Can we know voters' true intentions? Predicting stability and switching in vote choice before and after three General Elections

Stuart Perrett* & Jonathan Mellon

*British Election Study Research Associate
The University of Manchester

EPOP
12 September 2024

Predicting vote shares

1. How to link stated intentions and subsequent vote choices?
2. How well does the sample represent the electorate or subset of the electorate?

Predicting vote shares

1. **How to link stated intentions and subsequent vote choices?**
2. How well does the sample represent the electorate or subset of the electorate?

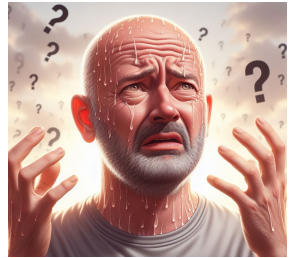
Our starting point

- Scope to improve our understanding of the link between intentions and choices
- By assessing what individual- or constituency-level information can be used to better predict vote shares from intentions
- Low hanging fruit – strong predictors may be easily or already collected by pollsters

Motivation – Two potential Conservative voters



- Intended to vote Conservative
- Lifelong Tory voter
- Strong Conservative identity
- Lives in Tunbridge Wells



- Intended to vote Conservative
- Voted Liberal Democrat in past
- Disliked Rishi Sunak
- Lives in Sheffield Hallam

Motivation

- The standard polling approach aggregates these two intentions as two Conservative votes – "nowcasters" (YouGov, Deltapoll, Lord Ashcroft, Redfield & Wilton, etc.)
- But we might think it is less likely the second respondent actually voted Conservative

→ Can additional information be used to weigh intentions and improve vote share predictions?

Data – BES Internet Panel

2015 Wave 4 (Pre-campaign) → Wave 6 (Post-election)
Wave 5 (Campaign) → Wave 6 (Post-election)

2017 Wave 11 (Pre-campaign) → Wave 13 (Post-election)
Wave 12 (Campaign) → Wave 13 (Post-election)

2019 Wave 17 (Pre-campaign) → Wave 19 (Post-election)
Wave 18 (Campaign) → Wave 19 (Post-election)

Methods – McFadden conditional multinomial logit models

We model the probability that voter i chooses party j among alternatives A as a function of alternative-specific covariates $\mathbf{X}_{ik}$

$$P(Y_i = j) = \frac{e^{\mathbf{X}_{ij}\mathbf{B}}}{\sum_{k \in A} e^{\mathbf{X}_{ik}\mathbf{B}}}$$

For example,

$$P(Y_i = Con) = \frac{e^{\mathbf{X}_{iCon}\mathbf{B}}}{e^{\mathbf{X}_{iCon}\mathbf{B}} + e^{\mathbf{X}_{iLab}\mathbf{B}} + e^{\mathbf{X}_{iLibDem}\mathbf{B}} + e^{\mathbf{X}_{iGreen}\mathbf{B}} + e^{\mathbf{X}_{iUKIP}\mathbf{B}} + e^{\mathbf{X}_{iOther}\mathbf{B}} + e^{\mathbf{X}_{iNone}\mathbf{B}}}$$

Probabilities for each party are then summed to produce our predicted vote share

Methods – Analytic approach

A voter's choice is modelled as a function of:

M1 their intention

M2 + an interaction with each party's distance from contention in their constituency

M3 + all other party-alternative predictors (party ID, past vote, like party leader, subjective certainty of intention, whether party is best on important issue, general likelihood of voting for party)

Methods – Distance from contention (DFC)

- Voters may make a tactical choice based on previous party support in their constituency
- We calculate each party's "distance from contention" – the difference between their vote share and the party with the highest other vote share at the previous election
- Positive values imply the party won, negative values that they were second or lower

Methods – Distance from contention (DFC)

Party	Sheffield Hallam 2019 Vote Share	Distance from Contention
Labour	34.6	1.2
Liberal Democrats	33.4	-1.2
Conservative	25.8	-8.8
Green	2.9	-31.7
Other	2.9	-31.7
UKIP	0.3	-34.3

Methods – Assessing vote share accuracy

We compare predicted vote shares with vote shares measured in each post-election wave using the Pedersen index:

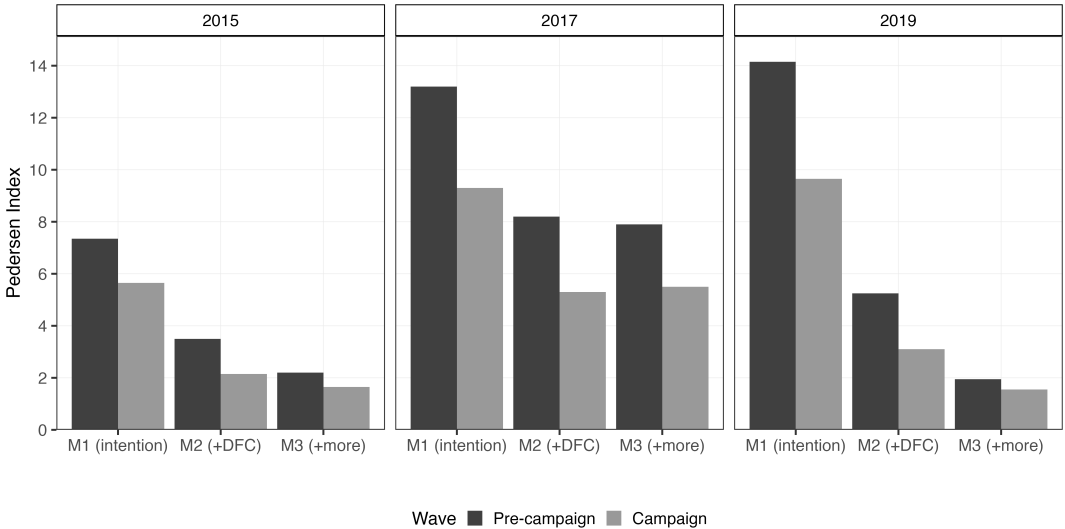
$$V = \frac{1}{2} \sum \mid p_{i,y} - p_{i,x} \mid$$

where $p_{i,t}$ is the percentage of party political party i in election wave y and $p_{i,x}$ their percentage in pre-campaign or campaign wave x

Results – 2015 vote shares

Party	Post-election share	Pre-campaign predictions M1 (intention)	M2 (+DFC)	M3 (+more)
Conservative	30.6	27.7	30.0	30.8
Labour	30.9	28.3	30.2	30.2
Liberal Democrats	7.9	8.2	8.0	7.9
SNP	8.9	7.3	6.9	7.7
Plaid Cymru	1.0	1.0	0.9	1.0
Green	2.7	3.8	3.0	3.0
UKIP	10.6	13.4	11.0	10.6
Other	0.9	4.2	2.6	1.9
None	6.3	6.1	7.4	7.0
Pedersen index		7.4	3.5	2.2

Results – Vote share predictions



Conclusions

- Information on a respondent's constituency can be used to substantially improve predictions of vote shares
- This information can be easily incorporated by pollsters

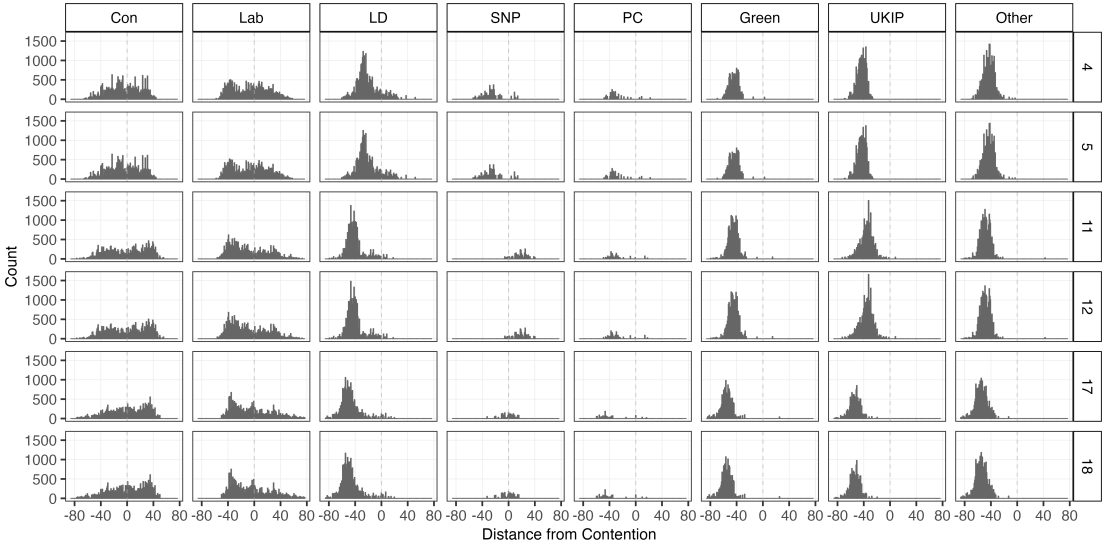
Moving forward

- Including respondents who answer "Don't know" when reporting their vote intention
- Separate models for England, Scotland, and Wales
- Using 2024 data
- Incorporating feedback on our exploratory models before out-of-sample prediction

Thank you!

stuart.perrett@manchester.ac.uk

DFC by wave and party



Observations and predictions

	id	bes_wave	alternative	choice	a_vote_intention	a_contention	a_party_id	a_past_vote	a_likeLeader	a_certainty	a_bestOnMII	a_ptv	prob_M1	prob_M2	prob_M3
1	1	4	None	FALSE	0	0.00	0	0	2	0	0	3	0.03	0.04	0.02
2	1	4	Con	FALSE	0	2.25	0	0	0	0	0	0	0.03	0.04	0.01
3	1	4	Lab	TRUE	1	−2.25	1	0	5	7	1	10	0.81	0.84	0.92
4	1	4	LD	FALSE	0	−19.91	0	0	0	0	0	0	0.03	0.03	0.00
5	1	4	Green	FALSE	0	−37.77	0	0	0	0	0	0	0.03	0.02	0.00
6	1	4	UKIP	FALSE	0	−36.23	0	1	6	0	0	5	0.03	0.02	0.04
7	1	4	Other	FALSE	0	−35.08	0	0	2	0	0	3	0.03	0.02	0.01