0.377	0.312	0.311	0.0187	1
Survation	4–6_May	0.310	0.310	0.380	0.377	0.312	0.311	0.0460	0
ICM	3–6_May	0.340	0.350	0.310	0.377	0.312	0.311	0.0253	0
Panelbase	1–6_May	0.310	0.330	0.360	0.377	0.312	0.311	0.0447	0
Opinium	4–5_May	0.350	0.340	0.310	0.377	0.312	0.311	0.0187	1
TNS_UK	30_April-4_May	0.330	0.320	0.350	0.377	0.312	0.311	0.0313	1
Ashcroft*1	5–6_May	0.330	0.330	0.340	0.377	0.312	0.311	0.0313	0
BMG*1	3–5_May	0.340	0.340	0.320	0.377	0.312	0.311	0.0247	0
SurveyMonkey*1	30_April-6_May	0.340	0.280	0.380	0.377	0.312	0.311	0.0460	1
							average	0.0301	0.4167

* = non-members of British Polling Council at May 2015

1992-04-09 election. Source: [0430a], Appendix 3

Pollster	Fieldwork	Con	Lab	Other 3pf	Con Actual	Lab Actual	Oth3PF Actual	MAE (3pf)	WIN
ASL	April_3-4	0.350	0.380	0.270	0.428	0.352	0.220	0.0520	0
Harris	April_7	0.380	0.400	0.220	0.428	0.352	0.220	0.0320	0
Harris	April_7	0.370	0.380	0.250	0.428	0.352	0.220	0.0387	0
MORI	April_7	0.370	0.400	0.230	0.428	0.352	0.220	0.0387	0
MORI	April_7-8	0.380	0.390	0.230	0.428	0.352	0.220	0.0320	0
NOP	April_7-8	0.390	0.420	0.190	0.428	0.352	0.220	0.0453	0

Gallup	April_7-8	0.390	0.380	0.230	0.428	0.352	0.220	0.0253	1
ICM	April_8	0.380	0.380	0.240	0.428	0.352	0.220	0.0320	0
							average	0.0370	0.1250

So the MAEs for 1992 and 2015 on a three-party-forced basis are 3.7% and 3% respectively. Using the same logic as Sparrow that would give us a threshold of 2.75% for three-party-forced. A quick sanity check pointed out that 1951 had a MAE of 2.71% so we rounded the threshold to 2.7%

So our threshold for acceptable MAEs are 2% or less on a four-party-forced basis (Con/Lab/Lib/Other) and 2.7% or less on a three-party-forced basis (Con/Lab/Oth)

APPENDIX 3: SOURCE RECONCILIATION FOR 1951,70,92,2015 AND 17

Sources do not necessarily agree for results. Here is a selection.

Code	Geo	Con 1951	Lab 1951	Con 1970	Lab 1970	Con 1992	Lab 1992	Con 2015	Lab 2015	Con 2017	Lab 2017
[0423c]	GB	47.8	49.4	46.2	43.8	42.8	35.2	37.7	31.2		
[0424d]	GB									43.5	41
[0424e]	GB									43.4	41
[0424f]	GB							37.8	31.2		
[0512e]	GB	47.8	49.4	46.2	43.8	42.8	35.2	37.7	31.2	43.4	41
[0423c]	UK									42.4	40
[0424e]	UK									42.3	40
[0512e]	UK	48	48.8	46.4	43	41.9	34.4	36.8	30.4	42.3	40

APPENDIX 4: HOUSE OF LORDS REPORT HL 106

The Political Polling and Digital Media Committee of the House of Lords was appointed on 29 June 2017[0506p] in response to perceived shortcomings in political opinion polling in the UK. It was wide ranging and took statements from the great and the good of academia, polling, researchers, journalists and the media, bookies and gamblers and others.[0506e], who are listed in Appendix 5. It published its report on 17 April 2018.[0506q]

The House of Lords report was HL Paper 106 and was given the title of "The politics of polling". The committee agreed with Jennings that polls had not become more inaccurate over time but nevertheless thought that for various reasons things may have taken a turn for the worse and may continue to deteriorate. To counter this it exhorted further effort and innovation from pollsters, especially on their weighting variables.

The committee considered the media. It thought that reportage of polls left something to be desired, with its habits of interpreting random variability as significant change and writing headlines that may not accurately reflect the poll being reported. To counter this, it recommended that the BPC expand its "Journalist's Guide to Opinion Polls", name-and-shame bad reportage, and in tandem with other organisations (RSS, MRS et al) develop training courses for journalist on how to report polls accurately.

When it came to oversight the committee refrained from recommending banning polls. Instead it recommended a rather dramatic expansion of the BPC's remit. It called upon the BPC (again with other organisations) to proactively oversee the conduct and reporting of polls, to advise, train,

monitor and judge pollsters and journalists both contemporaneously and retrospectively. This would radically change its role from the industry's adviser and representative to its overseer and enforcer.

For election periods, the committee advised that the Electoral Commission should be the body charged with ensuring that all polls during the campaign report their funding. It then made other recommendations regarding education and finished by recommending that the Government open multilateral dialogue with various international bodies about social media.[0424h]

The BPC were a little bemused by the report. It welcomed it but pointed out that it was a voluntary body with limited resources. [0424i]

APPENDIX 5: WITNESSES TO THE HOUSE OF LORDS COMMITTEE
The Political Polling and Digital Media Committee was appointed on 29 June 2017 and reported on 17 April 2018. During that time, it took statements from many people. Those who testified in public session are listed below. [0506p]

ACADEMICS
- Professor Will Jennings, Professor of Political Science and Public Policy, Southampton University
- Dr Benjamin Lauderdale, Associate Professor, London School of Economics
- Dr Jouni Kuha, Associate Professor of Statistics and Research Methodology, London School of Economics
- Professor Susan Banducci, Professor and Director of the Exeter Q-Step Centre, University of Exeter
- Professor Jane Green, University of Manchester
- Professor Chris Hanretty, Royal Holloway, University of London
- Professor Helen Margetts, Director, Oxford Internet Institute
- Professor Farida Vis, Professor of Digital Media, Manchester School of Art,Manchester Metropolitan University
- James Williams, Doctoral Candidate, Oxford Internet Institute

POLLSTERS AND RESEARCHERS
- Nick Moon, Moonlight Research
- Carl Miller, Research Director of the Centre for the Analysis of Social Media, Demos.
- Jean Pierre Kloppers, CEO at BrandsEye
- Deborah Mattinson, Co-Founder Britain Thinks
- Nicolas Sauger, Co-Director of the Laboratory for Interdisciplinary Evaluation of Public Policies (LIEPP)
- Professor John Curtice, President, BPC
- Simon Atkinson, Management Committee Member, BPC
- Johnny Heald, Managing Director, ORB
- Ben Page, CEO, Ipsos Mori
- Damian Lyons Lowe, Chief Executive, Survation
- Jane Frost, Chief Executive, Market Research Society

FACT CHECKERS, JOURNALISTS AND THE MEDIA
- Will Moy, Director, Full Fact
- David Cowling, Former Editor of Political Research BBC
- Jim Waterson, Politics Editor, Buzzfeed

- Sue Inglish, Former Head of Political Programmes, Analysis and Research, BBC
- Professor Richard Tait CBE, Professor of Journalism, Cardiff University
- Matt Tee, Chief Executive, Independent Press Standards Organisation (IPSO)
- Jonathan Levy, Director of Newsgathering and Operations, Sky News
- Professor Michael Thrasher, Co-Director of the Elections Centre and adviser to Sky
- David Jordan, Director of Editorial Policy and Standards, BBC
- Ric Bailey, Chief Adviser, Editorial Standards & Policy, BBC
- Ian Murray, Executive Director, Society of Editors

BOOKIES AND GAMBLERS
- Mike Smithson, Founder and Editor, Politicalbetting.com
- Matthew Shaddick, Head of Political Betting, Ladbrokes

OTHERS
- Jonathan Heawood, Chief Executive Officer, IMPRESS
- Lord Kinnock, Former Leader of the Labour party
- Claire Bassett, Chief Executive, Electoral Commission
- Bob Posner, Director of Political Finance and Regulation & Legal Counsel, Electoral Commission
- Sir Patrick McLoughlin, MP, Chancellor of the Duchy of Lancaster
- Matt Hancock MP, the Minister for Digital
- Peter Lee, Director of the Constitution Group

APPENDIX 6: GENERAL ELECTION FINAL POLLS

Year	Source	Link
1945-2010	NCRM Appendix3	http://eprints.ncrm.ac.uk/3789/1/Report_final_revised.pdf , appendix 3
2015	NCRM Table1	http://eprints.ncrm.ac.uk/3789/1/Report_final_revised.pdf , table1
2017	BPC 2017 Mark Pack	http://www.britishpollingcouncil.org/general-election-8-june-2017/ plus four others from Mark Pack's spreadsheet "PollBase-Q4-2017.xls" from https://www.markpack.org.uk/opinion-polls/

Year	Pollster	Publisher	Fieldwork	Sample	Con	Lab	Oth	Source
2017	ORB**	Sunday Telegraph	May 31 to 1 Jun	?	45	36	19	Mark Pack
2017	Norstat*	Sunday Express	May 31 to 1 Jun	?	39	35	26	Mark Pack
2017	Survey Monkey*	Sun	Jun 4 to 6	?	42	38	20	Mark Pack
2017	Qriously*	?	Jun 4 to 7	?	39	41	20	Mark Pack
2017	Opinium	?	Jun 4	3002	43	36	21	BPC 2017
2017	Survation	?	Jun 6-7	2798	41	40	19	BPC 2017
2017	Ipsos MORI	?	Jun 6-7	1291	44	36	20	BPC 2017
2017	ICM	?	Jun 6-7	1532	46	34	20	BPC 2017
2017	ComRes	?	Jun 5-7	2051	44	34	22	BPC 2017
2017	YouGov	?	Jun 5-7	2130	42	35	23	BPC 2017
2017	Panelbase	?	Jun 2-7	3018	44	36	20	BPC 2017
2017	Kantar Public	?	Jun 1-7	2159	43	38	19	BPC 2017
2017	BMG	?	Jun 6-7	1199	46	33	21	BPC 2017
2015	Populus	?	5–6 May	3917	34	34	32	NCRM Table1
2015	Ipsos-MORI	?	5–6 May	1186	36	35	29	NCRM Table1

2015	YouGov	?	4–6 May	10307	34	34	32	NCRM Table1
2015	ComRes	?	5–6 May	1007	35	34	31	NCRM Table1
2015	Survation	?	4–6 May	4088	31	31	38	NCRM Table1
2015	ICM	?	3–6 May	2023	34	35	31	NCRM Table1
2015	Panelbase	?	1–6 May	3019	31	33	36	NCRM Table1
2015	Opinium	?	4–5 May	2960	35	34	31	NCRM Table1
2015	TNS UK	?	30 Apr-4 May	1185	33	32	35	NCRM Table1
2015	Ashcroft*	?	5–6 May	3028	33	33	34	NCRM Table1
2015	BMG*	?	3–5 May	1009	34	34	32	NCRM Table1
2015	Survey Monkey*	?	30 Apr-6 May	18131	34	28	38	NCRM Table1
2010	TNS-BMRB	?	29 Apr–4 May	1864	33	27	40	NCRM Appendix3
2010	BPIX	Mail on Sunday	30 Apr-1 May	2136	34	27	39	NCRM Appendix3
2010	ICM	Guardian	3-4 May	1527	36	28	36	NCRM Appendix3
2010	Harris	Daily Mail	4-5 May	4014	35	29	36	NCRM Appendix3
2010	Angus Reid	Political Betting	4-5 May	2283	36	24	40	NCRM Appendix3
2010	ComRes	Independent/ITV News	4-5 May	1025	37	28	35	NCRM Appendix3
2010	Opinium	Daily Express	4-5 May	1383	35	27	38	NCRM Appendix3
2010	Populus	Times	4-5 May	2505	37	28	35	NCRM Appendix3
2010	YouGov	Sun	4-5 May	6483	35	28	37	NCRM Appendix3
2010	Ipsos MORI	Standard	5 May	1216	36	29	35	NCRM Appendix3
2005	ComRes	Independent on Sunday	23-28 Apr	1091	31	39	30	NCRM Appendix3
2005	BPIX	Mail on Sunday	?-29 Apr	?	33	37	30	NCRM Appendix3
2005	Populus	Times	29 Apr-2 May	1420	27	41	32	NCRM Appendix3
2005	ICM	Guardian	1-3 May	1532	32	38	30	NCRM Appendix3
2005	NOP	Independent	2-3 May	1000	33	36	31	NCRM Appendix3
2005	Populus	Times	2-3 May	2042	32	38	30	NCRM Appendix3
2005	Ipsos MORI	Evening Standard	3-4 May	1628	33	38	29	NCRM Appendix3
2005	Harris Interactive	?	3-4 May	4116	33	38	29	NCRM Appendix3
2005	YouGov	Telegraph	3-4 May	3962	32	37	31	NCRM Appendix3
2001	NOP	Sunday Times	May 31– Jun 1	1105	30	47	23	NCRM Appendix3
2001	Rasmussen	Independent	Jun 2-3	1266	33	44	23	NCRM Appendix3
2001	ICM	Evening Standard	Jun 2-3	1332	30	47	23	NCRM Appendix3
2001	ICM	Guardian	Jun 2-4	1009	32	43	25	NCRM Appendix3
2001	MORI	Economist	Jun 4-5	1010	31	43	26	NCRM Appendix3
2001	MORI	Times	Jun 5-6	1967	30	45	25	NCRM Appendix3
2001	Gallup	Daily Telegraph	Jun 6	2399	30	47	23	NCRM Appendix3
2001	NOP	Sunday Times	May 31– Jun 1	1105	30	47	23	NCRM Appendix3
2001	Rasmussen	Independent	Jun 2-3	1266	33	44	23	NCRM Appendix3
1997	Harris	Independent	27-29 Apr	1010	31	48	21	NCRM Appendix3
1997	NOP	Reuters	29 Apr	1000	28	50	22	NCRM Appendix3
1997	MORI	Times	29-30 Apr	2304	27	51	22	NCRM Appendix3
1997	ICM	Guardian	29-30 Apr	1555	33	43	24	NCRM Appendix3
1997	Gallup	Daily Telegraph	30 Apr	1849	33	47	20	NCRM Appendix3
1997	MORI	Evening Standard	30 Apr	1501	29	47	24	NCRM Appendix3
1992	ASL	?	Apr 3-4	1038	35	38	27	NCRM Appendix3
1992	Harris	ITN	Apr 7	2210	38	40	22	NCRM Appendix3
1992	Harris	Daily Express	Apr 7	1093	37	38	25	NCRM Appendix3
1992	MORI	Yorkshire TV	Apr 7	1065	37	40	23	NCRM Appendix3
1992	MORI	Times	Apr 7-8	1731	38	39	23	NCRM Appendix3

1992	NOP	Independent	Apr 7-8	1746	39	42	19	NCRM Appendix3
1992	Gallup	Daily Telegraph	Apr 7-8	2478	39	38	23	NCRM Appendix3
1992	ICM	Guardian	Apr 8	2186	38	38	24	NCRM Appendix3
1987	Gallup	Daily Telegraph	Jun 8-9	2505	41	34	25	NCRM Appendix3
1987	Marplan	Today	Jun 9	1086	43	35	22	NCRM Appendix3
1987	MORI	Times	Jun 9-10	1668	44	32	24	NCRM Appendix3
1987	Marplan	Guardian	Jun 10	1633	42	35	23	NCRM Appendix3
1987	NOP	Independent	Jun 10	1668	42	35	23	NCRM Appendix3
1987	ASL	Sun	Jun 9	1702	43	34	23	NCRM Appendix3
1987	Harris	TV-am	Jun 8-9	2122	42	35	23	NCRM Appendix3
1983	Gallup	Daily Telegraph	Jun 7-8	2003	45.5	26.5	28	NCRM Appendix3
1983	ASL	Sun	Jun 8	1100	46	23	31	NCRM Appendix3
1983	Marplan	Guardian	Jun 8	1335	46	26	28	NCRM Appendix3
1983	MORI	Evening Standard	Jun 8	1101	44	28	28	NCRM Appendix3
1983	Harris	Observer	Jun 8	576	47	25	28	NCRM Appendix3
1983	NOP	Northcliffe	Jun 8	1083	47	25	28	NCRM Appendix3
1979	MORI	Express	Apr 29-May 1	947	44.4	38.8	16.8	NCRM Appendix3
1979	Marplan	Sun	May 1	1973	45	38	17	NCRM Appendix3
1979	Gallup	Daily Telegraph	May 1-2	2348	43	41	16	NCRM Appendix3
1979	MORI	Evening Standard	May 2	1089	45	37	18	NCRM Appendix3
1979	NOP	Daily Mail	May 1-2	1069	46	39	15	NCRM Appendix3
1974 Oct	NOP	Daily Mail	Oct 2-5	1978	31	45.5	23.5	NCRM Appendix3
1974 Oct	Gallup	Daily Telegraph	Oct 3-7	954	36	41.5	22.5	NCRM Appendix3
1974 Oct	Marplan	Sun	Oct 8	1024	33.3	43.8	22.9	NCRM Appendix3
1974 Oct	ORC	Times	Oct 8	446	35	44	21	NCRM Appendix3
1974 Oct	Harris	Daily Express	Oct 5-9	678	34.6	43	22.4	NCRM Appendix3
1974 Oct	ORC	Evening Standard	Oct 5-9	1071	34.4	41.8	23.8	NCRM Appendix3
1974 Feb	Business Decisions	Observer	Feb 21	1056	36	37.5	26.5	NCRM Appendix3
1974 Feb	Gallup	Daily Telegraph	Feb 26-27	1881	39.5	37.5	23	NCRM Appendix3
1974 Feb	Harris	Express	Feb 26-27	3193	40.2	35.2	24.6	NCRM Appendix3
1974 Feb	NOP	Daily Mail	Feb 27	4038	39.5	35.5	25	NCRM Appendix3
1974 Feb	ORC	Evening Standard	Feb 27	2327	39.7	36.7	23.6	NCRM Appendix3
1974 Feb	ORC	Times	Feb 28	2327	38.7	34.9	26.4	NCRM Appendix3
1974 Feb	Marplan	London Weekend Television	Feb 28	2649	36.5	34.5	29	NCRM Appendix3
1970	Gallup	Daily Telegraph	Jun 14-16	2190	42	49	9	NCRM Appendix3
1970	NOP	Daily Mail	Jun 12-16	1562	44.1	48.1	7.8	NCRM Appendix3
1970	Marplan	Times	Jun 11-14	2267	41.5	50.2	8.3	NCRM Appendix3
1970	Harris	Daily Express	?	?	46	48	6	NCRM Appendix3
1970	ORC	Evening Standard	Jun 13-17	1840	46.5	45.5	8	NCRM Appendix3
1966	NOP	Daily Mail	Mar 27-29	1693	41.6	50.6	7.8	NCRM Appendix3
1966	Gallup	Daily Telegraph	Mar 24-28	?	40	51	9	NCRM Appendix3
1966	?	Daily Express	?	?	37.4	54.1	8.5	NCRM Appendix3

1966	Research Services	Observer	?	?	41.6	49.7	8.7	NCRM Appendix3
1964	Research Services	Observer	?	?	45	46	9	NCRM Appendix3
1964	NOP	Daily Mail	Oct 9-13	1179	44.3	47.4	8.3	NCRM Appendix3
1964	Gallup	Daily Telegraph	Oct 8-13	3829	43.5	46	10.5	NCRM Appendix3
1964	?	Daily Express	?	?	44.5	43.7	11.8	NCRM Appendix3
1959	NOP	Daily Mail	Oct 2-5	?	48	44.1	7.9	NCRM Appendix3
1959	?	Daily Express	?	?	49.1	45.4	5.5	NCRM Appendix3
1959	Gallup	News Chronicle	Oct 3-6	?	48.5	46.5	5	NCRM Appendix3
1959	Forecasting Statistics	Daily Telegraph	Oct 1-4	?	49	46	5	NCRM Appendix3
1955	Gallup (BIPO)	News Chronicle	May 21-24	?	51	47.5	1.5	NCRM Appendix3
1955	?	Daily Express	?	?	50.2	47.2	2.6	NCRM Appendix3
1951	Gallup (BIPO)	News Chronicle	Oct 22	?	49.5	47	3.5	NCRM Appendix3
1951	?	Daily Express	Oct 19-23	?	50	46	4	NCRM Appendix3
1951	Research Services	Daily Graphic	Oct 19	?	50	43	7	NCRM Appendix3
1950	Gallup (BIPO)	News Chronicle	Feb 17–20	?	43.5	45	11.5	NCRM Appendix3
1950	?	Daily Express	Feb 17–21	?	44.5	44	11.5	NCRM Appendix3
1950	?	Daily Mail	Feb 5–11	?	45.5	42.5	12	NCRM Appendix3
1945	Gallup	News Chronicle	Jun 24-27	?	41	47	12	NCRM Appendix3

* Non-BPC. BMG joined BPC in February 2016

** ORB is a BPC member but BPC did not include it as a final poll, see [0424d]

APPENDIX 7: GENERAL ELECTION VOTE SHARE RESULTS

- Source (paper) https://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7529
- Source (spreadsheet) http://researchbriefings.files.parliament.uk/documents/CBP-7529/CBP-7529-UK-election-stats-23.08.2017.DOWNLOAD.xls

THREE-PARTY-FORCED: LAB/CON/OTHER

Date	Election	Area	Lab 2dp	Con 2dp	Oth 2dp	Lab 3dp	Con 3dp	Oth 3dp
05/07/1945	1945	UK	0.40	0.48	0.12	0.397	0.477	0.126
23/02/1950	1950	UK	0.43	0.46	0.11	0.433	0.461	0.106
25/10/1951	1951	UK	0.48	0.49	0.03	0.480	0.488	0.032
26/05/1955	1955	UK	0.50	0.46	0.04	0.496	0.464	0.040
08/10/1959	1959	UK	0.49	0.44	0.07	0.494	0.438	0.068
15/10/1964	1964	UK	0.43	0.44	0.13	0.433	0.441	0.126
31/03/1966	1966	UK	0.42	0.48	0.10	0.419	0.479	0.102
18/06/1970	1970	UK	0.46	0.43	0.11	0.464	0.430	0.106
28/02/1974	1974Feb	UK	0.38	0.37	0.25	0.378	0.372	0.250
10/10/1974	1974Oct	UK	0.36	0.39	0.25	0.357	0.393	0.250
03/05/1979	1979	UK	0.44	0.37	0.19	0.439	0.369	0.192
09/06/1983	1983	UK	0.42	0.28	0.30	0.424	0.276	0.300
11/06/1987	1987	UK	0.42	0.31	0.27	0.422	0.308	0.270
09/04/1992	1992	UK	0.42	0.34	0.24	0.419	0.344	0.237
01/05/1997	1997	UK	0.31	0.43	0.26	0.307	0.432	0.261
07/06/2001	2001	UK	0.32	0.41	0.27	0.316	0.407	0.277

Date	Election	Area						
05/05/2005	2005	UK	0.32	0.35	0.33	0.324	0.352	0.324
06/05/2010	2010	UK	0.36	0.29	0.35	0.361	0.290	0.349
07/05/2015	2015	UK	0.37	0.30	0.33	0.368	0.304	0.328
08/06/2017	2017	UK	0.42	0.40	0.18	0.423	0.400	0.177
05/07/1945	1945	GB	0.39	0.49	0.12	0.393	0.488	0.119
23/02/1950	1950	GB	0.43	0.47	0.10	0.429	0.468	0.103
25/10/1951	1951	GB	0.48	0.49	0.03	0.478	0.494	0.028
26/05/1955	1955	GB	0.49	0.47	0.04	0.492	0.474	0.034
08/10/1959	1959	GB	0.49	0.45	0.06	0.488	0.446	0.066
15/10/1964	1964	GB	0.43	0.45	0.12	0.429	0.448	0.123
31/03/1966	1966	GB	0.41	0.49	0.10	0.414	0.487	0.099
18/06/1970	1970	GB	0.46	0.44	0.10	0.462	0.438	0.100
28/02/1974	1974Feb	GB	0.39	0.38	0.23	0.386	0.380	0.234
10/10/1974	1974Oct	GB	0.37	0.40	0.23	0.366	0.402	0.232
03/05/1979	1979	GB	0.45	0.38	0.17	0.449	0.377	0.174
09/06/1983	1983	GB	0.44	0.28	0.28	0.435	0.283	0.282
11/06/1987	1987	GB	0.43	0.32	0.25	0.432	0.315	0.253
09/04/1992	1992	GB	0.43	0.35	0.22	0.428	0.352	0.220
01/05/1997	1997	GB	0.32	0.44	0.24	0.315	0.443	0.242
07/06/2001	2001	GB	0.33	0.42	0.25	0.326	0.420	0.254
05/05/2005	2005	GB	0.33	0.36	0.31	0.332	0.361	0.307
06/05/2010	2010	GB	0.37	0.30	0.33	0.369	0.297	0.334
07/05/2015	2015	GB	0.38	0.31	0.31	0.377	0.312	0.311
08/06/2017	2017	GB	0.43	0.41	0.16	0.434	0.410	0.156

FOUR-PARTY-FORCED: LAB/CON/LIB/OTHER

Note: the paper includes the following in "the LD vote": Coalition Liberal Party for 1918; National Liberal for 1922; and Independent Liberal for 1931. Figures show Liberal/SDP Alliance vote for 1983-1987 and Liberal Democrat vote from 1992 onwards.

Date	Election	Area	Lab 2dp	Con 2dp	LD 2dp	Oth 2dp	Lab 3dp	Con 3dp	LD 3dp	Oth 3dp
05/07/1945	1945	UK	0.40	0.48	0.09	0.03	0.397	0.477	0.090	0.036
23/02/1950	1950	UK	0.43	0.46	0.09	0.02	0.433	0.461	0.091	0.015
25/10/1951	1951	UK	0.48	0.49	0.03	0.00	0.480	0.488	0.026	0.006
26/05/1955	1955	UK	0.50	0.46	0.03	0.01	0.496	0.464	0.027	0.013
08/10/1959	1959	UK	0.49	0.44	0.06	0.01	0.494	0.438	0.059	0.009
15/10/1964	1964	UK	0.43	0.44	0.11	0.02	0.433	0.441	0.112	0.014
31/03/1966	1966	UK	0.42	0.48	0.09	0.01	0.419	0.479	0.085	0.017
18/06/1970	1970	UK	0.46	0.43	0.08	0.03	0.464	0.430	0.075	0.031
28/02/1974	1974Feb	UK	0.38	0.37	0.19	0.06	0.378	0.372	0.193	0.057
10/10/1974	1974Oct	UK	0.36	0.39	0.18	0.07	0.357	0.393	0.183	0.067
03/05/1979	1979	UK	0.44	0.37	0.14	0.05	0.439	0.369	0.138	0.054
09/06/1983	1983	UK	0.42	0.28	0.25	0.05	0.424	0.276	0.254	0.046
11/06/1987	1987	UK	0.42	0.31	0.23	0.04	0.422	0.308	0.226	0.044
09/04/1992	1992	UK	0.42	0.34	0.18	0.06	0.419	0.344	0.178	0.059
01/05/1997	1997	UK	0.31	0.43	0.17	0.09	0.307	0.432	0.168	0.093
07/06/2001	2001	UK	0.32	0.41	0.18	0.09	0.316	0.407	0.183	0.094
05/05/2005	2005	UK	0.32	0.35	0.22	0.11	0.324	0.352	0.220	0.104

06/05/2010	2010	UK	0.36	0.29	0.23	0.12	0.361	0.290	0.230	0.119
07/05/2015	2015	UK	0.37	0.30	0.08	0.25	0.368	0.304	0.079	0.249
08/06/2017	2017	UK	0.42	0.40	0.07	0.11	0.423	0.400	0.074	0.103
05/07/1945	1945	GB	0.39	0.49	0.09	0.03	0.393	0.488	0.092	0.027
23/02/1950	1950	GB	0.43	0.47	0.09	0.01	0.429	0.468	0.093	0.010
25/10/1951	1951	GB	0.48	0.49	0.03	0.00	0.478	0.494	0.026	0.002
26/05/1955	1955	GB	0.49	0.47	0.03	0.01	0.492	0.474	0.028	0.006
08/10/1959	1959	GB	0.49	0.45	0.06	0.00	0.488	0.446	0.060	0.006
15/10/1964	1964	GB	0.43	0.45	0.11	0.01	0.429	0.448	0.114	0.009
31/03/1966	1966	GB	0.41	0.49	0.09	0.01	0.414	0.487	0.086	0.013
18/06/1970	1970	GB	0.46	0.44	0.08	0.02	0.462	0.438	0.076	0.024
28/02/1974	1974Feb	GB	0.39	0.38	0.20	0.03	0.386	0.380	0.198	0.036
10/10/1974	1974Oct	GB	0.37	0.40	0.19	0.04	0.366	0.402	0.188	0.044
03/05/1979	1979	GB	0.45	0.38	0.14	0.03	0.449	0.377	0.141	0.033
09/06/1983	1983	GB	0.44	0.28	0.26	0.02	0.435	0.283	0.260	0.022
11/06/1987	1987	GB	0.43	0.32	0.23	0.02	0.432	0.315	0.231	0.022
09/04/1992	1992	GB	0.43	0.35	0.18	0.04	0.428	0.352	0.183	0.037
01/05/1997	1997	GB	0.32	0.44	0.17	0.07	0.315	0.443	0.172	0.070
07/06/2001	2001	GB	0.33	0.42	0.19	0.06	0.326	0.420	0.188	0.066
05/05/2005	2005	GB	0.33	0.36	0.23	0.08	0.332	0.361	0.226	0.081
06/05/2010	2010	GB	0.37	0.30	0.24	0.09	0.369	0.297	0.236	0.098
07/05/2015	2015	GB	0.38	0.31	0.08	0.23	0.377	0.312	0.081	0.230
08/06/2017	2017	GB	0.43	0.41	0.08	0.08	0.434	0.410	0.076	0.080

APPENDIX 8: WIN VS BRIER SCORE

These series of articles use the WIN score, where

- For 1 prediction: WIN_t = 1 if prediction t predicted the correct winner, 0 if it did not
- For N predictions: $WIN = \frac{1}{N}\sum_{t=1}^{N} WIN_t$

How does that relate to the Brier score? The Brier score is the recommended metric of the Good Judgement project[0602a] and was recommended in our interview with Stephen Fisher[IF1]. The Brier score (BS) looks like this:

- For 1 prediction $BS_t = \sum_{i=1}^{R}(f_{ti} - o_{ti})^2$

   where f_{ti} = prediction t's predicted probability that party i will win,
   and o_{ti} = 1 if party i wins, 0 if party i does not win
   and R = R_t = number of parties being predicted by prediction t
   and i ∈ {1,..R}

- For N predictions: $BS = \frac{1}{N}\sum_{t=1}^{N} BS_t$
   $$= \frac{1}{N}\sum_{t=1}^{N}\sum_{i=1}^{R}(f_{ti} - o_{ti})^2 \qquad \text{[0602b]}$$

   where f_{ti} = prediction t's predicted probability that party i will win,
   and o_{ti} = 1 if party i wins, 0 if party i does not win
   and R = R_t = number of parties being predicted by prediction t
   and i ∈ {1,..R}

and N = the number of times you make a prediction

and $t \in \{1,..N\}$

This is the formulation used by the Good Judgement Project [0602a] who quote thus:

The Brier score is the squared error of a probabilistic forecast. To calculate it, we divide your forecast by 100 so that your probabilities range between 0 (0%) and 1 (100%). Then, we code reality as either 0 (if the event did not happen) or 1 (if the event did happen). For each answer option, we take the difference between your forecast and the correct answer, square the differences, and add them all together. For a yes/no question where you forecasted 70% and the event happened, your score would be $(1 - 0.7)^2 + (0 - 0.3)^2 = 0.18$. For a question with three possible outcomes (A, B, C) where you forecasted A = 60%, B = 10%, C = 30% and A occurred, your score would be $(1 - 0.6)^2 + (0 - 0.1)^2 + (0 - 0.3)^2 = 0.26$. The best (lowest) possible Brier score is 0, and the worst (highest) possible Brier score is 2.

If we restrict ourselves to a prediction of the winner with no possibility of a draw, then there are only two possible outcomes: party i wins and "somebody else" loses, or party i loses and "somebody else" wins. Under these circumstances, then there are two parties – the party you are predicting and "somebody else" - so R = 2 and $f_{t2} = (1 - f_{t1})$ and $o_{t2} = (1 - o_{t1})$. Substituting these in we obtain

- For 1 prediction $BS_t = [\,(f_{t1} - o_{t1})^2 \;+\; (f_{t2} - o_{t2})^2\,]$
$$= [\,(f_{t1} - o_{t1})^2 \;+\; ((1 - f_{t1}) - (1 - o_{t1}))^2\,]$$
$$= [\,(f_{t1} - o_{t1})^2 \;+\; (o_{t1} - f_{t1})^2\,]$$
$$= [\,(f_{t1} - o_{t1})^2 \;+\; (f_{t1} - o_{t1})^2\,]$$
$$= [\,2\,(f_{t1} - o_{t1})^2\,]$$
$$= 2(f_{t1} - o_{t1})^2$$

But for a binary prediction of the winner with no possibility of a draw, f_{t1} is either 1 ("prediction t predicts party i wins") or 0 ("prediction t predicts party i loses"), and o_{t1} = 0 or 1, depending on f_{t1}. Plugging in those values gives us

- For 1 prediction $BS_t = 2(1 - 0)^2 = 2$ if you predict party i wins and it loses
or $BS_t = 2(1 - 1)^2 = 0$ if you predict party i wins and it wins

But for a binary prediction of the winner with no possibility of a draw, then we know that

- For 1 prediction $WIN_t = 0$ if you predict party i wins and it loses
or $WIN_t = 1$ if you predict party i wins and it wins

Or to put it simply. For a binary prediction of the winner with no possibility of a draw, then

- For 1 prediction $BS_t = 2 * (1 - WIN_t)$
- For N predictions: $BS = 2 * (1 - WIN\ \ )$