

FORECAST ERROR: PREDICTORS OF THE 2017 UK GENERAL ELECTION

Author: Timothy Martyn Hill, 28 June 2017

INTRODUCTION

On 18 April 2017 British Prime Minister Theresa May announced [0507s] a snap general election for 8 June 2017, hoping the country would deliver her the strong parliamentary majority she believed she needed to successfully conduct the UK's Brexit negotiations. The result on 9 June – a hung parliament, with no party having an overall majority – was not what she would have expected seven weeks before. Nor was it what many of the pollsters, betting firms, modellers and academics were anticipating. In this respect, it was one of the most interesting elections in a long while.

In this article, we will list the predictors of this election and how they performed between the announcement and election day.

METHODOLOGY

In reviewing previous elections, we outlined various problems and methods of coping with them. But this election was called at short notice and so there were other problems to contend with.

One obvious problem was that, due to the short notice, there weren't that many academic predictors in the beginning – plenty of odds and polls, but building models takes time and academics and analysts needed a few days to catch up. To cope with this backloading, we tried to select predictors that extended from just after the date the election was called to election day itself, and preferably changed over time, to a maximum of five for each category.

The short notice also required a change in technique: instead of painstakingly trying to reconstruct every prediction, we took samples every Wednesday from Wednesday 26 April 2017, using the latest value at that point. This made data-gathering much easier and the flood of predictors on the last day could then be handled separately.

To comparably assess accuracy we will use four-party-forced: all predictions will be forced into of four categories corresponding to three parties – Conservative (Con), Labour (Lab) and Liberal Democrats (Lib) – plus all others (Other). These four-party-forced predictions are then measured against the result. This matches what was done in 2010 [0618b][0618c] and 2015 [0618a], and was computationally easier and enabled us to compare bets to odds to models.

As is our practice, we selected mean absolute error (MAE) to measure accuracy. It or its variants are widely understood by pollsters and modellers, and can be used for any finite prediction that can be handled arithmetically. But there are problems with using MAE: it has no direction and doesn't differentiate between an underestimate and an overestimate. To get around these problems, we will also note whether the predictor actually predicted the winner (WIN). For a four-party-forced prediction, the definitions of MAE and WIN are given below.

MAE = Mean Absolute Error = $[|prediction A - result A| + |pred B - result B| + |pred C - result C| + |pred D - result D|]/4$

WIN = Did it predict the winner? = "1" if it did, "0" if it didn't

An example MAE calculation for a four-party-forced prediction of the 2015 election is given below.

Table 1. Example calculation for mean absolute error

Party	Prediction		2015 GB Result		Absolute error
	%	Prop'n	%	Prop'n	
Con	40%	0.4	37.8%	0.378	0.022
Lab	30%	0.3	31.2%	0.312	0.012
Lib	20%	0.2	8.1%	0.081	0.119
Other	10%	0.1	22.9%	0.229	0.129
Total absolute error					0.022 + 0.012 + 0.119 + 0.129 = 0.282
Mean absolute error					0.282 / 4 = 0.0705

THE CAMPAIGN

The campaign started when the election was announced on 18 April and ended when the polls closed at 10pm on 8 June. It was not a homogenous process and can be divided into distinct acts: Joe Twyman of YouGov went for three acts [0612a], but this article will subdivide his first and third acts into two parts each. And the story goes like this...

18 APRIL TO 27 APRIL: CONSERVATIVE ASCENDANCY

This period was marked by a rapid increase in the Conservative party's poll rating from the high 30s/low 40s to the high 40s, occasionally hitting 49% or 50%. Labour remained moribund in the mid to high 20s, and the talk was of a landslide, discussing not just whether Labour would be beaten but whether there would be anything left of the party afterwards. Parliament prorogued on 27 April.

27 APRIL TO 15 APRIL: LABOUR ENTRENCHMENT

On 27 April Labour's poll ratings crept up to 30%. They weren't getting much better but they weren't falling back either, and during this period neither party was budging much: Labour stayed in the high 20s/low 30s, the Conservatives stayed in the mid/high 40s. Labour was still losing but no longer dying in place.

16 MAY TO 22 MAY: MANIFESTOS ARE RELEASED

Between 16 May and 22 May, Labour and the Conservatives both launched their manifestos. Some of Labour's policies were popular, but a Conservative manifesto commitment to change funding for elderly social care proved profoundly unpopular. It proposed paying for long-term care by using the proceeds of the estate after death and was instantly characterised by its opponents as the "dementia tax". The commitment was rapidly disavowed on 22 May but the damage had been done: Labour had begun to creep up to the mid-30s and Conservatives down to the low 40s

22 MAY TO 2 JUNE: A TERRIBLE DISTRACTION

On 22 May 2017 a suicide bomber detonated a bomb outside an Ariana Grande concert in Manchester Arena, killing himself and 22 adults and children. This generated a brief pulse in the Conservative lead lasting two-to-three days. But as the effects dissipated, the polls began to pick up a rise in Labour: 38%, 36%, 37%, settling in the mid-to-high 30s. The Conservatives were still in the mid-to-low 40s but the days of discussing landslide victories were long gone.

On 31 May a leader's debate was held on the BBC, gathering party leaders from the major GB parties (the NI parties did not attend). Corbyn and May had separately decided not to attend but, at the last minute, Corbyn decided to turn up and the debate took place, with the Prime Minister conspicuous by her absence. Two days later, Labour hit 40% in the polls.

3 JUNE TO 8 JUNE: ANOTHER OUTRAGE

On 3 June 2017 three men went on a rampage in London, driving into pedestrians on London Bridge and stabbing random passers-by, killing eight people and injuring 48. They were shot dead. Again there was a brief pulse in the Conservative lead lasting two-to-three days, but when it had ended the polls simply hadn't moved much: Labour in the mid-to-high 30s, Conservatives in the mid-to-low 40s, with Labour briefly in the lead in one poll but not by much.

THE RESULTS

To measure accuracy of predictors against the results we have to define what the results are. We will use the figures from the British Polling Council's (BPC) initial response [0612b] as the GB vote share results, and the figures from the House of Commons Library Commons Briefing Paper CBP-7979 [0622b] as the UK and GB seat results. We use GB vote share because despite making an effort to poll UK-wide for the EU referendum, pollsters yet again mostly polled GB-only [0622a] for the 2017 General Election. We use UK seat results because no Labour, Conservative nor Liberal Democrat were elected to seats in Northern Ireland, so we can switch from UK to GB by just subtracting 18 from the "Other" category. The situation is further complicated by the Speaker (currently the Conservative John Bercow), who is counted as "Other" by some sources and Conservative by others. This article will adhere to the convention that the Speaker is "Other" which drops the number of Conservative seats by 1 to 317. This gives us the following:

Table 2. Results used by this article as the results of the 2017 General Election

Category	Con	Lab	Lib	Oth	Total	Source
GB vote share	43.5%	41.0%	7.6%	7.9%	100.0%	[0622d]
UK seats	317	262	12	60	651	[0622b]

Having established the methodology and results, we will now examine how each category of predictor performed.

THE POLLSTERS

Pollsters are people who do opinion polling in the UK. They range in size from wealthy individuals commissioning polls to multinational companies for whom political opinion polling is a loss-leader attracting clients to their main operations. Just before the Prime Minister's announcement, the individual pollsters were focussing on polling and the changes required since the 2015 polling debacle. Martin Boon of ICM had just published a thought piece on emotional measurement [0604a]. Joe Twyman of YouGov was advertising a course at Essex Summer School on "Achieving Impact through Quantitative Social Science" [0604b]. His colleague Anthony Wells had just posted an item on his blog about that day's YouGov poll. Ben Page of Ipsos Mori was spending Easter in Palermo [0604d], and his colleague Bobby Duffy had polled on society [0604e] the previous week.

Whereas in previous years a poll took days to set up and complete, nowadays pollsters poll frequently on a wide variety of matters and can organise a new one very quickly. They adjusted to the announcement rapidly and, before the day was out, Bobby Duffy had published an item in the *Independent* [0604f], Anthony Wells had put up another blog entry [0604g], and ICM had released a flash poll [0604g].

In total, 14 pollsters polled on the 2017 General Election. We selected the five most frequent - Panelbase, Opinium, Survation, ICM, and YouGov - and tracked them through the campaign. Their MAE and WIN over time looked like this:

Figure 1: MAE for vote share polls between Wednesday 26 April 2017 and Wednesday 7 June 2017

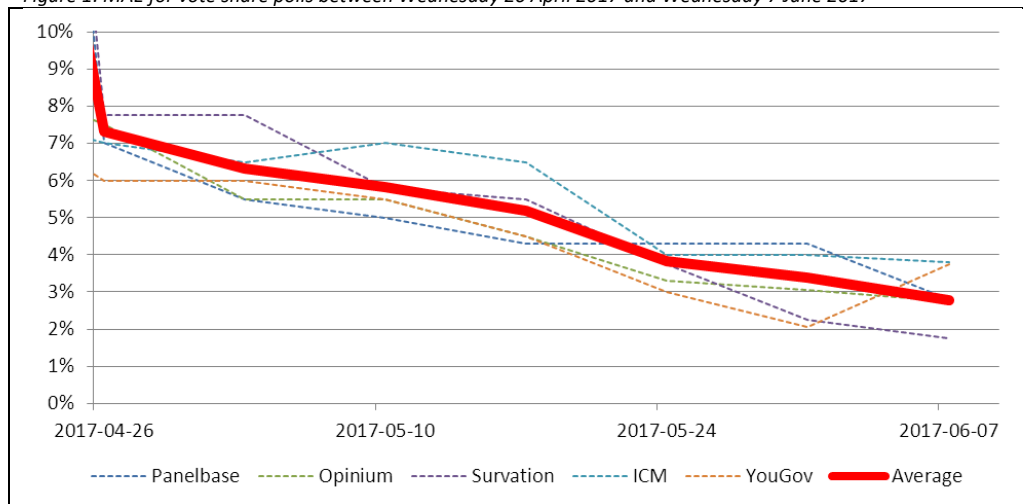


Table 3: Total WIN for vote share polls between Wednesday 26 April 2017 and Wednesday 7 June 2017

Date	27 Apr	3 May	10	17	24	31	7 Jun
Number correctly predicting Conservative most votes	5/5	5/5	5/5	5/5	5/5	5/5	5/5

This was not a good performance by the pollsters and in different circumstances, such as the 2015 election, they would have been considered a failure. But the short notice, the fact that they predicted the winner correctly and got the parties broadly right (except for Labour) muted criticism. This phenomenon of forgiving poor predictors if they correctly predict the winner has been noted before by us with the 2014 Scotland Referendum [0625a] and it will be seen again in this article.

THE BOOKMAKERS

Betting on political events has become more popular in the UK in recent years, and a wide variety of bookmakers now offer bets, whether fixed-odds or spreads. They range from longstanding High Street bookmakers to online operations based in Gibraltar [0610c]. The best-known website is probably the summary website Oddschecker.com or Mike Smithson's PoliticalBetting.com, better known as "PB". Just before the announcement, the individual bookies were considering the upcoming first round of the French Presidential election on 23 April, while Mike Smithson's site was discussing the upcoming Manchester Gorton by-election [0610a] and the chances of Francois Fillon in the French Presidential [0610b].

Bookmakers can react very quickly to upcoming events and many hold open political betting markets throughout the year, so the announcement was transparent from their point of view: the odds were already up or (in the case of constituency betting) would be up shortly afterwards

In total about 25 bookies offered on the 2017 General Election. For ease of calculation, we selected 10 Bet, Bet365, NetBet and Winner, and tracked them through the campaign. Their MAE and WIN over time looked like this:

Figure 2: MAE for odds of overall majority between Wednesday 26 April 2017 and Wednesday 7 June 2017

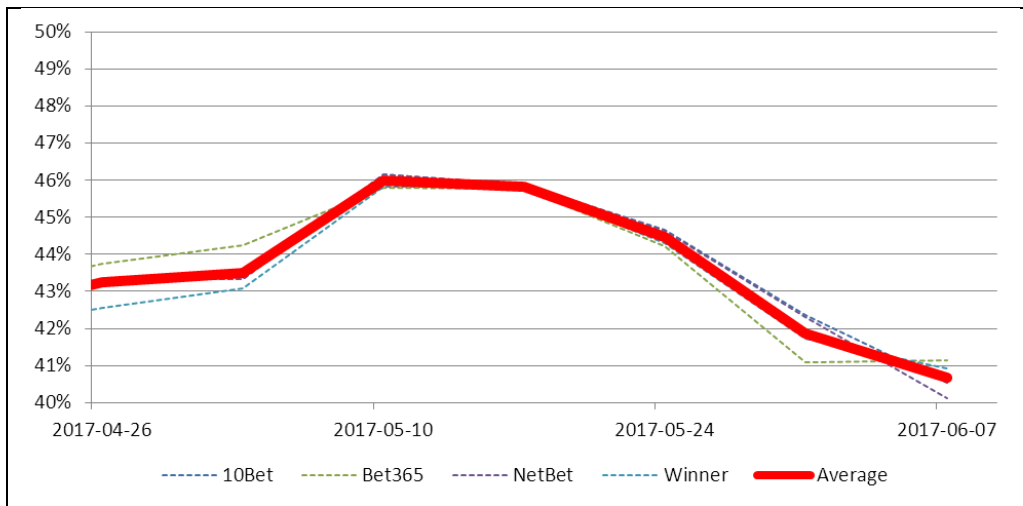


Figure 3: MAE for odds of most seats between Wednesday 26 April 2017 and Wednesday 7 June 2017

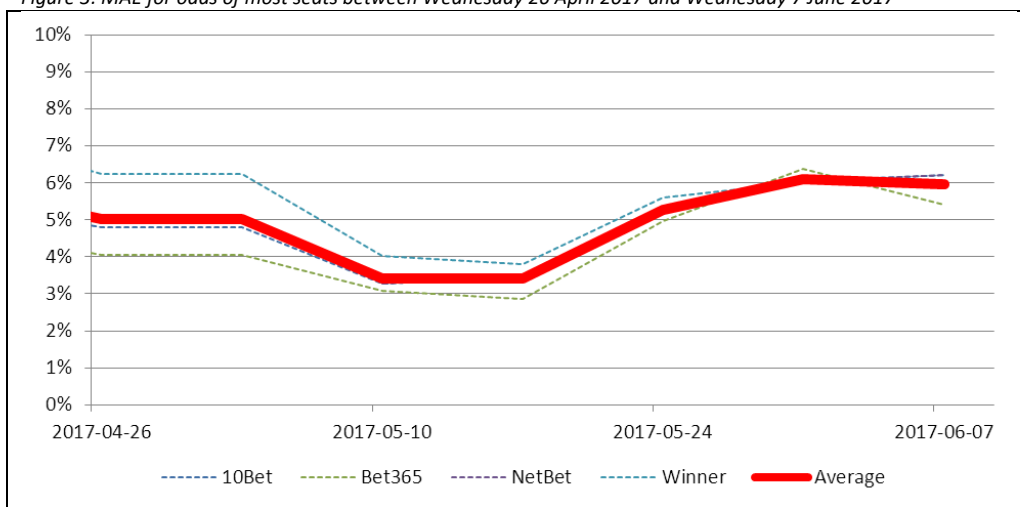


Table 4: Total WIN for selected odds between Wednesday 26 April 2017 and Wednesday June 7 2017

Date	27 Apr	3 May	10	17	24	31	7 Jun
Number correctly predicting no overall majority (NOM)	0/4	0/4	0/4	0/4	0/4	0/4	0/4
Number correctly predicting Conservatives most seats	3/3	3/3	4/4	4/4	4/4	4/4	4/4

As repeatedly stated, the bookmaker odds are not good predictors. The only thing stopping the MAE for overall majority climbing higher was that the bookmakers correctly identified the probability of a Labour or any other party majority to be effectively zero, so reducing the average error. The odds on most seats performed better, but were still worse than the non-bookmaker predictors.

THE ACADEMICS AND ANALYSTS

Over the years an entire ecology of people has grown up to study elections and polls, and to make predictions. They range from academics such as Stephen Fisher at Oxford and Will Jennings at Southampton, to analysts such as Matt Singh and Martin Baxter. The best-known website is probably Martin Baxter's ElectoralCalculus.co.uk, or Fisher's ElectionsEtc.com or The Polling Observatory's posts at <https://sotonpolitics.org/tag/polling-observatory/>. Just before the announcement, the academics and analysts were making preparations for the scheduled 2020 election or thinking about other upcoming elections.

Academics and analysts generally took a few days to get their things in order to start making predictions, though some predictions came in pretty fast: Matthew Lebo (Stony Brook University) had coincidentally made a prediction the previous day [0611a] and Martin Baxter keeps his predictions live and had one up by the end of the 19 April [0611e]. The rest trickled in thereafter, with the Polling Observatory putting the band back together (Ford/Jennings/Pickup/Wlezien) and Chris Hanretty resurrecting the "ElectionForecast" site as a solo project without Benjamin Lauderdale and Nick Vivyan.

There weren't that many academics and analysts making predictions, so we captured what we could. We seriously considered Polling Observatory but they didn't do seat predictions. So we selected Baxter, Fisher, Lebo, and Hanretty/Election Forecast, and tracked them through the campaign. Their MAE and WIN over time looked like this:

Figure 4: MAE for seat predictions between Wednesday 26 April 2017 and Wednesday 7 June 2017

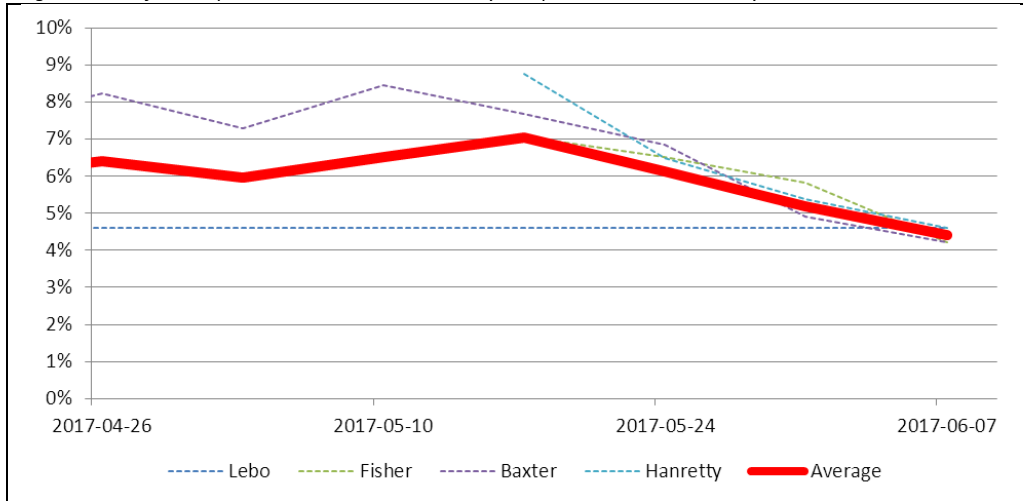


Table 5: Total WIN for seat predictions between Wednesday 26 April 2017 and Wednesday 7 June 2017

Date	27 Apr	3 May	10	17	24	31	7 Jun
Number correctly predicting Conservatives most seats	4/4	4/4	4/4	4/4	4/4	4/4	4/4

This was an indifferent performance by the academics and analysts. We were torn between congratulating them for getting the predictions set up and running in such a short time, and criticising the final outcome. We settled for acknowledging a danger that is not often remarked upon: it is relatively straightforward to make vote share predictions, but it is far more difficult to make seat predictions.

THE PREDICTION MARKETS AND SPREAD BETTING

A prediction market is a market for predictions: a concept is introduced and a buy price and sell price are assigned. The buy and sell prices vary with market demand. A spread betting firm (in the meaning commonly held in Britain) looks very similar, with a varying buy and sell price, with the concept of “financial spread betting” becoming increasingly important. A third category – exchange betting – has a back and lay price and, again, appears similar but has differences. The distinction between the categories deserves an article of its own but for the moment we will simply note that although the differences may be important due to differing structure, legislation, oversight and taxation, for practical purposes this article will not distinguish between them. The best-known spread betting website is probably SpreadEx.com or SportingIndex.com, the latter commonly known as “SPIN”. The best-known prediction market is probably PredictIt.com. The best known exchange betting market is probably Betfair Exchange, not to be confused with the fixed-odds Betfair Sportsbook. Just before the announcement they were considering the French Presidential election and currency and share movements [0609a][0609b].

The firms by their nature can open markets quickly, and they put markets up on the election almost immediately. Unfortunately we could only capture a few, so our choices were constrained to what we could get. PredictIt did buy/sells on specific seat ranges and “next PM”, but both these markets were difficult to handle (“next PM: Boris Johnson?”) so we discarded it. Odds on number of seats from Betfair Exchange were also difficult to handle, so we discarded that as well. Consequently we selected SpreadEx, SPIN, and IG Index. Their MAE and WIN over time looked like this:

Figure 5: MAE of seat ranges between Wednesday 26 April 2017 and Wednesday 7 June 2017

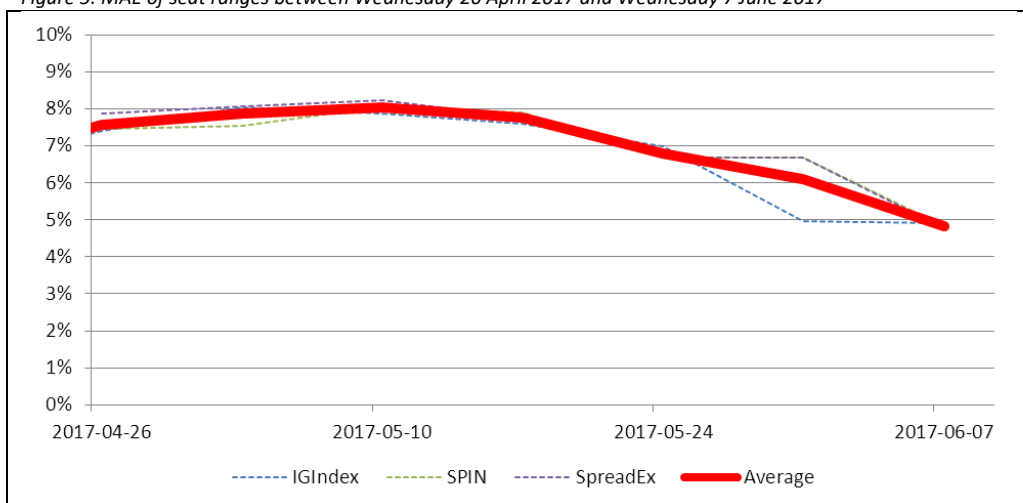


Table 6: Total WIN of seat ranges between Wednesday 26 April 2017 and Wednesday 7 June 2017

Date	27 Apr	3 May	10	17	24	31	7 Jun
Number correctly predicting Conservatives most seats	3/3	3/3	3/3	3/3	3/3	3/3	3/3

Again, this was not a good performance. But it does illustrate the advantages and disadvantages of the “wisdom of crowds approach”: without a singular conscious oversight, the mass of punters converged to a solution that was very like the academic and analyst predictions. But similarly it wasn’t that accurate: the crowd is not as wise as its reputation suggests. The question lingers: if the academics and analysts had not predicted or the pollsters not polled, would this have been different? We will never know.

THE POLLSTER-MODELLERS

During the campaign an interesting hybrid appeared – the pollster-modellers. Models for election prediction were hardly unknown, but pollsters had usually issued... well, polls, as the name implies. But in this campaign YouGov and Lord Ashcroft both issued Multilevel Regression and Poststratification models: “MRP” for short, or “Mister P” [0617a] colloquially. MRP is a modelling technique that takes the output from an opinion poll, traces the responses back to the people polled, notes their demographic and geographic characteristics, assigns a probability to those characteristics (“90% likelihood of male over 75 earning 60K in Surrey responding Con”), then weights that by the number of people with those characteristics in each area [0617b] to come up with a prediction.

Because of interest generated by this methodology we separated the YouGov model and Ashcroft model out. Unlike the other predictors, they were not present at the start of this election period, and their predictions were not sampled prospectively but reconstructed retrospectively. Their MAE and WIN over time looked like this:

Figure 5: MAE for seat predictions between Wednesday 26 April 2017 and Wednesday 7 June 2017

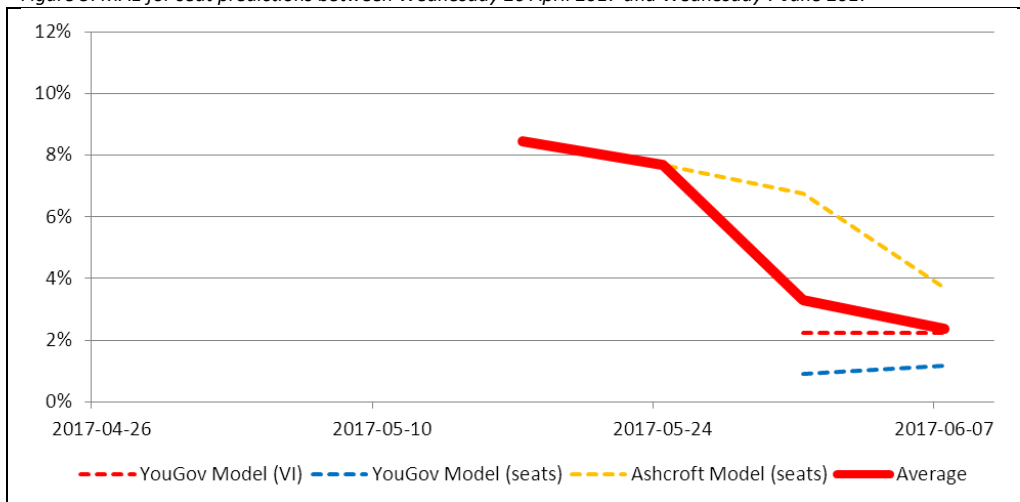


Table 7: Total WIN for seat predictions between Wednesday 26 April 2017 and Wednesday 7 June 2017

Date	27 Apr	3 May	10	17	24	31	7 Jun
Number correctly predicting Conservative most seats	2/2	2/2	2/2	2/2	2/2	2/2	2/2
Number correctly predicting Conservatives most votes	1/1	1/1	1/1	1/1	1/1	1/1	1/1

This is much better: the YouGov model was a good predictor for vote share and an adequate one for seats. However praise must be counterbalanced by noting the somewhat poorer performance of the Ashcroft model. MRP is not a panacea and decisions such as the method of turnout prediction remain vital.

THE LAST DAY

By polling day, all the predictions were in. A full list of the final predictions is in Appendix 2. Displayed graphically, they look like this:

Figure 7: Average MAE of latest predictors for each category at Thursday 8 June 2017

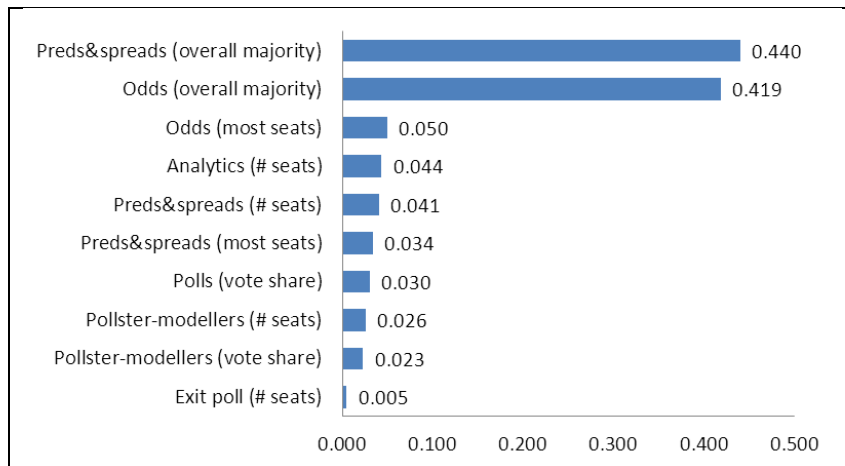


Table 8: Total number of predictors correctly predicting the winner for each category at Thursday 8 June 2017

Predictor	Number of predictions	Average MAE	WIN
Exit poll (# seats)	1	0.005	1/1
Pollster-modellers (vote share)	1	0.023	1/1
Pollster-modellers (# seats)	2	0.026	2/2
Polls (vote share)	11	0.030	11/11
Preds&spreads (most seats)	2	0.034	2/2
Preds&spreads (# seats)	3	0.041	3/3
Analytics (# seats)	11	0.044	11/11
Odds (most seats)	16	0.050	16/16
Odds (overall majority)	17	0.419	0/17
Preds&spreads (overall majority)	2	0.440	0/2

We can only repeat what we said earlier: this was not a good performance. There were some bright points, such as the performance of the pollster-modellers, and the outstanding performance of the exit poll (commissioned by the UK's main broadcasters). The failure of the overall majority odds is not a new phenomenon: it is a blunt instrument that should not be relied upon for close elections. The polls weren't actually that bad. But even so, you would not describe this as a spectacular success.

CONCLUSIONS

In 2010 the pollsters were reasonably settled as to their methodologies and achieved a decent result. In 2015 that comfort was shaken by their collective failure, and the 2016 EU referendum performance did little to help calm nerves. By 2017, the Sturgis inquiry had reported and the pollsters had a good idea of what they intended to do, and the snap election gave them a legitimate excuse for failure if things went wrong. It was an atmosphere that encouraged experimentation, and they promptly obliged [0617f] with differing methods, a wide range of predictions, no herding and no bad-tempered Twitter spats. The day before was enlivened by a "stooshie" (their word) in the Survation offices over their final poll numbers: in the end they published and were not damned, getting it broadly right [0617h].

The bookies were expanding into a vacuum. The expansion of political betting in recent years meant that many odds were offered on many combinations: overall majority, most seats, next PM, constituency betting, financial indices, spread betting, prediction markets, fixed-odds gambling – a rich ecology where, whatever you wanted, there was a price to match and somebody willing to take your money. The problem was not acquiring data, it was deciding which data to discard.

By comparison, the academics didn't have a bad election, but it wasn't a good one either. It takes time to set up a model and they had short notice, so filler metrics such as polls-of-polls occasionally cropped up. But eventually there was a reasonable spread of predictions.

The most obvious trend was the triumph of Big Data. A 1980s poll would have been instantly recognisable to a classical statistician: a questionnaire taken face-to-face by a researcher, answers accumulated over several days and a mean produced, with the more daring assaying a confidence interval. But since then, face-to-face gave way to telephone polling, which is now giving way to online panel polling. The word "statistics" became conflated with "modelling". The increased availability of open-source software and advances in storage and data access meant the amount of data coming in became bigger and the techniques and skilled workforce needed to handle them became available. The statistician running proprietary software, such as SAS or SPSS on a desktop, is being displaced by the modeller running open-source R on a laptop. The implications of the election were twofold, one good, one bad.

The first was the success of YouGov's Multilevel Regression and Poststratification model. This model caused a negative reaction from partisans who wished to disbelieve its prediction of a hung parliament, and puzzlement from laymen who didn't know what it meant ("So...it's a poll, yes?" "Not quite"). But it was unambiguously successful.

The second was sadder: the failure of turnout models. There had been quite the fad for turnout models since 2015, with ComRes sticking to its ComRes Voter Turnout Model despite its performance in the referendum, and ICM also using a turnout model. But models will

always have an Achilles' heel: they are vulnerable to historically rare changes in behaviour. So when people began to notice that turnout models were giving the Conservatives a higher lead over Labour than self-reported likelihood to vote, it led to speculation regarding a mode effect. Sturgis and Jennings thought there might be such an effect [0617c], as did Anthony Wells [0617d][0617f], although he also thought it depended on whether non-BPC companies like SurveyMonkey were included [0617e]. In the event, turnout models did not perform well, but given that there are known problems with self-reported likelihood to vote we assume that turnout models will not be abandoned but simply recalibrated.

Was this another polling debacle? The BPC acknowledges that the final polls were not ideal [0618e] and we note that our MAEs were over the 2% threshold deemed acceptable for a four-party-forced. But is this sufficient? In previous articles we hypothesised that MAE increases with the number of parties predicted, although it's more accurate to say the MAE changes as the number of parties predicted changes. So the concept of a constant threshold disappears. But we can compare the BPC's own figures from 2017 to its own figures in 2015, as follows:

Table 9: BPC self-assessment of final polls for 2015 and 2017. See Appendix 4 and 5 for calculations

Format	MAE of BPC self-assessment 2015	MAE of BPC self-assessment 2017
Six-party-forced	1.47	1.7
Four-party-forced	2.1	2.6
Source	[0618d]	[0618e]

So regardless of forced format, the poll predictions have gotten worse.

As stated at the outset, this was one of the most interesting elections in a long while. The implications for the country will be important, as the UK Government can now only command a majority in the new Parliament with the help of Northern Ireland's DUP, so any Brexit negotiated settlement will need to be broadly palatable. But from a non-partisan point of view, this year's vote gave indications to the future course of election prediction as new techniques and markets open up. Regardless of the political outcomes, the future for election prediction appears bright... although whether it will also be accurate is yet to be determined.

NOTES AND REFERENCES

[0507]

- [0507s] "Theresa May's general election statement in full", 18 April 2017 <http://www.bbc.co.uk/news/uk-politics-39630009>

[0604]

- [0604a] 12 APRIL 2017 HEARTS AND MINDS Martin Boon <https://www.research-live.com/article/opinion/hearts-and-minds/id/5021112>
- [0604b] <https://twitter.com/JoeTwyman/status/849217665067606016> or <https://t.co/ebXKhebDQp> or <https://yougov.co.uk/find-solutions/sectors/political/summer-school/>
- [0604c] <http://ukpollingreport.co.uk/blog/archives/9832#comments>
- [0604d] <https://twitter.com/benatipsosmori/status/853841891376365568>
- [0604e] <https://twitter.com/BobbyIpsosMORI/status/851015012961353728>
- [0604f] <http://www.independent.co.uk/voices/general-election-brexit-polls-pollsters-most-unpredictable-a7689551.html>
- [0604g] <http://ukpollingreport.co.uk/blog/archives/9833#comments>

[0609]

- [0609a] <https://twitter.com/IGSquawk/status/854065772385185793>
- [0609b] <https://twitter.com/IGSquawk/status/854061973859962880>

[0610]

- [0610a] <http://www2.politicalbetting.com/index.php/archives/2017/04/17/the-canvas-data-that-proved-to-be-spot-on-in-richmond-suggests-labour-could-be-in-trouble-in-manchester-gorton/>
- [0610b] <http://www2.politicalbetting.com/index.php/archives/2017/04/17/fillon-moves-to-second-favourite-in-the-french-race-after-new-poll-has-him-within-a-point-of-le-pen/>
- [0610c] <http://www.independent.co.uk/money/tax/jackpot-bookies-avoid-1bn-tax-8476143.html>
- [0610d] <http://www2.politicalbetting.com/index.php/archives/2017/06/11/the-2017-general-election-the-most-profitable-general-election-ever-for-the-bookies/>

[0611]

- [0611a] <https://twitter.com/MattLebo4/status/854417336278949889>
- [0611b] <https://twitter.com/ProfJaneGreen/status/852181899065909252>
- [0611c] <https://twitter.com/drjennings/status/853989959656771585>
- [0611d] <https://twitter.com/StephenDFisher/status/849552592195682308>
- [0611e] <http://archive.is/JOPEs>

[0612]

- [0612a] <https://yougov.co.uk/news/2017/05/30/election-campaign-story-three-acts-so-far/>
- [0612b] <http://www.britishpollingcouncil.org/general-election-8-june-2017/>

[0617]

- [0617a] <https://thepoliticalmethodologist.com/2013/10/14/being-careful-with-multilevel-regression-with-poststratification/> see also <http://andrewgelman.com/2013/10/09/mister-p-whats-its-secret-sauce/>
- [0617b] http://www.princeton.edu/~jkastell/MRP_primer/mrp_primer.pdf, see also www.princeton.edu/~jkastell/mrp_primer.html
- [0617c] <https://sotonpolitics.org/2017/06/04/will-turnout-weighting-prove-to-be-the-pollsters-achilles-heel-in-ge2017/>
- [0617d] <http://ukpollingreport.co.uk/blog/archives/9899#comments> see also <http://ukpollingreport.co.uk/blog/archives/9895>
- [0617e] <http://ukpollingreport.co.uk/blog/archives/9915#comments>
- [0617f] <https://yougov.co.uk/news/2017/06/01/pollsters-experimental-election/>
- [0617g] <https://twitter.com/DamianSurvation/status/872573142186815488>
- [0617h] <http://survation.com/wp-content/uploads/2017/06/Survation-GE2017-Final-Poll-2d7l9l8.pdf>, see also <http://ukpollingreport.co.uk/blog/archives/9909#comments>

[0618]

- [0618a] "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/>

- [0618b] <http://www.britishpollingcouncil.org/accuracy-of-the-final-polls/>
- [0618c] <http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/firth/exit-poll-explainer/>
- [0618d] <http://www.britishpollingcouncil.org/general-election-7-may-2015/>
- [0618e] <http://www.britishpollingcouncil.org/general-election-8-june-2017/>

[0625]

- [0625a] "Forecast error: European referenda, past and present". Contents by Hill sole, Significance, 25 May 2016. See <https://www.significancemagazine.com/politics/252-forecast-error-european-referenda-past-and-present> or https://www.statslife.org.uk/files/Forecast_Error_-_EU_referenda2.pdf

APPENDIX 1: ALL PREDICTIONS ON WEDNESDAY 26 APRIL 2017

POLLSTERS

A selection of polls on/before Wednesday 26 April 2017 looks like this:

Table 10: Polls with fieldwork ending between 2017-04-18 and 2017-04-26

End of fieldwork	Pollster/client	CON	LAB	LIB	OTH
26/04/2017	YouGov/The Times	45%	29%	10%	16%
25/04/2017	Ipsos MORI/Evening Standard	49%	26%	13%	12%
24/04/2017	ICM/The Guardian	48%	27%	10%	15%
24/04/2017	Panelbase	49%	27%	10%	14%
24/04/2017	Kantar Public	46%	24%	11%	19%
23/04/2017	Norstat/Sunday Express	42%	26%	10%	22%
22/04/2017	Survation/Mail on Sunday	40%	29%	11%	20%
21/04/2017	YouGov/Sunday Times	48%	25%	12%	15%
21/04/2017	ICM/ITV	48%	26%	10%	16%
20/04/2017	ORB/Daily Telegraph	44%	29%	8%	19%
20/04/2017	ComRes/Sunday Mirror	50%	25%	11%	14%
20/04/2017	Opinium/Observer	45%	26%	11%	18%
20/04/2017	YouGov	44%	25%	12%	19%
19/04/2017	YouGov/The Times	48%	24%	12%	16%
18/04/2017	ICM/The Guardian	46%	25%	11%	18%

BOOKMAKERS

A selection of bookies on/before Wednesday 26 April 2017 looks like this:

Table 11: Odds on an overall majority at 2017-04-26: see <http://archive.is/glgeB>. All odds in decimal format

	Con	NOM	Lab	AOM	Lib	Gre
10Bet	1.12	7	34	151		
Bet365	1.11	7.5	41	151		
Betfair Exchange	1.13	9	52	342		
Betfair Sportsbook	1.13	6	17		51	301
Betfred	1.13	7.5	34		101	
Betstars	1.11	6.5	41	101		
BetVictor	1.14	5.5				
Betway	1.12	5.5	26		51	
Boylesports	1.13	6	21		51	301
Coral	1.12	8	21		101	1001
Ladbrokes	1.13	7	21		67	1001
Marathon		6	26			
Netbet	1.12	7	34	151		
PaddyPower	1.07	7	21		101	1001
SkyBet	1.11	7	34		151	1001
Stan James		6	34	101		
WilliamHill	1.08	7	21		41	
Winner	1.12	6	26	51		

Table 12: Odds on most seats at 2017-04-26: see <http://archive.is/DPORP>. All odds in decimal format

	Con	Lab	Lib	Gre	UKI	AOP
10Bet	1.05	13	51	401	501	
32Red	1.06	13	41	501	251	
888Sport	1.06	13	41	501	251	
Bet365	1.04	15	67	501	751	
Betfair Exchange	1.06	17.6	114		903	950
Betfair Sportsbook	1.04	13	67	501	501	
Betfred	1.05	15	41	501	501	
Betstars	1.04	15	34		126	
BetVictor	1.05	11	23	201	101	
Betway	1.05	13	23	501	251	
Boylesports	1.06	9	34	201	51	
Bwin	1.05	13	26	501	81	
Coral	1.05	13	51	501	201	
Ladbrokes	1.06	9	34	501	101	
Marathon	1.05	10	26	501	501	
PaddyPower	1.04	13	67	501	501	
SkyBet	1.04	15	51	501	501	
SportingBet	1.05	13	41	501	501	
Stan James	1.05	15	34	501	251	
Unibet	1.06	13	41	501	251	
WilliamHill	1.04	13	26	201	67	
Winner	1.05	11	26	301	301	

ACADEMICS AND ANALYSTS

Table 13: Predictions made by academics and analysts by 2017-04-26.

	Date made	CON	LAB	LIB	Oth	Source
Martin Baxter vote	19/04/2017	0.442	0.245	0.102	0.211	http://archive.is/JOPEs
Martin Baxter seats	19/04/2017	392	170	9	80	http://archive.is/JOPEs
Matthew Lebo	18/04/2017	372	202	10? ("maybe a bit more")	67	http://archive.is/mY6Hr
Stephen Fisher	24/04/2017	390	181	9	71	http://archive.is/9eill
Martin Baxter vote	26/04/2017	0.477	0.254	0.110	0.159	http://archive.is/JC4Xj
Martin Baxter seats	26/04/2017	414	158	9	70	http://archive.is/JC4Xj
Polling Observatory	19/04/2017	0.43	0.257	0.105	0.208	http://archive.is/QARWR

PREDICTION MARKETS AND SPREAD BETTING

The prediction markets on/before Wednesday 26 April 2017 looked like this:

Table 14: preds and spreads at 2017-04-26.

Number of seats	Con	Lab	Lib	UKI	SNP	Note
Spreadex seats	384-390	158-164	25-28	0.25-1.25	46-49	At 27/04/2017
Spin seats	384-390	162-168	26-29	0.25-1.25	47-50	At 27/04/2017
IG Index seats	386	165.5	27.8	?	?	Closing price
Most seats	Con	Lab	Lib	UKI	AOP	
Spreadex most seats	91.6-98.1	3.4-6.7	0.3-0.7	0-0.2	0-0.2	At 27/04/2017
Spin most seats	90-96	4.5-7.5	0.5-1.5	0-0.1	0-0.1	At 27/04/2017
Overall majority	Con	Lab	AOP	NOM		
Spreadex overall majority	85.7-93.4	0.9-2.5	0-0.3	6.6-11.8		At 27/04/2017
Spin overall majority	84-91	1-2.5	0.1-0.5	8-13		At 27/04/2017
PM post-election	May	Corbyn	Farron	Boris		
PredictIt	92	8	0.5	0.5		

APPENDIX 2: ALL PREDICTIONS ON LAST DAY

The final predictions are the latest ones on or before the polling stations closed at 10pm on Thursday June 8 2017

POLLSTERS

Table 15: Latest polls on/before 2017-06-08. ORB, Norstat not included because latest poll was on 2017-06-01

Pollster	Last day of fieldwork	Con	Lab	UKIP	Lib Dem	SNP	Green	Others	Source
BMG	07/06/2017	0.46	0.33	0.05	0.08	0.04	0.03	0.01	[0619a]
ComRes	07/06/2017	0.44	0.34	0.05	0.09	0.04	0.02	0.02	[0619b]
ICM	07/06/2017	0.46	0.34	0.05	0.07	0.05	0.02	0.01	[0619c]
Ipsos MORI	07/06/2017	0.44	0.36	0.04	0.07	0.05	0.02	0.02	[0619d]
Kantar Public	07/06/2017	0.43	0.38	0.04	0.07	0.04	0.02	0.02	[0619e]
Opinium	07/06/2017	0.43	0.36	0.05	0.08	0.05	0.02	0.01	[0619f]
Panelbase	07/06/2017	0.44	0.36	0.05	0.07	0.04	0.02	0.02	[0619g]
Qriously	07/06/2017	0.39	0.41	0.04	0.06	0.03	0.07	0.00	[0619h]
Survation	07/06/2017	0.41	0.40	0.02	0.08	0.04	0.02	0.03	[0619i]
SurveyMonkey	06/06/2017	0.42	0.38	0.04	0.06	0.03	0.02	0.05	[0619j]
YouGov	07/06/2017	0.42	0.35	0.05	0.10	0.05	0.02	0.01	[0619k]

- [0619a] http://www.heraldsotland.com/news/15334427.Britain_at_crossroads_as_eve_of_poll_survey_suggests_May_heading_for_landslide_victory/
- [0619b] http://www.comresglobal.com/wp-content/uploads/2017/06/Independent_FINAL-VI-and-Political-Poll_7th-June-2017_130392.pdf
- [0619c] https://www.icmunlimited.com/wp-content/uploads/2017/06/2017_guardian_prediction_PRELIM_1500.pdf
- [0619d] <https://www.ipsos.com/sites/default/files/2017-06/pm-election-2017-final-tables.pdf>
- [0619e] <http://uk.kantar.com/ge2017/2017/conservatives-heading-for-election-win/>
- [0619f] <http://opinium.co.uk/political-polling-4th-june-2017/>
- [0619g] <http://www.panelbase.com/media/polls/W10470w7tablesforpublication.pdf>
- [0619h] <http://www.wired.co.uk/article/election-polls-labour-conservative-winner>
- [0619i] <http://survation.com/wp-content/uploads/2017/06/Survation-GE2017-Final-Poll-2d7f9f8.pdf>, see also <https://twitter.com/Survation/status/872586205006888961>
- [0619j] <https://www.thesun.co.uk/news/3739747/theresa-may-poll-terrorism-election-london-bridge/>
- [0619k] https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/d8zsb99eyd/TimesResults_FINAL%20CALL_GB_June2017_W.pdf

BOOKMAKERS

Table 16: Odds on an overall majority at approx. 8:30pm 2017-06-08.

Bookmaker	Date	Con	NOM	Lab	AOM	Gre	Lib	UKI	Source
Bet365	08/06/2017	1.16	6.5	17	1001				[0619l]
Betdaq	08/06/2017	1.15	7.21	25	97				[0619l]
Betfair Exchange	08/06/2017	1.17	7.46	29					[0619l]
Betfair Sportsbook	08/06/2017	1.16	5.5	19		1001	501	1001	[0619l]
Betfred	08/06/2017	1.2	6.5	17			1001	2501	[0619l]
Betstars	08/06/2017	1.2	4.75	15	1001				[0619l]
BetVictor	08/06/2017	1.17	5.5	15	1001				[0619l]
Boylesports	08/06/2017	1.17	5.5				1001		[0619l]
Coral	08/06/2017	1.2	6	13					[0619l]
Ladbrokes	08/06/2017	1.2	6	13					[0619l]
Marathon	08/06/2017	1.2	5	15		501	201	401	[0619l]
PaddyPower	08/06/2017	1.17	5	17		1001	501	1001	[0619l]
SkyBet	08/06/2017	1.18	5.5	17		2501	1501	2501	[0619l]
SportingBet	08/06/2017	1.17	4.5	15			1501	1501	[0619l]
Stan James	08/06/2017	1.2	6	17	1001				[0619l]
Unibet	08/06/2017	1.2	6	18		2501		2501	[0619l]
WilliamHill	08/06/2017	1.22	5	9			501	1001	[0619l]

- [0619l] <http://archive.is/vbSEk>

Table 17: Odds on most seats at approx. 8:30pm 2017-06-08. Unibet & Winner discarded due to unusual combinations

Bookmaker	Con	Tie	Lab	Lib	UKI	Gre	AOP	Source
Betdaq	1.05		16	970	970		970	[0619m]
Betfair Exchange	1.06		15.7	950	950		950	[0619m]
Betfair Sportsbook	1.05		11	1501	2001	2001		[0619m]
Betfred	1.07		9	751	2001	2001		[0619m]
Betstars	1.06		9.5	501	1001	1001		[0619m]
BetVictor	1.05		10	1001	1001	1001		[0619m]
Betway	1.07		7	501	1001	1001		[0619m]

Coral	1.07		8	501	1001	1001		[0619m]
Ladbrokes	1.07		8	501	1001	1001		[0619m]
Marathon	1.06		9	301	301	1001		[0619m]
Matchbook	1.05		15.9	476	476		298	[0619m]
PaddyPower	1.05		10	501	1001	1001		[0619m]
SkyBet	1.05		10	1001	2001	2001		[0619m]
SportingBet	1.05		9.5	1001	1251	1501		[0619m]
Stan James	1.05		10	501	1001	1001		[0619m]
Unibet	1.06	6.5	9.5	1001	11	1.2		{0619m}
WilliamHill	1.06		9	501	1001	1001		[0619m]
Winner	1.06		8					{0619m}

- [0619m] <http://archive.is/IKtNI>

ACADEMICS AND ANALYSTS

Table 18: Latest seat predictions made by academics & analysts before 10pm 2017-06-08.

Academics and analytics	Date	Con	Lab	SNP	LIB	UKI	Source
Michael Thrasher	05/05/2017	349	215	54	9	0	[0619n]
Stephen Fisher et al (polls-based)**	08/06/2017	349	223	47	9	0	[0619y]
Stephen Fisher et al (combined)*	08/06/2017	358	214	47	8	0	[0619z]
Britain Elects	08/06/2017	356	219	43	9	0	[0619n]
Lord Ashcroft*	08/06/2017	363	217	44	5	0	[0619o]
Electoral Calculus*	08/06/2017	358	218	49	3	0	[0619p]
Chris Hanretty	08/06/2017	366	207	46	7	0	[0619n]
Matt Singh	07/06/2017	374	207	43	7	0	[0619n]
Nigel Marriott	06/06/2017	375	202	48	2	0	[0619n]
Iain Dale	05/06/2017	386	178	47	12	0	[0619n]
Matthew Lebo	05/06/2017	372	202	about the same	maybe a bit better	n/a	[0619q]

* In his tweet Hix gave the penultimate prediction. We have used the latest one before the polls closed

** Fisher et al issued two seat predictions on 8 June. In his tweet Hix captured only one of them

- [0619n] <https://twitter.com/simonjihix/status/872724388768092160>
- [0619o] <http://lordashcroftpolls.com/2017/06/estimated-conservative-majority-rises-final-ashcroft-model-update/>
- [0619p] <http://archive.is/b2aSM>
- [0619q] <https://twitter.com/MattLebo4/status/871884896473690112>
- [0619y] <https://electionsetc.com/2017/06/08/polls-based-forecast-for-the-2017-british-general-election/>
- [0619z] <https://electionsetc.com/2017/06/08/final-combined-forecast-for-ge2017/>

PREDICTION MARKETS AND SPREAD BETTING

Table 19: Latest preds and spreads of number of seats before polls closed on 2017-06-08

Number of seats	Con	Lab	Lib	UKI	SNP	Source
Spreadex seats	356-362	206-212	11.5-13.5	0.1-0.5	43.5-45.5	[0619r]
Spin seats	357-362	206-211	10.5-12.5	0.2-0.6	44.5-46.5	[0619r]
IG Index seats	357-363	209-214	10 to 12	0.05-0.35	43.5-45.5	[0619s]

- [0619r] Contemporaneous notes by author
- [0619s] <http://archive.is/6F966>

Table 20: Latest preds and spreads of most seats before polls closed on 2017-06-08

Most seats	Con	Lab	Lib	UKI	AOP	Source
Spreadex most seats	90-98.1	3.8-5.9	0-0.2	0-0.2	0-0.2	[0619r]
Spin most seats	88-95	05 to 11	0-0.5	0-0.1	0-0.1	[0619r]

Table 21: Latest preds and spreads of overall majority before polls closed on 2017-06-08

Overall majority	Con	Lab	AOP	NOM	Source
Spreadex overall majority	81.8-90	2.4-4.8	0-0.2	8.3-14.3	[0619r]
Spin overall majority	81-87	01 to 05	0-0.5	10 to 16	[0619r]

Table 22: Latest preds and spreads of next PM before polls closed on 2017-06-08

PM post-election	May	Corbyn	Farron	Boris	Source
PredictIt	89-91	08 to 09	0-1	0-1	[0619r]

POLLSTER-MODELLERS

Table 23: Latest MRP models before polls closed on 2017-06-08

MRP model	Date	Con	Lab	UKIP	LD	SNP	PC	Green	Others	Source
YouGov (VI)	06/06/2017	42%	38%	3%	9%				7%	[0620a]
YouGov (midpoint of seat range)	06/06/2017	301.5	270.0	0.5	13.0	41.5	1.5	1.0	22.0	[0620a]
Ashcroft (number of seats)	08/06/2017	363	217	0	5	44	1	1	20	[0620r]

- [0620a] <https://twitter.com/YouGov/status/872432058773233664>
- [0620r] <http://lordashcroftpolls.com/2017/06/estimated-conservative-majority-rises-final-ashcroft-model-update/#more-15319>

EXIT POLL

Table 24: Ipsos MORI / GfK Exit Poll for BBC News, ITV News and Sky News at 10pm 2017-06-08.

Pollster	Last day of fieldwork	Con	Lab	Lib Dem	SNP	Others	Speaker	Source
Ipsos MORI / GfK Exit Poll	08/06/2017	314	266	14	34	22	1	[0620z]

- [0620z] https://www.ipsos.com/ipsos-mori/en-uk/ipsos-mori-and-2017-general-election?language_content_entity=en-uk

APPENDIX 3: TWITTER FEEDS

- <https://twitter.com/benatipsosmori>
- <https://twitter.com/martinboon>
- <https://twitter.com/JoeTwyman>
- <https://twitter.com/anthonyjwells>
- <https://twitter.com/BobbyIpsosMORI>
- <https://twitter.com/MattLebo4>
- <https://twitter.com/ProfJaneGreen>
- <https://twitter.com/drjennings>
- <https://twitter.com/StephenDFisher>

APPENDIX 4: BPC 2015 MAE ON 4- AND 6-PARTY-FORCED BASIS

In 2015 The British Polling Council listed the final polls for the 2015 General Election and calculated the error on a six-party-forced basis (Con/Lab/LD/UKIP/Green/Other) against the 2015 GB results. Below we calculate the MAE and repeat the calculation on a four-party-forced basis (Con/Lab/LD/Other).

Table 25: 6PF MAE of 2015 opinion polls: see <http://www.britishpollingcouncil.org/general-election-7-may-2015/>

	CON	LAB	LD	UKIP	GREEN	OTHER
Average	33.6	33.6	9	13	4.8	6.1
Result	37.8	31.2	8.1	12.9	3.8	6.3
Error	-4.2	2.4	0.9	0.1	1	-0.2
Absolute error	4.2	2.4	0.9	0.1	1	0.2

MAE = 1.47

Table 26: 4PF MAE of the same 2015 opinion polls:

	CON	LAB	LD	OTHER
Average	33.6	33.6	9	23.8
Result	37.8	31.2	8.1	22.9
Error	-4.2	2.4	0.9	0.9
Absolute error	4.2	2.4	0.9	0.9

MAE = 2.1

APPENDIX 5: BPC 2017 MAE ON 4- AND 6-PARTY-FORCED BASIS

In 2017 The British Polling Council listed the final polls for the 2017 General Election and calculated the error on a six-party-forced basis (Con/Lab/LD/UKIP/Green/Other) against the 2015 GB results. Below we calculate the MAE and repeat the calculation on a four-party-forced basis (Con/Lab/LD/Other).

Table 27: 6PF MAE of 2017 opinion polls: see <http://www.britishpollingcouncil.org/general-election-8-june-2017/>

	CON	LAB	LD	UKIP	GREEN	OTHER
Average	43.7	35.8	7.9	4.4	2.1	5.9
Result	43.5	41	7.6	1.9	1.6	4.4

Error	0.2	-5.2	0.3	2.5	0.5	1.5
Absolute error	0.2	5.2	0.3	2.5	0.5	1.5

MAE =1.7

Table 28: 4PF MAE of the same 2017 opinion polls:

	CON	LAB	LD	OTHER
Average	43.7	35.8	7.9	12.6
Result	43.5	41	7.6	7.9
Error	0.2	-5.2	0.3	4.7
Absolute error	0.2	5.2	0.3	4.7

MAE = 2.6

APPENDIX 6: RESULTS

Table 29. Results used by this article as the results of the 2017 General Election

Def'n	Date	CON	LAB	LIB	UKIP	GREEN	OTHER	Total	Source	Speaker
GB_VI	2017-06-12	43.5%	41.0%	7.6%	1.9%	1.7%	4.3%	100.0%	[0622e]	?
GB_VI	2017-06-09	43.5%	41.0%	7.6%	1.9%	1.6%	4.4%	100.0%	[0622a]	Con
GB_VI	2017-06-09	43.5%	41.0%	7.6%	1.9%	1.6%	4.4%	100.0%	[0622d]	?
UK_VI	2017-06-19	42.3%	40.0%	7.4%	1.8%	1.6%	6.9%	100.0%	[0622b]	Oth
GB_seats	?	318	262	12	0	1	39	632	[0622c]	Con
GB_seats	2017-06-09	318	261	12	0	1	40	632	[0622a]	Con
GB_seats	2017-06-19	317	262	12	0	1	40	632	[0622b]	Oth
UK_seats	?	318	262	12	0	1	58	651	[0622c]	Con
UK_seats	2017-06-19	317	262	12	0	1	59	651	[0622b]	Oth

- [0622a] <https://yougov.co.uk/news/2017/06/09/how-yougovs-election-model-compares-final-result/>
- [0622b] <http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7979>
- [0622c] <http://www.bbc.co.uk/news/election/2017/results>
- [0622d] <http://www.britishpollingcouncil.org/general-election-8-june-2017/>
- [0622e] https://www.ipsos.com/ipsos-mori/en-uk/ipsos-mori-and-2017-general-election?language_content_entity=en-uk

For completion's sake we also note the postfacto probabilities. The Conservatives obtained most seats and most votes, but nobody had an overall majority, so:

Table 30. Postfacto probabilities (part 1) of the 2017 General Election

Def'n	CON	LAB	LIB	OTHER
P(most seats)	1	0	0	0
P(most votes)	1	0	0	0

For the overall majority odds it was computationally easier to use Con/Lab/Any other party/No overall majority as the four-party-forced, so:

Table 31. Postfacto probabilities (part 2) of the 2017 General Election

Def'n	CON	LAB	AOP	NOM
P(overall majority)	0	0	0	1