Election Forecasting

Stephen Fisher

University of Oxford

Outline

- Approaches to election forecasting
- UK General election 2015
- EU membership referendum 2016
- US General election 2016

Types of election forecasts

- Structural models
  - Election cycle, economics, leadership, …
- Vote intention poll based
- Synthetic models
- Betting and prediction markets
- Citizen forecasts
- Experts and "superforecasters"
- Social media and other big data
- Combining forecasts

Seats forecasting

- Uniform change
- Probabilistic uniform change
- Constituency level models
  - Reconciliation of votes at different levels
- Multilevel forecasts
  - EU and local as well as national
- Modelling seats outcomes directly
- Government formation

What do forecasts represent?

- Estimates as to what will happen
- Expressions of expectations based on past experience
  - From the range of data used (c.f. Black Swan)
- Ways of analysing effects of different factors
- Aggregations of information
  - Of the instincts of the forecasters
  - Of public polls and forecasts

2015 GENERAL ELECTION

Long range forecast approach

- Regress votes on polls x days before, and smooth the relationship over x
  - Picks up bias (to Labour) and regression to mean
  - Works better on change since previous election than levels
  - Later added adjustments for UKIP and SNP
- Probabilistic seat prediction
  - Later added adjustments for constituency polls and models of YouGov data at the constituency level

Tory underestimate within 0.4pts of 1992 miss

Eve-of-election polling minus popular vote share.
Data source: Mark Pack
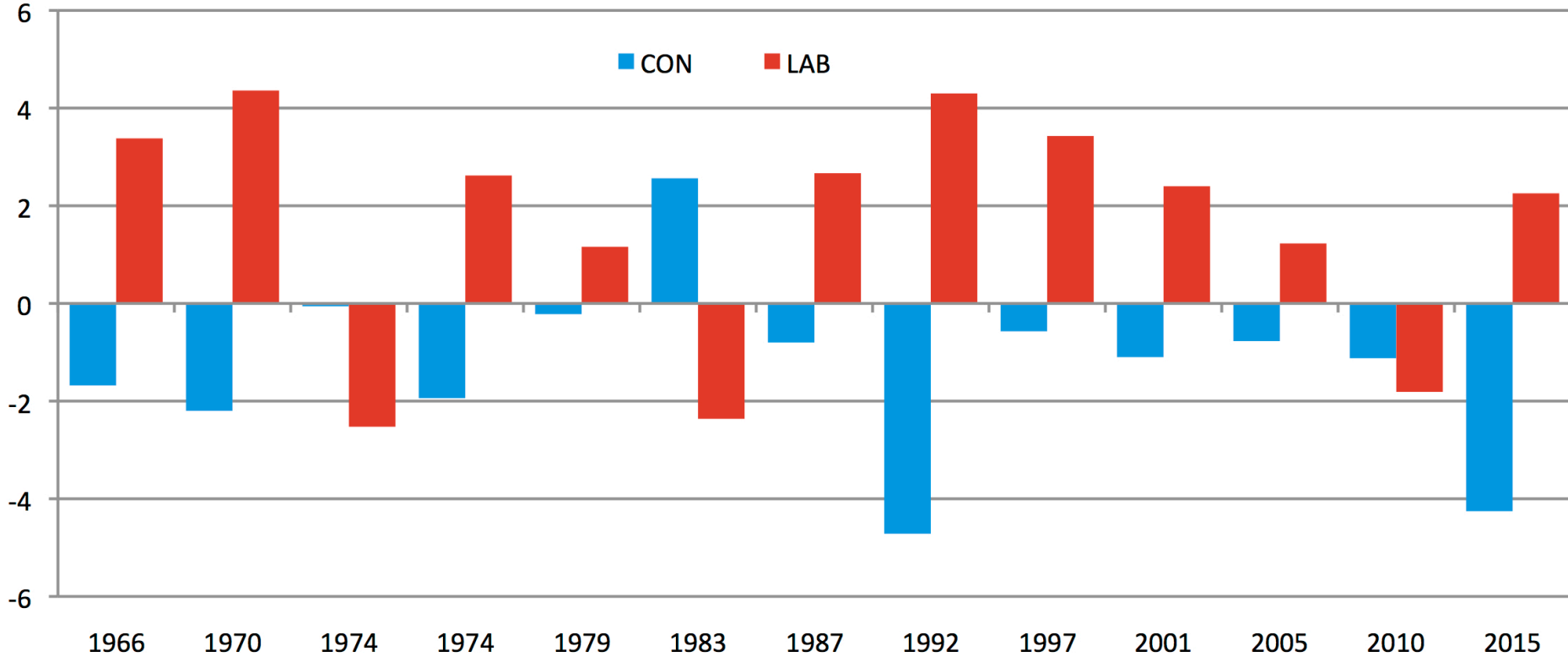
My first election forecast

Date of forecast: 25.10.2013
Days till the election: 559

Inputted current average poll shares
Con : 33
Lab : 38
LD : 11

Forecast Election Day Shares and 95% Prediction Intervals
Con : 40.2 plus or minus 11.8 i.e. between 28 and 52
Lab : 31.8 plus or minus 6.6 i.e. between 25 and 38
LD : 11.8 plus or minus 14.5 i.e. between 0 and 26

Forecast Election Day Seats
Con : 337
Lab : 265
LD : 21
Con majority of 24

Forecast Election Day Seat approx. 95% Prediction Intervals
Assuming LD share at 11.8 and allowing Con and Lab to vary as per interval above.
Con between 219 and 471
Lab between 140 and 376
LD between 14 and 30

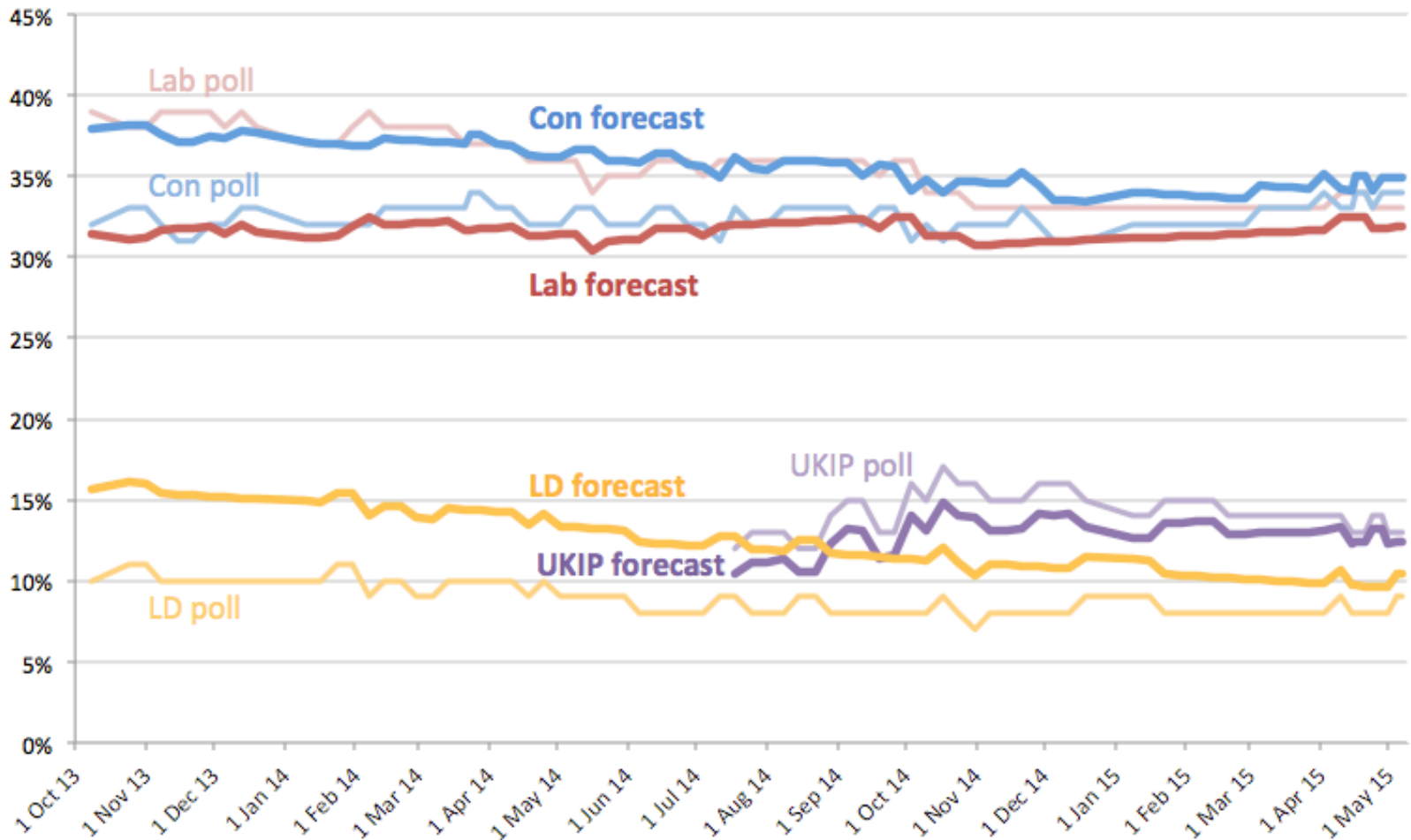**Probabilities of key outcomes**
Pr(Con majority) = 57%
Pr(Lab majority) = 15%
Pr(Hung parliament) = 28%
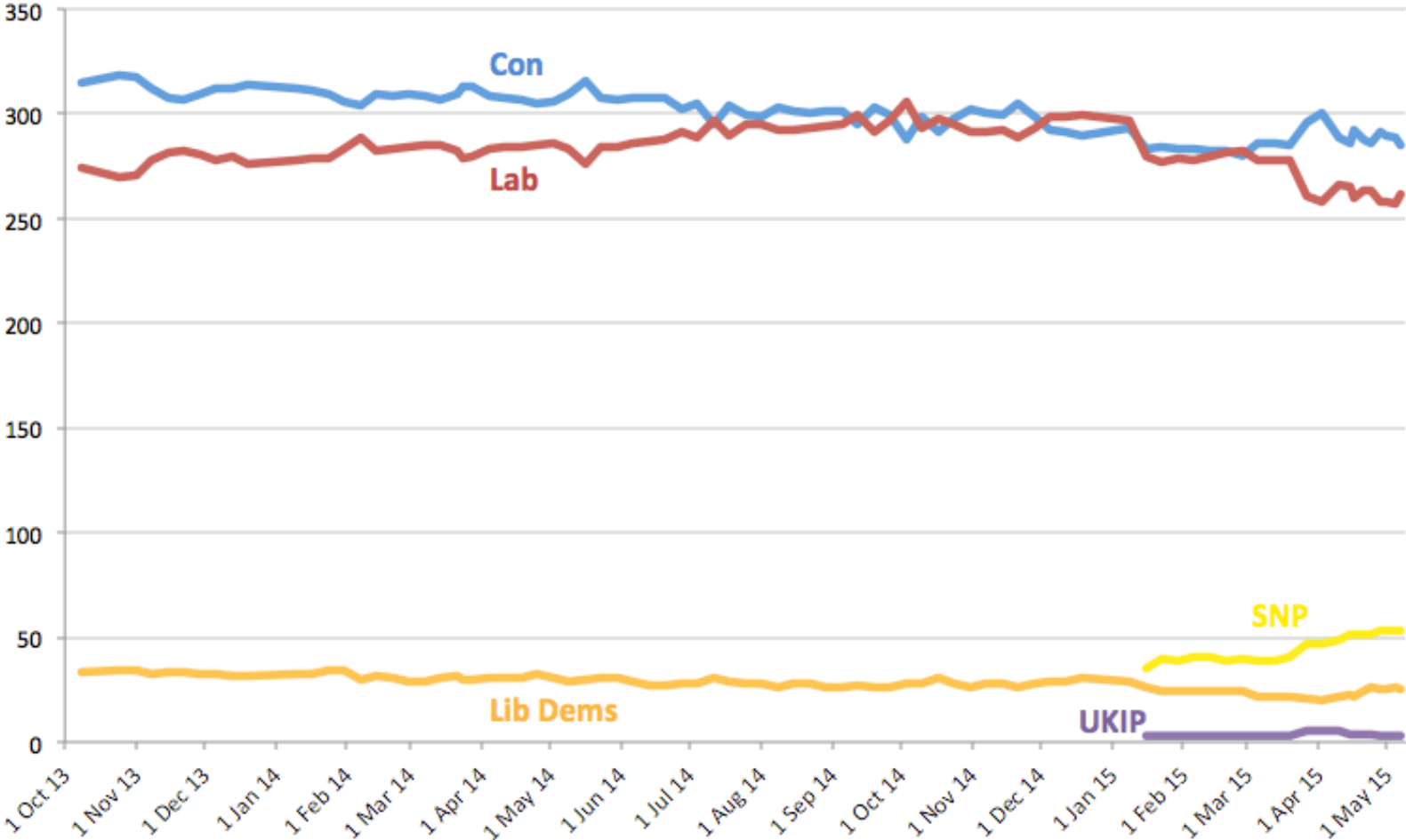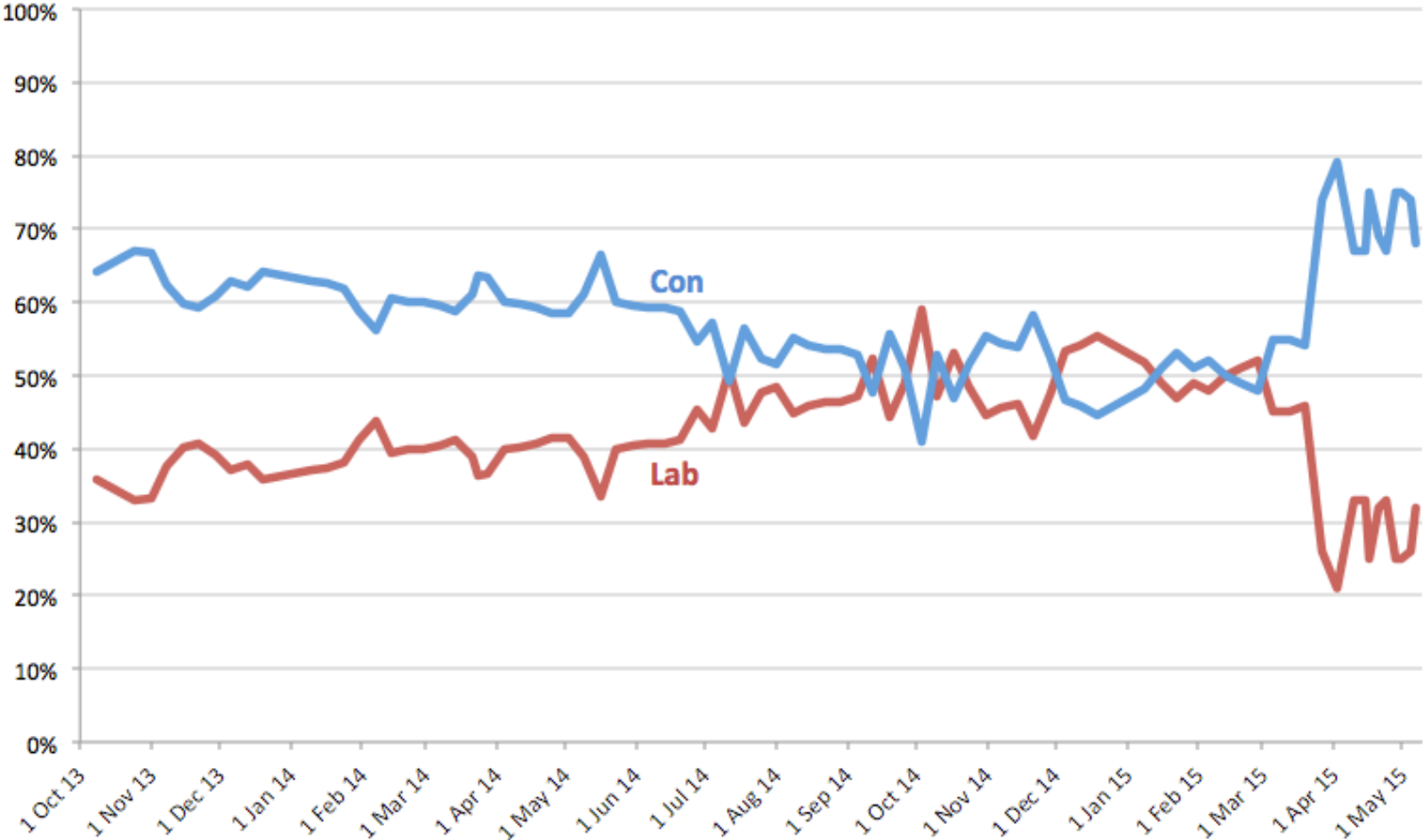Pr(Con largest party) = 88%
Pr(Lab largest party) = 12%

http://electionsetc.blogspot.co.uk/2013/10/a-long-range-forecast-for-2015-british.html
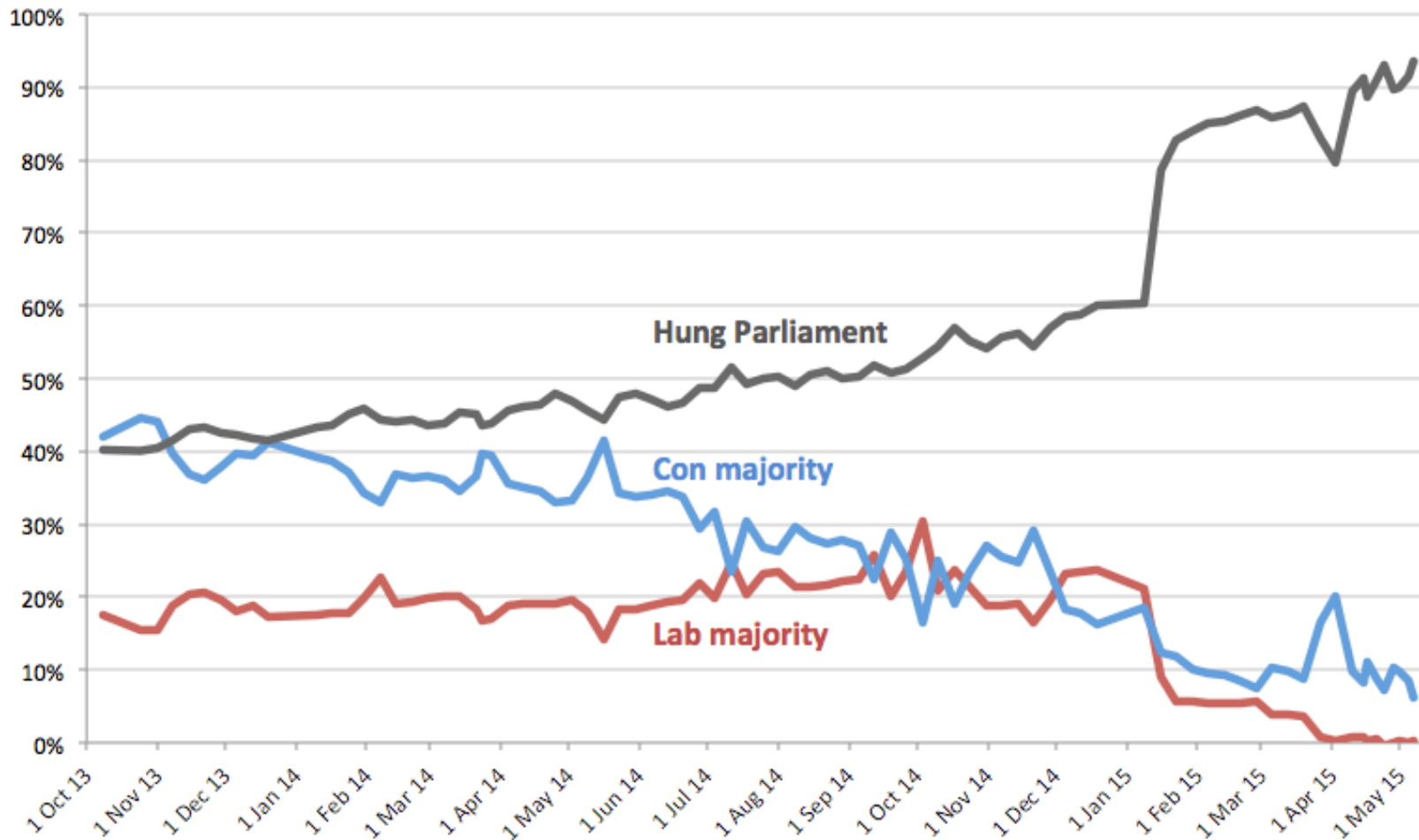
Polls and forecast shares

Seats forecasts

Largest party on seats probability

Key events probabilities

2015 UK General Election forecast, 7 May 2015

GB vote shares

Con 35%	Lab 32%	LD 10%	UKIP 12%	others 10%

UK seats

285	262	25	SNP 53	26
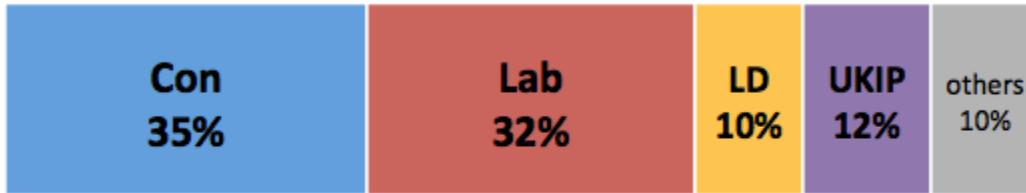
Conservatives 38 seats short of a majority

Probabilities

Lab maj, 0%
Lab+DUP, 1%

- Con maj 6%
- Con+DUP 5%
- Con+LD 16%
- Con+LD +DUP 14%
- Con largest, but Lab +SNP+DUP +others maj 5%
- Con largest, but Lab+SNP +other left maj 15%
- Con largest, but Lab+SNP maj, 6%
- Lab largest, Lab+SNP maj 8%
- Left+LD +DUP 9%
- Lab+LD +DUP 7%
- Lab +LD 7%

Lab PM 58%
Con PM 42%
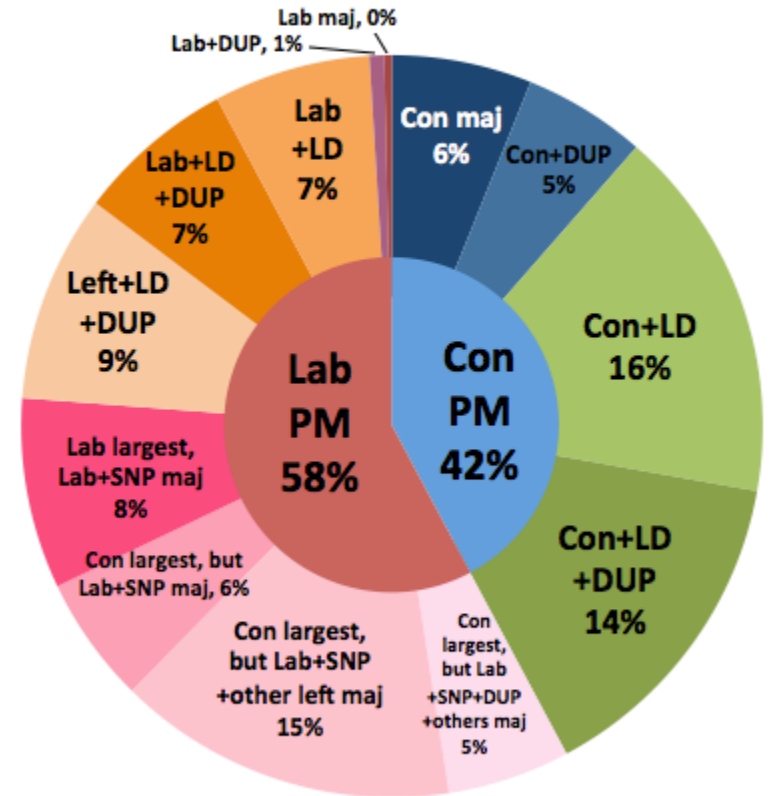
electionsetc.com   @electionsetc

Just a 3.3% chance of Con winning 331 or more

Seats outcome

	GB	E&W	E&W 2010 Seats:		
			Con	Lab	LD
CON	331	330	295	8	27
LAB	232	231	10	209	12
LD	8	7			7
UKIP	1	1	1		
SNP	56				
PC	3	3			
GRN	1	1			

Seats outcome with uniform change within England and Wales separately

	Actual	UNS	E&W 2010 UNS Seats:		
	E&W	E&W	Con	Lab	LD
CON	330	320	294		26
LAB	231	240	12	217	11
LD	7	9			9
UKIP	1				
PC	3	3			
GRN	1	1			
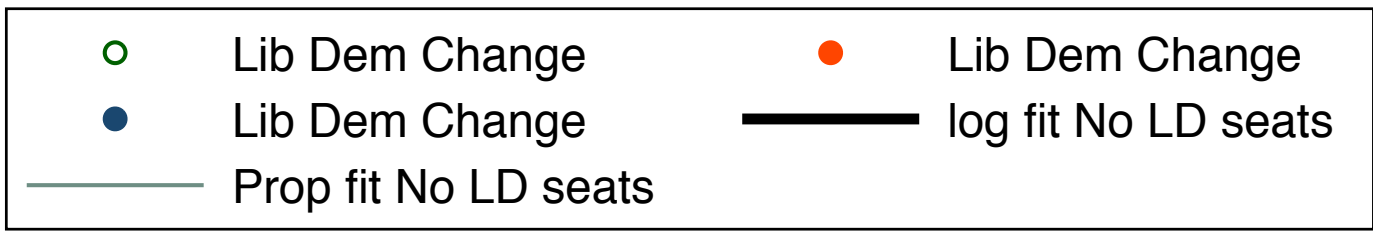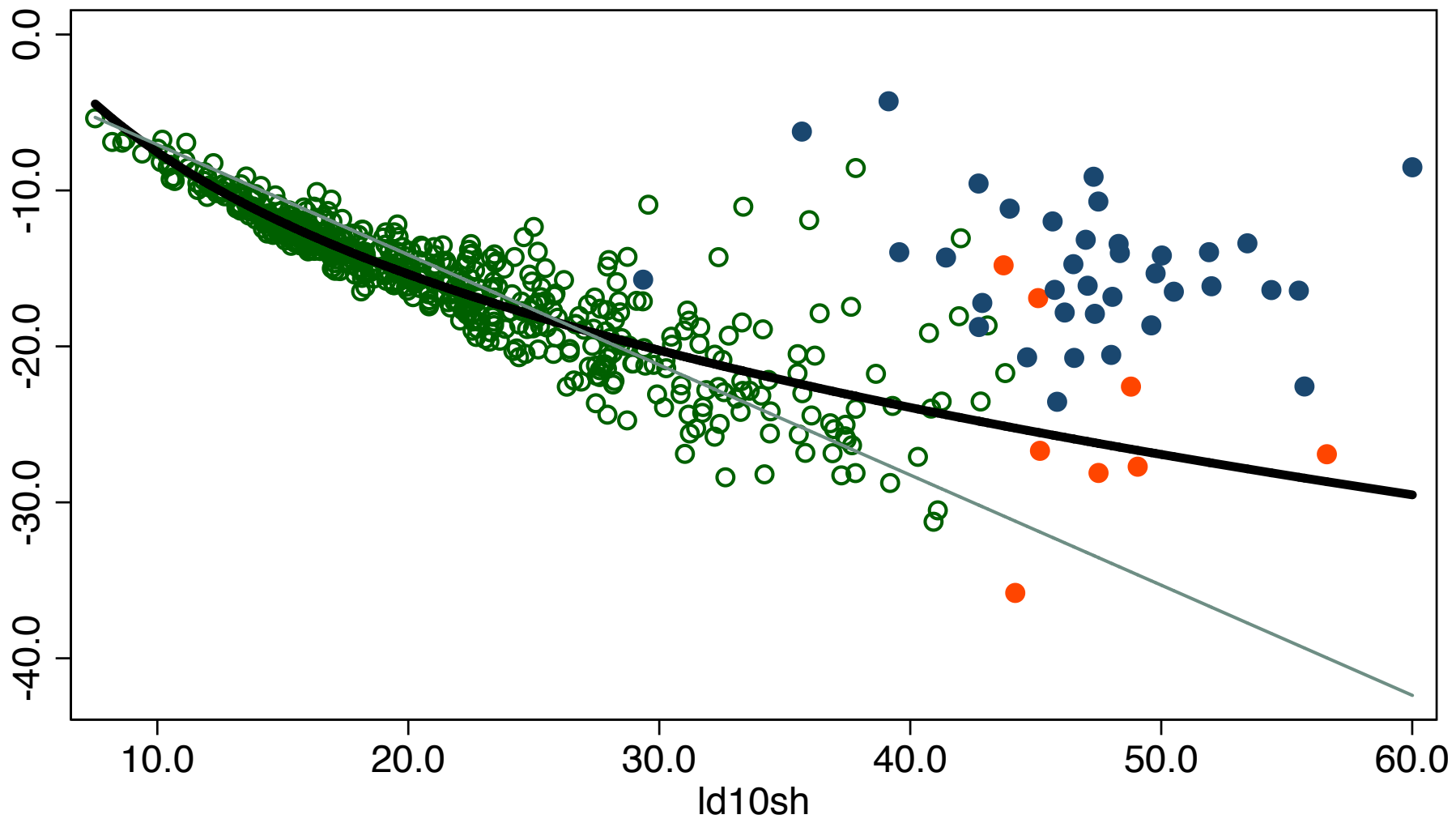
How come ESW uniform change with Con incumbency works so well?

- LD incumbency bonus (+11pts) relative to baseline pattern of collapse meant LD MP performance close to national average

- Conservative first-time incumbency effects important for explaining an extra 10 seats above nation-by-nation uniform change

- 8 Con gains from Lab offset 8 Lab gains from Con

- Uniform change in Scotland predicts SNP 55

Implications for accuracy of pre-election forecasts *this time*

- Getting national vote shares, especially Con-Lab lead in Eng & Wales right was most important

- Helpful to pick up Con first term incumbency effect

- Allow for Clacton as an obvious "special"

- Otherwise no need to worry about constituency variation to get accurate seat totals
  - By coincidence uniform change worked well for LD MPs

Academic pre-election forecasters (LSE, March 2015)

	Methodology
Burnap et al	Twitter
Fisher	Historical votes and polls; sub-GB polls; probabilistic prediction
Ford et al	Historical votes and polls; sub-GB polls; probabilistic prediction
Hanretty et al	Historical votes and polls; sub-GB polls; probabilistic prediction
Lebo et al	Electoral pendulum plus PM approval
Lewis-Beck et al	Gov approval, vote intention and economic growth
Mellon et al	BES voter transition model at the constituency level
Murr	Citizen constituency forecasts
Prosser	2013 and 2014 local election rounds
Rallings et al	Local by-elections
Stegmaier et al	CMS party support 3mth lag with Vote, PID, Econ Eval, PM approval, events
Whiteley et al	Modified cube-rule with lagged vote share but LD time-series model

Vote Share Predictions
(*Electoral Studies*, pre noon 7th May)

	Con	Lab	LD	SNP	PC	UKIP	Green	C-L
Actual	38	31	8	5	1	13	4	+6
Final Polls	34	34	9	5	1	13	5	+0
Forecast Ave	34	32	10	6	0.4	15	3	+2
Burnap et al	30	29	5	9	0	24	2	+1
Fisher	35	32	10	4	-	12	-	+3
Ford et al	35	33	9	-	-	-	-	+2
Hanretty et al	34	33	12	4	1	11	4	+2
Lebo et al	-	-	-	-	-	-	-	+3
Lewis-Beck et al	34	32	8	-	-	-	-	+3
Mellon et al	33	31	10	-	-	14	-	+2
Murr	-	-	-	-	-	-	-	-
Prosser	35	31	16	-	-	13	-	+4
Rallings et al	33	33	10	-	-	13	-	+0
Stegmaier et al	31	35	8	-	-	-	-	-4
Whiteley et al	-	-	-	-	-	-	-	-

Seat Predictions (#)

	Con	Lab	LD	SNP	PC	UKIP	Green
Actual	331	232	8	56	3	1	1
Opinion Polls	270	285	19	55	2	0	1
Forecast Ave	279	278	25	43	3	3	1
Burnap et al	285	306	21	9	3	5	1
Fisher	285	262	25	53	3	3	1
Ford et al	274	272	24	53	-	2	-
Hanretty et al	278	267	27	53	4	1	1
Lebo et al	287	263	-	41	-	-	-
Lewis-Beck et al	286	274	10	-	-	-	-
Mellon et al	274	278	23	52	2	0	-
Murr	292	262	28	40	4	5	2
Prosser	296	287	38	-	-	-	-
Rallings et al	276	280	22	48	3	2	1
Stegmaier et al	245	299	26	-	-	-	-
Whiteley et al	271	281	34	-	-	-	-

Summary evaluation of pre-election forecasts

- All were pretty poor
- Pre-election vote intention polls under-estimated Con by 4 points and over-estimated Lab by 2
  - Anything which improved on this tended to do better
  - including long-term electoral pendulum effects; PM approval; economic data and 2014 local elections (but not more recent local by elections)
- By coincidence uniform change worked well for LD MPs
  - But those who modelled constituency variation typically overestimated them

EU MEMBERSHIP REFERENDUM

# Theory for Dynamics

- Risk aversion kicks in.
  - The change option will decline in popularity
  - Final polls will tend to overestimate support for change
    - Don't Knows (DK) will split disproportionately for the status-quo between the final polls and the actual result
- These processes will be stronger for political reform and citizen initiated referendums.
- But reverse if the change option is portrayed as the real status quo

Forecasting model components

- Estimation of current public opinion
  - Poll aggregation
- Expectation of change in polls up to the final polls
- Estimation of "error" in the final polls
- Estimation of uncertainty in the above

Poll aggregation

- Simple poll averaging preferred to state-space models because of dramatic late changes in house effects: Scotland 2014, GE 2015
  - Somewhat similar story for EUref 2016
  - Use most recent poll from each company within last two weeks.
- But, Brexit polls showing big mode effects
  - Telephone polls had more Remain voters than online
    - Diff of 8.9 points on Remain share excluding DK in March
    - Declined to 2.7 point gap for final forecast
  - Online polls did best but it was not clear which was more likely to be right *ex ante*
  - Many more online polls than telephone polls
- So, estimate mode effect and benchmark to the mid-point of online and telephone
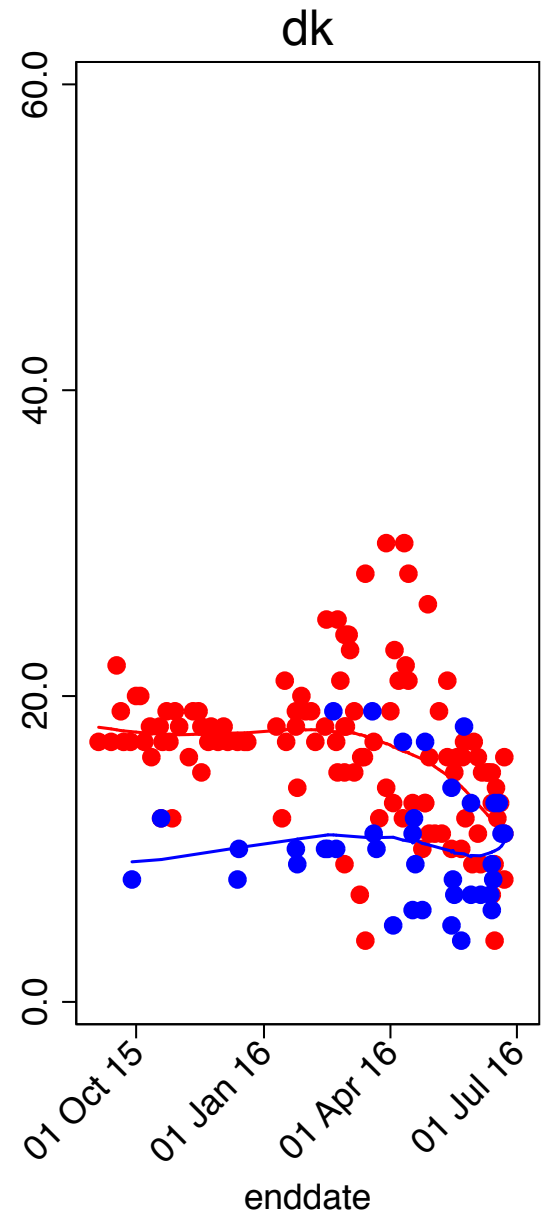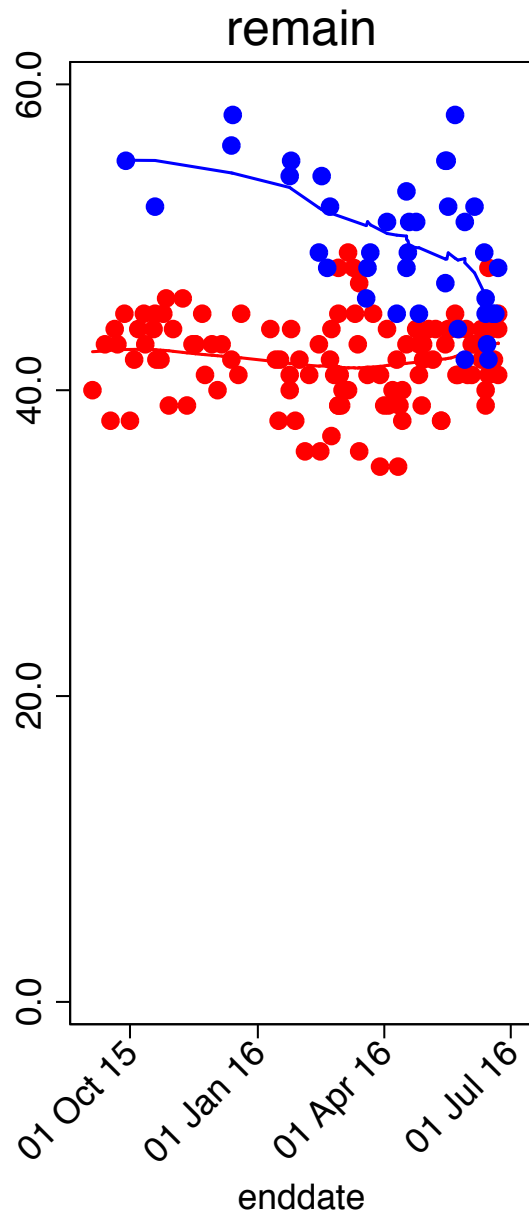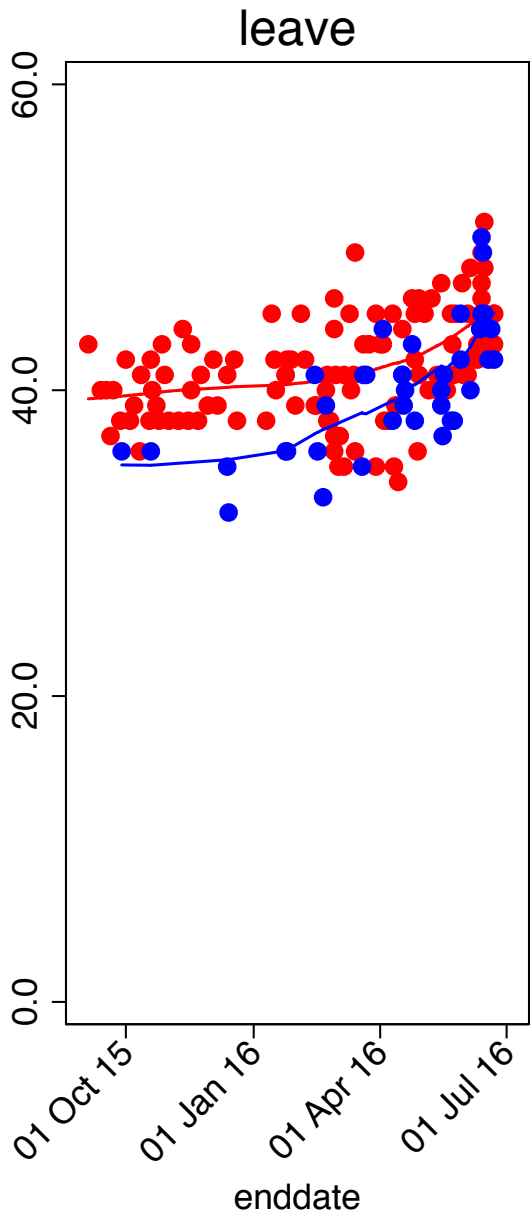
Online = red, Telephone = blue

Historical Data

- Use referendums in UK or on EU from a broader dataset
- Polls from national-level referendums in stable democracies since 1990
  - countries continuously rate 'Free' by Freedom House since 1991
  - Excluding Switzerland
  - excluding countries with population < 300,000
- Plus all referendums in UK at regional or national level
- 1686 polls from 283 referendums

Forecasting campaign trend

- Multilevel model for %Change in polls
- Use polls up to D+5 days out where D is number of days to the referendum
- Predictors: time to referendum interacted with political reform, citizen initiated and reversion point reversal dummies
- Predicted change is coefficient of time*D
- Just +1.15 point rise for Remain was expected over final 3 months, with s.e. 4.83 !

Forecasting the difference between the final polls and outcome

- Referendum polls internationally not very good:
  - Polls from last 10 days out on average by 6.7 points
  - This improves to just 4.5 if there were at least 3 such polls, and 2.8 for the 15 cases with 5 polls or more in final 10 days.
- In the 10 previous UK referendums with polls,
  - result outside range of final month polls in 5 cases
  - most extreme best in 7
- Estimated error in final polls for Remain (exc DK) to underestimate the eventual Remain vote by 1.51 points
 - Prediction interval was +/-10

Referendum	Poll average for change (%)	Share for change (%)	Difference	Passed	Change > 50%
AUT13 Conscription	45.2	40.3	-4.9		
AUS99 Republic	45.5	45.1	-0.3		
MLT15 Spring hunting	45.7	49.6	3.9		
AUS99 Constit. preamble	46.1	39.3	-6.8		
FRA05 EU Constitution	46.7	45.3	-1.3		
NOR94 EU accession	46.9	47.8	0.9		
DNK00 Euro	47.2	46.8	-0.4		
ITA16 Constitution	**47.3**	**40.9**	**-6.4**		
DNK15 EU opt out	48	46.9	-1.1		
GRC15 Bailout proposals	48	38.7	-9.3		
GBR16 Brexit	**48.4**	**51.9**	**3.5**	**Yes**	**Yes**
IRL08 EU Lisbon	49	46.6	-2.4		
SVN05 State broadcasting	49.8	50.7	0.9	Yes	Yes
SVN10 Croatian Border	51.2	51.5	0.4	Yes	Yes
SWE94 EU accession	51.9	52.7	0.9	Yes	Yes
MLT11 Divorce	52.1	53.2	1	Yes	Yes
FRA92 EU Maastricht	52.3	51	-1.2	Yes	Yes
DNK92 EU Maastricht	52.4	49.3	-3.1		
HUN04 Dual citizenship	52.8	51.6	-1.2		Yes
PRT98 Abortion	53	49.1	-3.9		
IRL02 Abortion	53	49.6	-3.5		
IRL95 Divorce	53.2	50.3	-2.9	Yes	Yes

Referendums in the UK: Final 30 days polls and outcomes
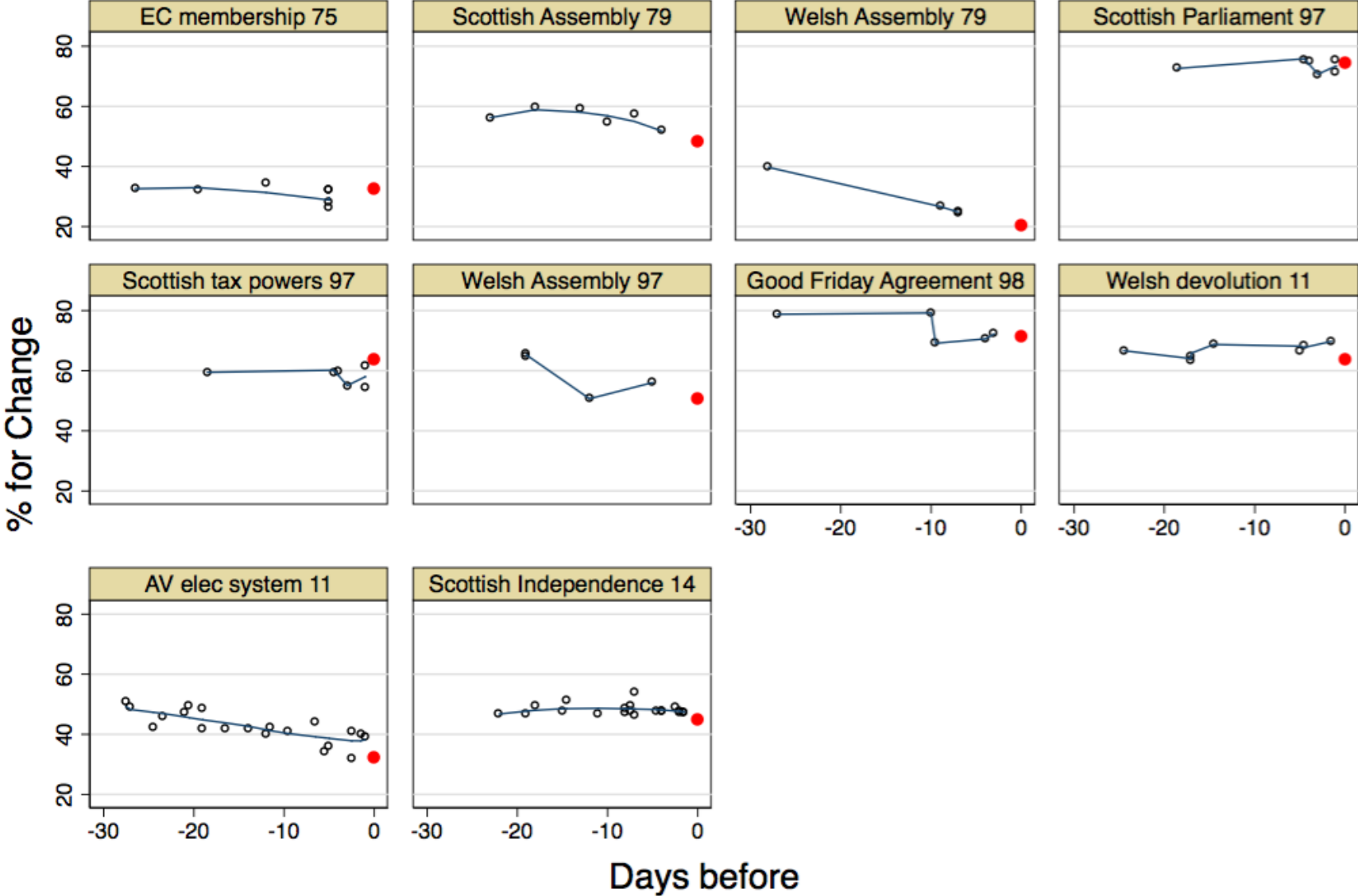
First and final forecasts

- First forecast, 14th March:
  - Polling average of 55% Remain
  - forecast of 58% Remain vote share
  - 95% prediction interval: 44% to 72% Remain
  - probability that Remain wins: 87%
- Final forecast, 23rd June:
  - Polling average of 51% Remain
  - forecast of 52% Remain vote share
  - 95% prediction interval: 42% to 62% Remain
 - probability that Remain wins: 64%
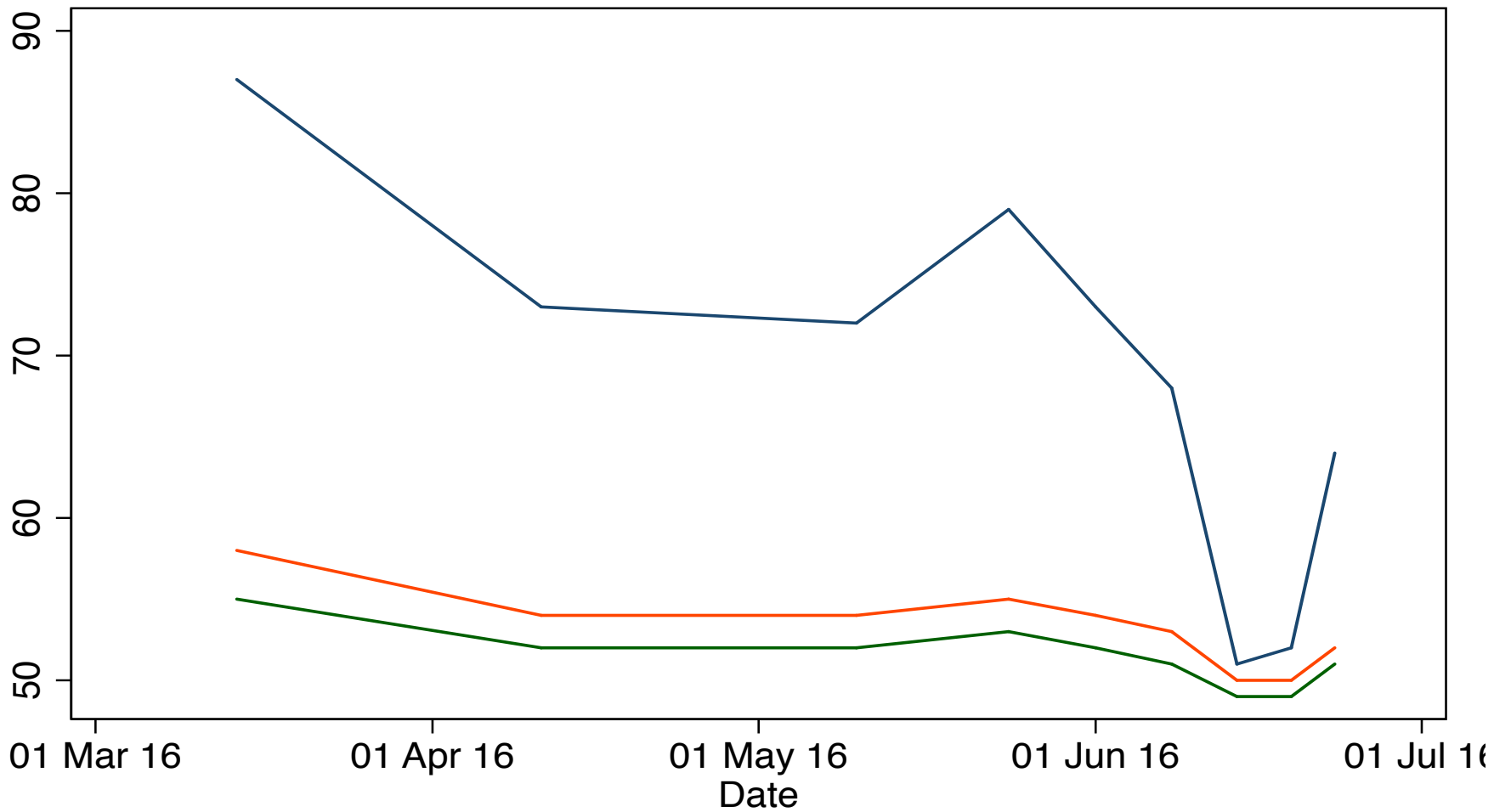
Combined Forecast Motivation

- Weak foundations, both empirically and conceptually, for forecasting EUref from other quite different referendums

- Not a good idea to rely on polls only, especially after GE 2015 polling miss.

- Many advocates of betting and prediction markets, superforecasters, citizen forecasters etc

- Combine different kinds of forecast

- Inspired by pollyvote.com

Final Forecast

	Remain % share	Leave % share	Probability Remain wins
Betting markets	53.5	46.5	76.7
Prediction markets			73.4
Citizen forecasts	52	48	55.2
Expert forecasts	55.1	44.9	62
Volunteer forecasts	54	46	74
Polls	50.6	49.4	55.6
Poll based models	52.5	47.5	68.5
Non-poll based models	55.6	44.4	
Combined forecast (mean)	**53.3**	**46.7**	**66.5**

Polls may have been out but they were best indicators.
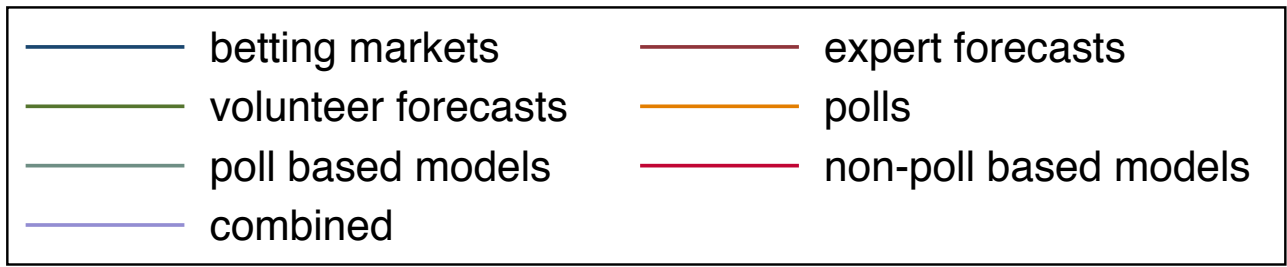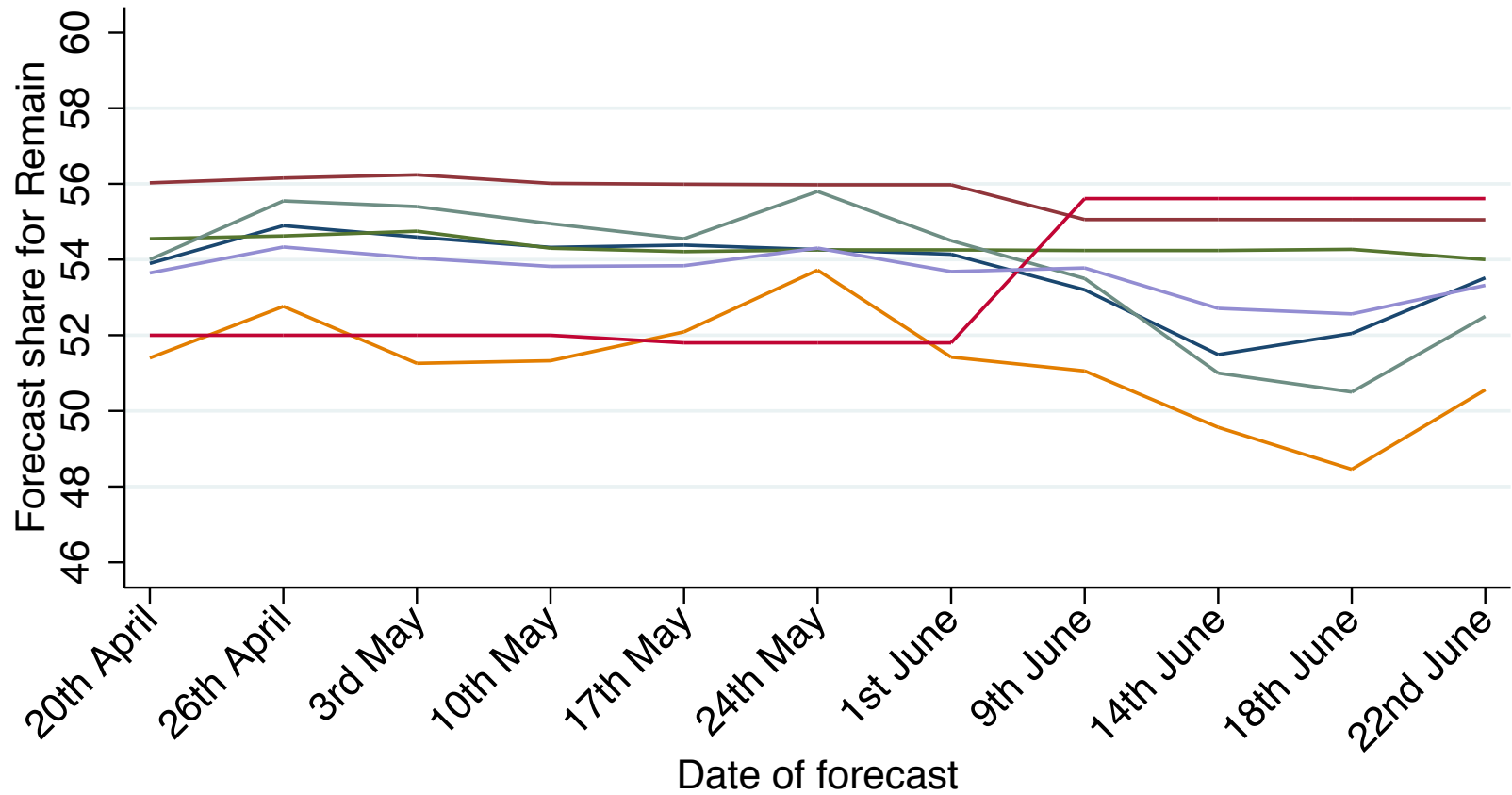
Source: Fisher-Shorrocks Combined Forecast method at ElectionsEtc.com
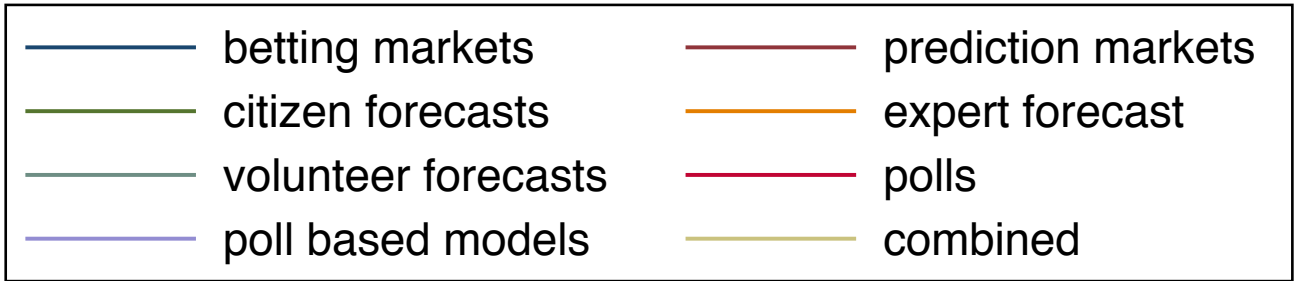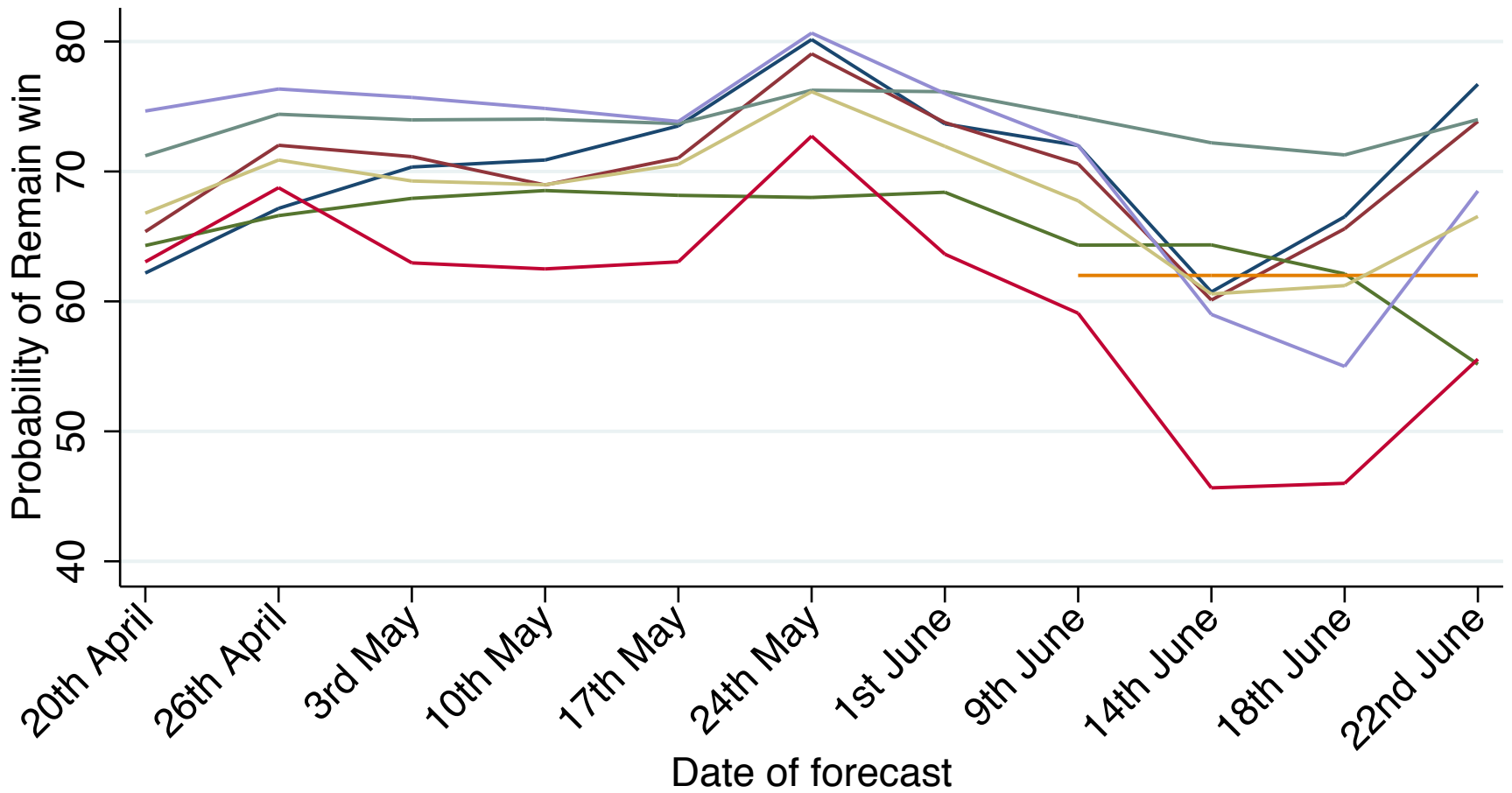
Source: Fisher-Shorrocks Combined Forecast method at ElectionsEtc.com

Brexit outcome was a surprise to many because…

- Polls initially and generally suggested Remain
- Support for change increased during the campaign when usually it drops
- People expected a late swing to Remain
- Only one previous referendum where vote for change against PM recommendation
- Only two previous referendums where change behind in final polls but went on to win.
- (+ various political and pejorative points)

Brexit should not have been a surprise because…

- Polls do no have a good track record in referendums, either in UK or elsewhere
- Internet polls were consistently showing it was close
- Clear potential for a nationalist bandwagon
- Final polls were close (51:49)
  - And this was partly because of pollster adjustments in favour of Remain
- (ElectionsEtc.com was saying a 36% chance of a Leave win)

So, public opinion change and poll performance for Brexit were not typical, but neither were they unusual

US 2016 presidential election

- Probabilities of Clinton win:
  - Poll based forecasters: Linzer 87%, HuffPost 98%, Wang 99%, Silver 67%.
  - Prediction markets: PredictIt 81%, Predictwise 89%
  - Citizen forecasts: NYT polls 5/3; Murr et al 6 pt lead
  - Pollyvote combined: 5 pt lead and 323 college
  - Abramowitz and Norpoth – Trump win

US Congressional elections 2016

- Few forecasts but House consistently expected to stay Republican
  - PredictIt 93%
- Senate forecasts all on the Democrat side with probabilities ranging from 53% to 91%
- Probability of GOP control of all three branches: 16% for PredictIt
  - But 7% of Dem control of all three

Conclusion

- Despite high profile failures, election forecasting is not in crisis
  - Brexit and Trump both had 30%+ probabilities
  - Alternatives to traditional methods still no better
- What's needed to do better
  - Polling accuracy to improve
  - Forecasters to understand and improve robustness
  - More attention to uncertainty by consumers
  - Appreciation of limitations
    - How much do we know and can we say?