Election Forecasting

GRAD-E1234

Introduction and Overview

Simon Munzert

Spring Semester 2017
Humboldt-University of Berlin

Introduction

About me

- postdoctoral research fellow at Mannheim Centre for European Social Research
- postdoctoral research and teaching fellow at HU Berlin (starting in March)
- email: simon.munzert@gmail.com
- office hours: Fridays, after the course. Also available by appointment on Skype.

Session outline

- Organizational Matters
  - Goals
  - Requirements
  - Course Outline
- Introduction
  - Forecasting the 2013 German General Election
  - Potential future applications
- Forecasting corner
- Discussion

Organizational Matters

Goals

After attending this course, you will. . .

1. have acquired solid knowledge about all major flavors of election forecasting models,
2. be able to assess the value and trustworthiness of existing and new models based on a set of hard criteria,
3. be able to adapt existing models to new settings and design your own forecasting approaches, and
4. have the technical skills to implement these models using statistical software.

Prerequisites and Requirements

Prerequisites

1. decent understanding of statistics (Statistics I and II)
2. basic familiarity with statistical software (Stata or R)

Requirements

1. regular attendance
2. preparation of readings
3. mid-term exam
4. forecasting design paper
5. forecasting project

Requirements: regular attendence

- you cannot miss more than two sessions
- if you are unable to attend, please advice me by email with as much notice as possible
- participation in class will not be subject to grading

Requirements: preparation of texts

- usually two **required readings** per session
- invest roughly 1-2 hours for each of these papers
- be able to answer these questions after reading:
  - What is the idea behind the used forecasting model?
  - What are the core assumptions for this model to work?
  - How does the model relate to other forecasting approaches?
  - In which setting is the model evaluated/applied, and are you convinced by the application?
  - Which parts of the paper did you fail to understand?
  - What do you think of the paper?
- usually two to three **additional readings** per session
- invest no more than 15 minutes per additional reading, enabling you to give a summary of the core idea in a few sentences
- some additional readings in German → won't be covered in exam

General readings

- Tetlock, Philip E., and Dan Gardner, 2015: *Superforecasting. The Art and Science of Prediction.* New York: Crown.
- Silver, Nate, 2012. *The Signal and the Noise: Why So Many Predictions Fail–But Some Don't.* London: Penguin.
- Groß, Jochen. 2010. *Die Prognose von Wahlergebnissen. Ansätze und empirische Leistungsfähigkeit.* Wiesbaden: VS-Verlag.

Requirements: mid-term exam

- test on the contents of the course (only what was covered up until the mid-term exam week)
- 60 minutes; mix of open and closed questions
- will take place in the mid-term exam week

Grading

The exam makes for 30% of the final grade.

Requirements: forecast design paper

- develop your own forecasting approach
- discuss a scenario
- discuss benefits and drawbacks of chosen approach
- provide road map of the implementation of the model
- max. 3,000 words
- submit via Moodle
- *deadline*: April 12, 11.59p.m.

Grading

The design paper makes for 30% of the final grade.

Requirements: forecasting project

- implement forecast method of your choice with statistical software
- offer a transparent report of the procedure
- max. 6,000 words including (core) code
- collaboration possible (but you have to explicitly state who contributed to which part of the project)
- write in RMarkdown and submit via Moodle
- *deadline*: May 12, 11.59p.m.

Grading

The design paper makes for 40% of the final grade.

Course Outline

Please consult the syllabus now...

Session	Date	Topic
		Introduction
1	Feb 10	Introduction and Overview
2	Feb 17	Why Forecasting Elections is a Hard Case
3	Feb 24	NEW: Guest lecture by Jonathan Mellon
		Theories and Models
4	Mar 03	Fundamentals Models
5	Mar 10	Survey-Based Models
6	Mar 17	Prediction Markets
	Mar 24	**Exam Week**
7	Mar 31	Digital Trace Data Models
8	Apr 07	Combinations of Forecasts
9	Apr 14	Do Forecasts Matter? (date subject to change)
		Election Forecasting in Practice
10	Apr 21	A Primer to R
11	Apr 28	Data Collection and Model Setup
12	May 05	Out-of-Sample Testing and Publication of Forecasts
13	May ??	On demand: Presentation of Students' Model Ideas

Three more things

Three more things

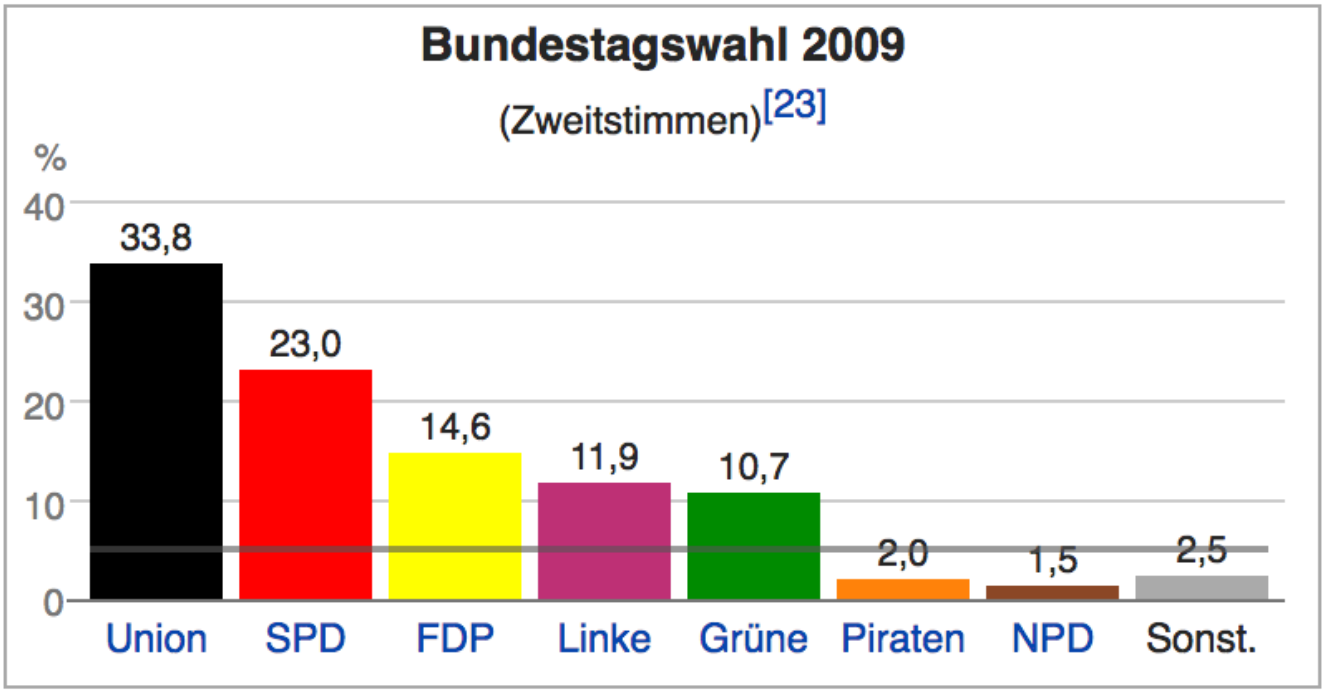
1. consume information on upcoming elections and take a critical view of horserace journalism
2. document your expectations about future events
3. task:
   - everyone, write down your current expectation of the distribution of vote shares across major parties for the upcoming Bundestag election.
   - defend your expectation in 2-3 sentences.

Introduction

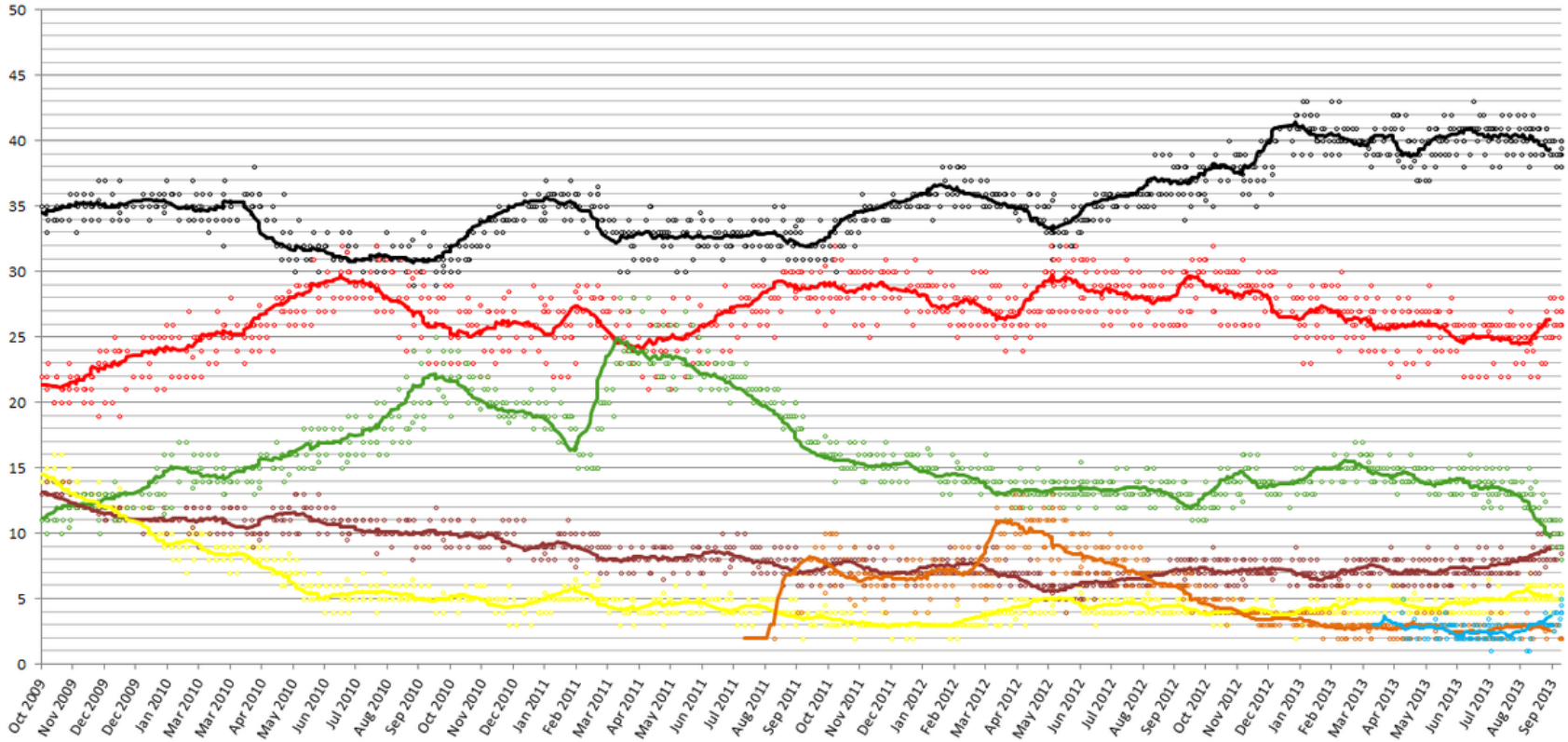
German General Election 2013: Scenario

German General Election 2013: Scenario

- chancellor-candidates: Merkel vs. Steinbrück
- junior coalition partner FDP extraordinarily unpopular
- Euro crisis / Financial crisis among more significant topics
- modifications to electoral law: overhang seats to be compensated through apportioned leveling seats



Polls before the election



By Humongous125 - Created in Excel using the polling data from Opinion polling for the German federal election, 2013., CC BY-SA 3.0, https://en.wikipedia.org/w/index.php?curid=48044259

Polls before the election

Institut	Datum	CDU/CSU	SPD	Grüne	FDP	Linke	Piraten	AfD	[Einklappen] Sonstige
Endergebnis[62]	*22.09.2013*	*41,5 %*	*25,7 %*	*8,4 %*	*4,8 %*	*8,6 %*	*2,2 %*	*4,7 %*	*4,1 %*
Emnid[63]	22.09.2013	39 %	26 %	9 %	6 %	9 %	—	4 %	7 %
Allensbach[64]	20.09.2013	39,5 %	27 %	9 %	5,5 %	9 %	2 %	4,5 %	3,5 %
Emnid[65]	20.09.2013	39 %	26 %	9 %	6 %	9 %	—	4 %	7 %
Forsa[59]	20.09.2013	40 %	26 %	10 %	5 %	9 %	2 %	4 %	4 %
Forschungsgruppe Wahlen[66]	19.09.2013	40 %	27 %	9 %	5,5 %	8,5 %	—	4 %	6 %
INSA[67]	19.09.2013	38 %	28 %	8 %	6 %	9 %	2 %	5 %	4 %

Polling forecast

Allensbach (Institut für Demoskopie) [Umfragen]

Wenn am nächsten Sonntag Bundestagswahl wäre …

1998–2002 | 2002–2005 | 2005–2009 | **2009–2013** | Aktuell

Gesamt | Ost | West

	CDU/CSU	SPD	GRÜNE	FDP	LINKE	PIRATEN	AfD	Sonstige	Befragte	Zeitraum
22.09.2013	41,5 %	25,7 %	8,4 %	4,8 %	8,6 %	2,2 %	4,7 %	4,0 %	Bundestagswahl	
20.09.2013	39,5 %	27,0 %	9,0 %	5,5 %	9,0 %	2,0 %	4,5 %	3,5 %	1.070	16.09.–19.09.
18.09.2013	39,0 %	26,0 %	11,0 %	6,0 %	9,0 %	2,0 %	3,5 %	3,5 %	1.269	03.09.–13.09.
04.09.2013	40,0 %	25,0 %	12,5 %	6,0 %	7,5 %	2,5 %	3,0 %	3,5 %	1.548	20.08.–31.08.
21.08.2013	39,0 %	24,5 %	13,5 %	7,0 %	7,0 %	2,0 %	3,0 %	4,0 %	1.553	03.08.–15.08.
12.07.2013	40,0 %	25,5 %	12,5 %	6,5 %	6,0 %	2,0 %	3,5 %	4,0 %	1.583	01.07.–11.07.
19.06.2013	38,0 %	26,0 %	14,0 %	6,0 %	7,0 %	2,0 %	3,0 %	4,0 %	1.490	01.06.–14.06.
22.05.2013	39,0 %	27,0 %	13,0 %	6,0 %	6,0 %	2,5 %	3,5 %	3,0 %	1.503	04.05.–17.05.
03.05.2013	38,5 %	26,0 %	15,0 %	5,5 %	6,0 %	3,0 %	3,5 %	2,5 %	?.???	15.04.–30.04.
17.04.2013	38,5 %	28,0 %	15,0 %	5,0 %	7,0 %	3,0 %	–	3,5 %	1.548	27.03.–12.04.
20.03.2013	39,5 %	26,5 %	15,5 %	6,0 %	6,0 %	2,5 %	–	4,0 %	1.5??	01.03.–14.03.
20.02.2013	39,0 %	27,0 %	15,0 %	6,0 %	6,5 %	3,0 %	–	3,5 %		01.02.–14.02.
23.01.2013	39,0 %	28,0 %	14,0 %	5,0 %	7,0 %	3,0 %	–	4,0 %	1.653	05.01.–18.01.
19.12.2012	37,5 %	30,0 %	14,0 %	4,5 %	6,5 %	3,5 %	–	4,0 %	1.611	01.12.–13.12.
21.11.2012	37,5 %	31,0 %	12,5 %	4,0 %	6,5 %	4,0 %	–	4,5 %	1.606	26.10.–08.11.

Polling forecast

Step 1: Assessing the Poll–Vote Link Using Historical Data

In step 1, we use historical data to predict election outcomes from polling results. To this end, we conceive of the vote share of party $j = 1, 2, \ldots, J$ as a linear function of a constant term α, the party's polling result in survey $i = 1, 2, \ldots, N$ conducted by institute $k = 1, 2, \ldots, K$, weighted by slope coefficient β, and a series of error terms that are specific to parties (ω) and an interaction of party and polling firm (ξ), plus an idiosyncratic residual (ψ), for all of which we impose the usual distributional assumptions:[11]

$$\text{vote}_j = \alpha + \beta \text{poll}_{ijk} + \omega_j + \xi_{jk} + \psi_{ijk}. \tag{1}$$

Step 3: Extrapolating to Current Elections

Equipped with these parameter estimates, we can now plug values of current polls into Equation (2) to arrive at poll-specific forecasts of party vote shares at the upcoming election that account for the types of biases described above:

$$\widehat{\text{vote}}_{ijk} = \hat{\alpha} + \hat{\beta}\,\text{poll}_{ijk} + \hat{\omega}_j + \hat{\xi}_{jk}. \tag{2}$$

Step 4: Combining Forecasts from Various Polls

For each party, we then combine the predicted vote shares from Equation (2), weighting by the reciprocal of the variance V of the individual quantities,

$$\widehat{\text{vote}}_{jm} = \frac{\sum_{i=1}^{N_m} \frac{1}{\hat{V}_{ijk}} \widehat{\text{vote}}_{ijk}}{\sum_{i=1}^{N_m} \frac{1}{\hat{V}_{ijk}}}, \tag{3}$$

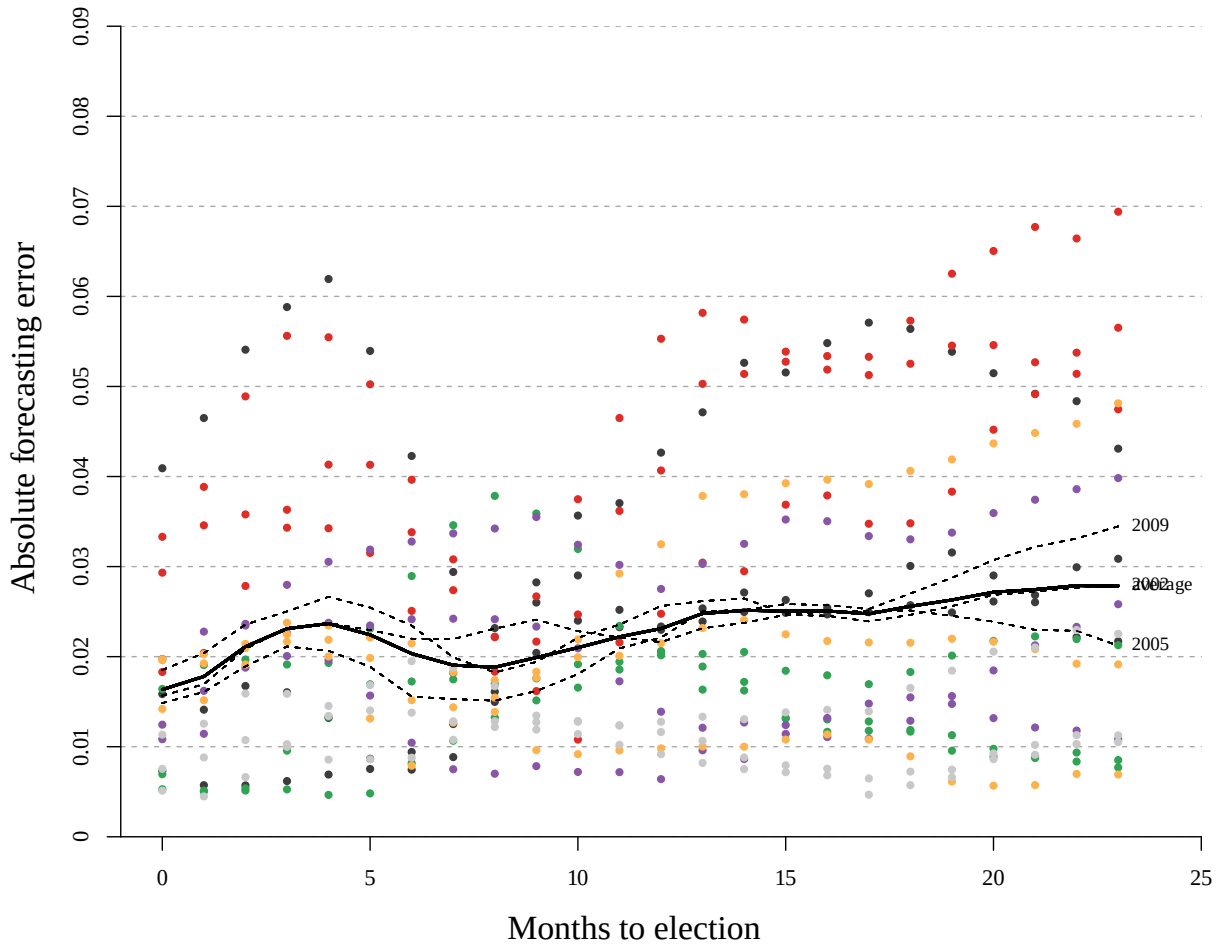
Polling forecast

TABLE 1
REML ESTIMATES OF THE MODEL OF PARTY VOTE SHARES IN PAST ELECTIONS, SEE EQUATION (1)

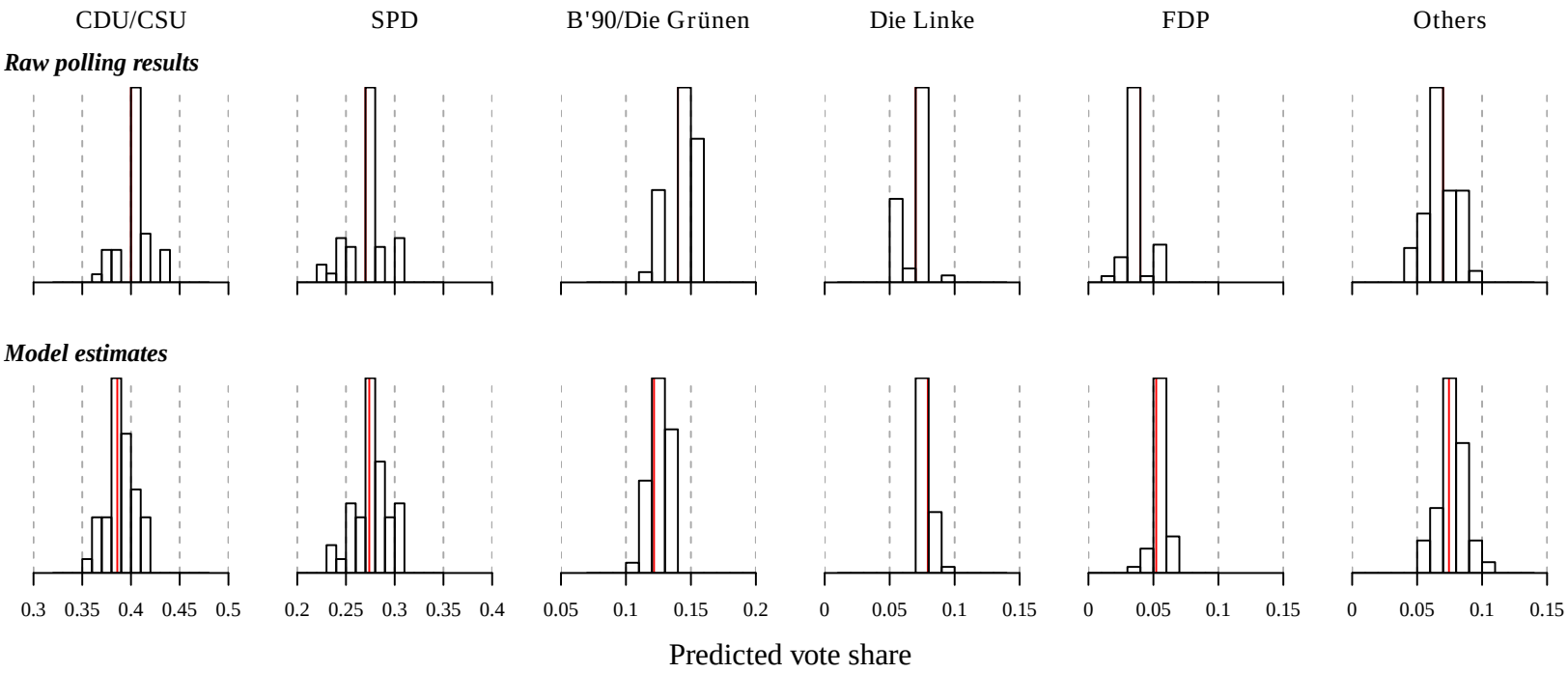
Coefficient	**Estimate (SE)**
Intercept α	0.0201 (0.0072)
Poll β	0.8742 (0.0216)
Party-level variance σ^2_ω	0.0002
Party–institute-level variance σ^2_ξ	0.0000
Residual variance σ^2_ψ	0.0004

Note: 123 polls conducted 8 to 10 months before the 1998 to 2009 elections are included.

Polling forecast



Polling forecast

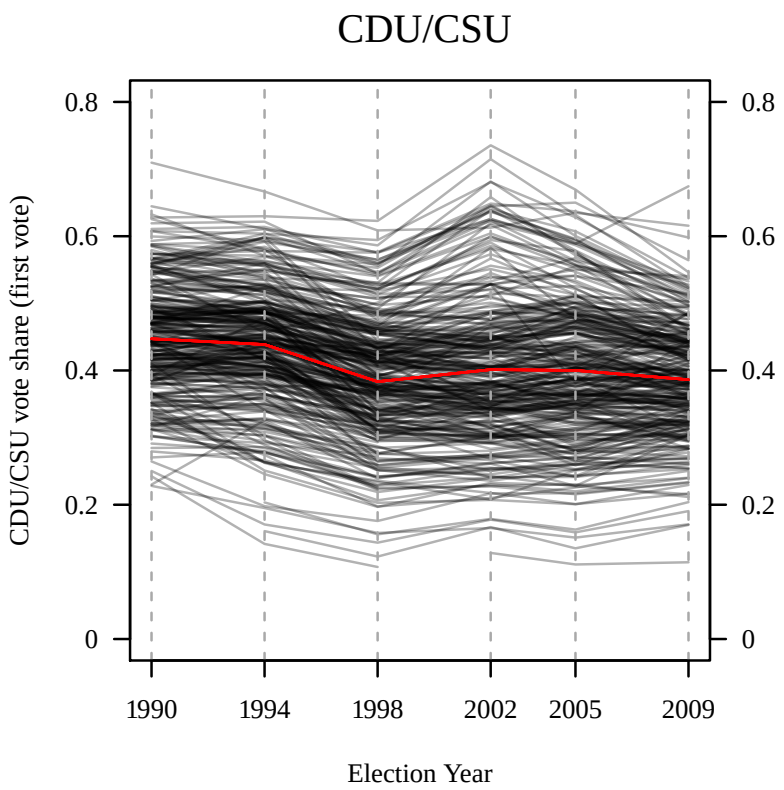


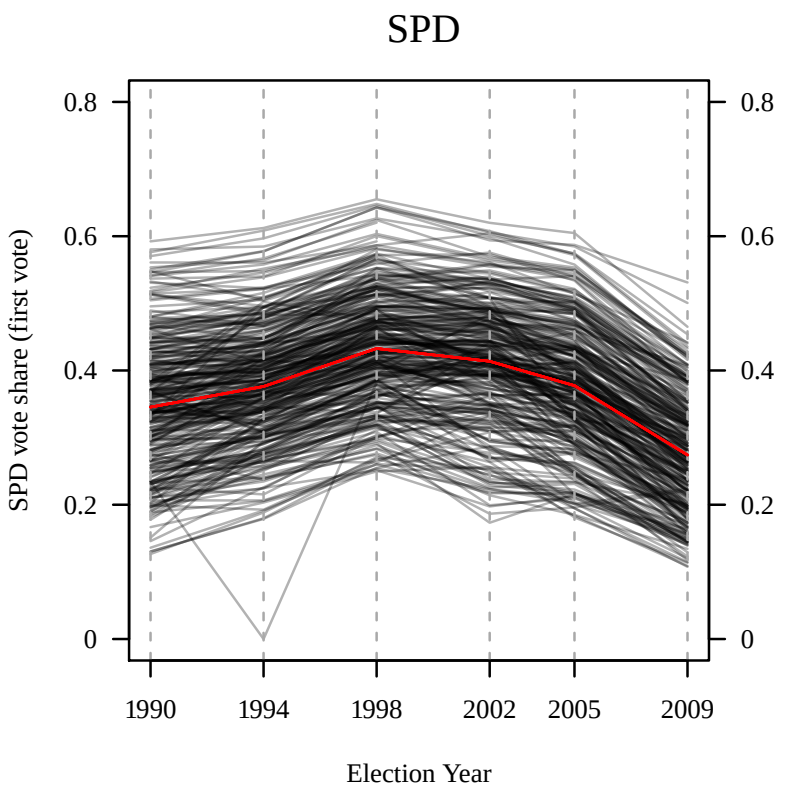
Forecasting the 2013 German Bundestag Election Using Many Polls and Historical Election Results

PETER SELB and SIMON MUNZERT

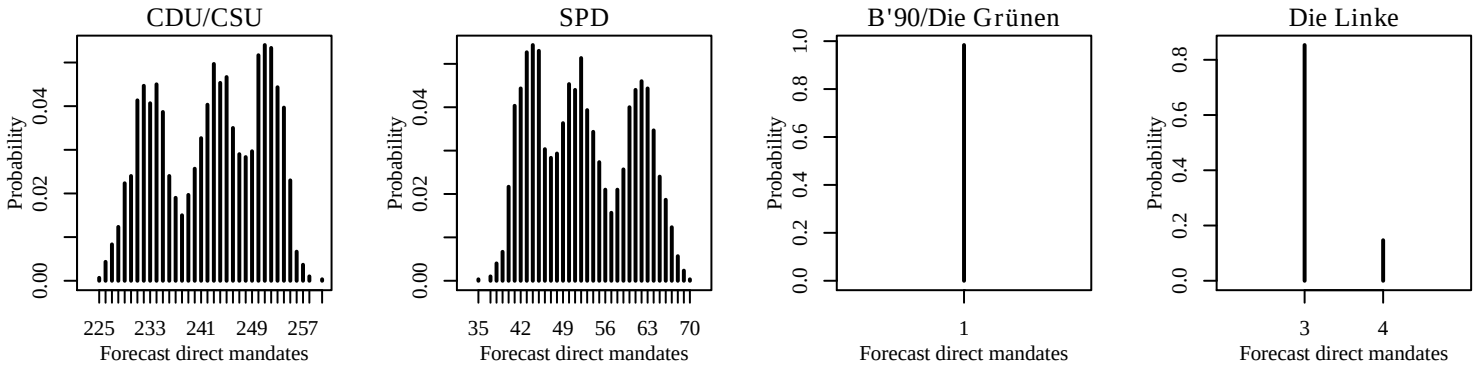
This article reports on an attempt to forecast the outcome of the 2013 election to the German Bundestag. In contrast to the predominant academic approach to forecast incumbent vote shares from measures of government popularity, economic conditions and other fundamental variables, we entirely relied on data from published trial heat polls. Opposite to common practice in the news media, we did not take isolated polls as election forecasts in their own right. Instead, we used historical data to assess empirically the relationship between polls and election outcomes, and combined extrapolations from current polls in a Bayesian manner. The forecast was published one month ahead of the election. The retrospective evaluation of our method was added after the election. While our method is parsimonious and provides a large lead time, the performance at the 2013 election was underwhelming. We offer additional suggestions how the approach can be improved in future scenarios.

Constituency forecast





Constituency forecast



Polling forecast

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting elections at the constituency level: A correction–combination procedure

Simon Munzert

University of Mannheim, Mannheim Centre for European Social Research (MZES), 68131 Mannheim, Germany

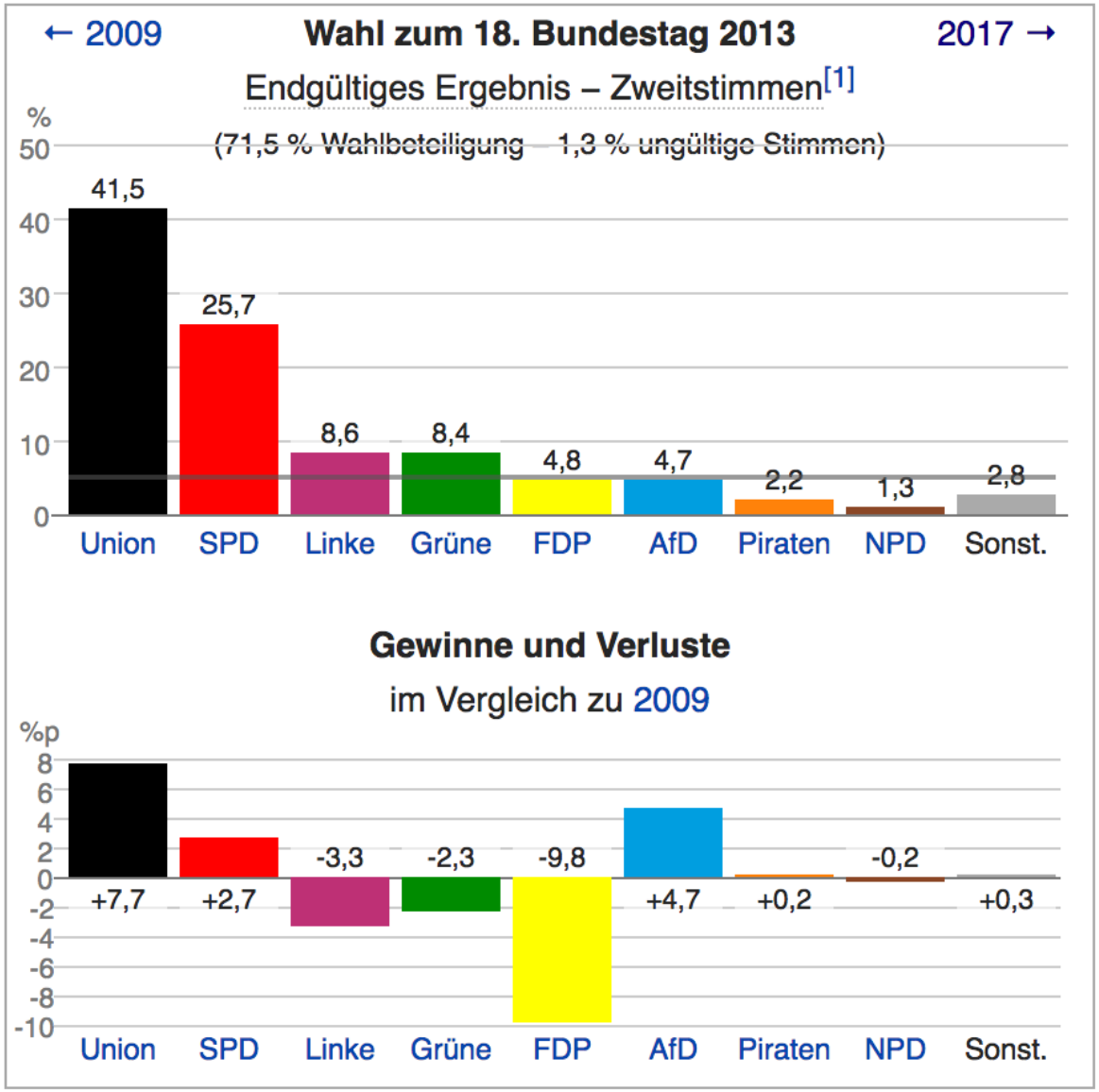
ARTICLE INFO

Keywords:
Election forecasting
Parliamentary elections
Constituency
Bias
Correction
Combination
Germany

ABSTRACT

Scholarly efforts to forecast parliamentary elections have targeted the national level predominantly, disregarding the outcomes of constituency races. In doing so, they have frequently failed to account for systematic bias in the seats–votes curve, and been unable to provide candidates and campaign strategists with constituency-level information. On the other hand, existing accounts of constituency-level election forecasting suffer from data sparsity, leading to a lack of precision. This paper proposes a correction–combination procedure that allows for the correction of individual constituency-level forecast models for election-invariant bias, then combines these models based on their past performances. I demonstrate the use of this procedure through out-of-sample forecasts of 299 district races at the 2013 German federal election.

Results


← 2009
Wahl zum 18. Bundestag 2013
2017 →
Endgültiges Ergebnis – Zweitstimmen[1]
(71,5 % Wahlbeteiligung – 1,3 % ungültige Stimmen)
%
50
40
30
20
10
0
41,5
25,7
8,6
8,4
4,8
4,7
2,2
1,3
2,8
Union
SPD
Linke
Grüne
FDP
AfD
Piraten
NPD
Sonst.
Gewinne und Verluste
im Vergleich zu 2009
%p
8
6
4
2
0
-2
-4
-6
-8
-10
+7,7
+2,7
-3,3
-2,3
-9,8
+4,7
+0,2
-0,2
+0,3
Union
SPD
Linke
Grüne
FDP
AfD
Piraten
NPD
Sonst.

Forecast aftermath

Tabelle 2: Das offizielle Endergebnis und die Abweichungen der Prognosen und letzten Umfragen vor der Wahl vom tatsächlichen Wahlergebnis

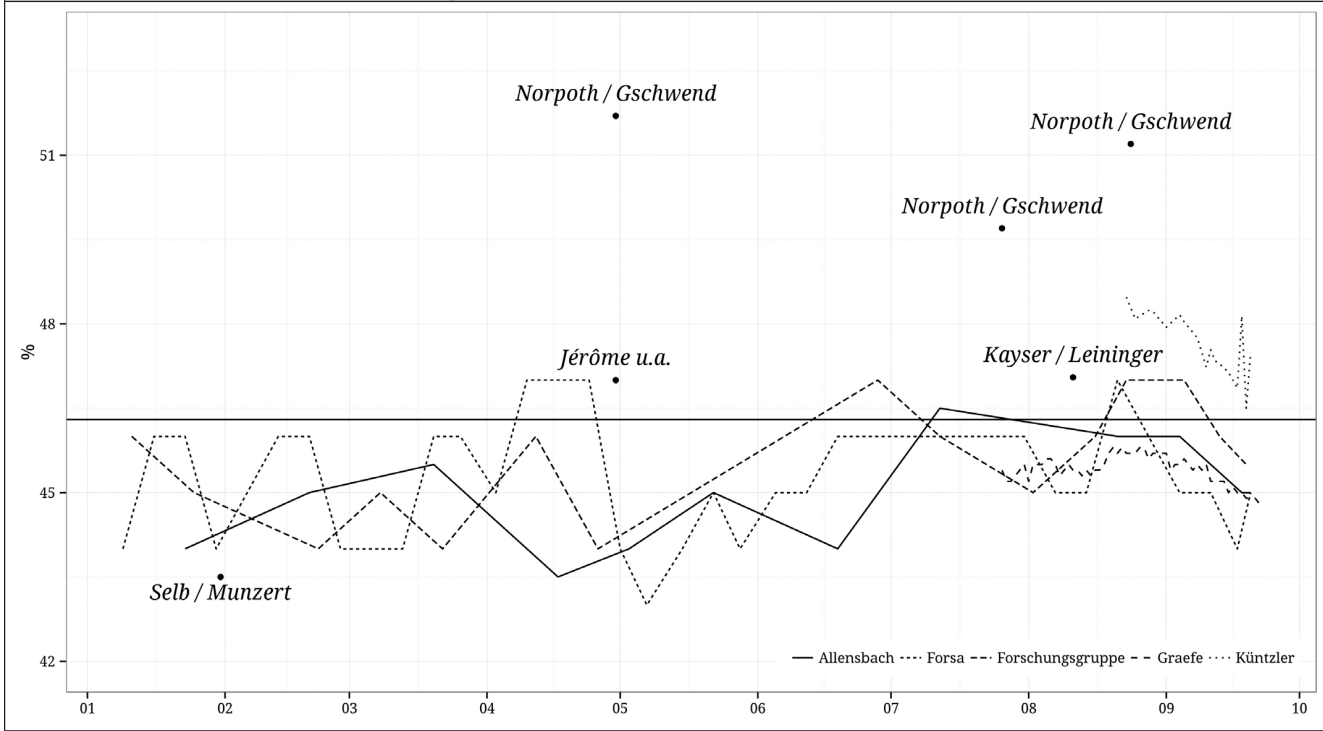
		Strukturelle Modelle				Synthetische Modelle		Umfragen		
	Endgültiges Ergebnis	*Jérôme* u.a.	*Norpoth / Gschwend*	*Kayser / Leininger*	*Selb / Munzert*	*Graefe*	*Küntzler*	Forschungsgruppe Wahlen (19.9.)	Forsa (20.9.)	Allensbach (20.9.)
Genauigkeit (MAE)	–	1,82	4,9	0,75	2,82	1,15	0,98	0,84	0,89	0,89
Vorlaufzeit (Monate)	–	5	5	2	8	2	1	3 Tage	3 Tage	3 Tage
Koalition	46,3	0,7	4,9	0,75	-2,8	1,5	0,98	-0,8	-1,3	-1,3
CDU/CSU	41,5	-0,5			3,4	-2,5		-1,5	-1,5	-2,0
FDP	4,8	1,2			0,6	1,0		-0,7	0,2	0,7
SPD	25,7	2,3			2,5	0,5		1,3	0,3	1,3
Grüne	8,4	1,6			5,1	1,9		0,6	1,6	0,6
Die Linke	8,6	0,4			-0,9	-0,2		0,1	0,4	0,4
Sonstige	10,9	-4,9			-4,4	-0,8		-0,9	-0,9	-0,9

Quelle: Eigene Zusammenstellung.

Leininger, Arndt. 2015. Wissenschaftliche Wahlprognosen - Alternative oder Ergänzung zu Umfragen? Zeitschrift für Parlamentsfragen 2015(4): 675–691.

Forecast aftermath

Abbildung 1: Zeitreihe der Wahlprognosen für die Summe der von Koalitionsparteien CDU/CSU und FDP zu erwartenden Stimmenanteile (Januar 2013 bis zur Wahl im September 2013)



Anmerkung: Modellbasierte Prognosen sind zum frühestmöglichen Zeitpunkt der Prognose eingetragen. Die Prognosen *Graefes* und *Küntzlers* sowie die durch Allensbach, Forsa und Forschungsgruppe Wahlen erhobenen Umfragewerte sind, um deren zeitliche Veränderungen darzustellen, als Linien abgetragen. Die horizontale Linie stellt das tatsächliche Wahlergebnis dar.
Quelle: Eigene Darstellung.

Leininger, Arndt. 2015. Wissenschaftliche Wahlprognosen - Alternative oder Ergänzung zu Umfragen? Zeitschrift für Parlamentsfragen 2015(4): 675–691.

Forecast aftermath: constituency forecast

		Seats forecast (direct mandate), by party					Σ dev
		CD(S)U	SPD	FDP	B'90	Linke	
Uncorrected	Uniform swing	224	70	0	1	4	24
	Polling	290	9	0	0	0	108
	Combined (eq. wgt.)	264	31	0	1	4	55
	Combined	229	65	0	1	4	14
Corrected	Uniform swing	223	71	0	1	4	26
	Polling	270	27	0	1	1	68
	Combined (eq. wgt.)	245	50	0	1	3	18
	Combined	242	53	0	1	3	12
Other	*election.de* (14.09.13)	224	69	0	1	5	24
	spiegel.de (21.09.13)	181	89	0	3	13	97
Actual result		236	58	0	1	4	

Forecast aftermath: constituency forecast

		MAE	% correct	% correct, marg. dist.
Uncorrected	Uniform swing	0.031	92.3	78.8
	Polling	0.046	81.6	55.6
	Combined (eq. wgt.)	0.031	89.6	68.7
	Combined	0.028	93.3	80.8
Corrected	Uniform swing	0.029	92.0	77.8
	Polling	0.030	88.3	66.7
	Combined (eq. wgt.)	0.026	92.0	75.8
	Combined	0.025	93.0	78.8

Potential future applications

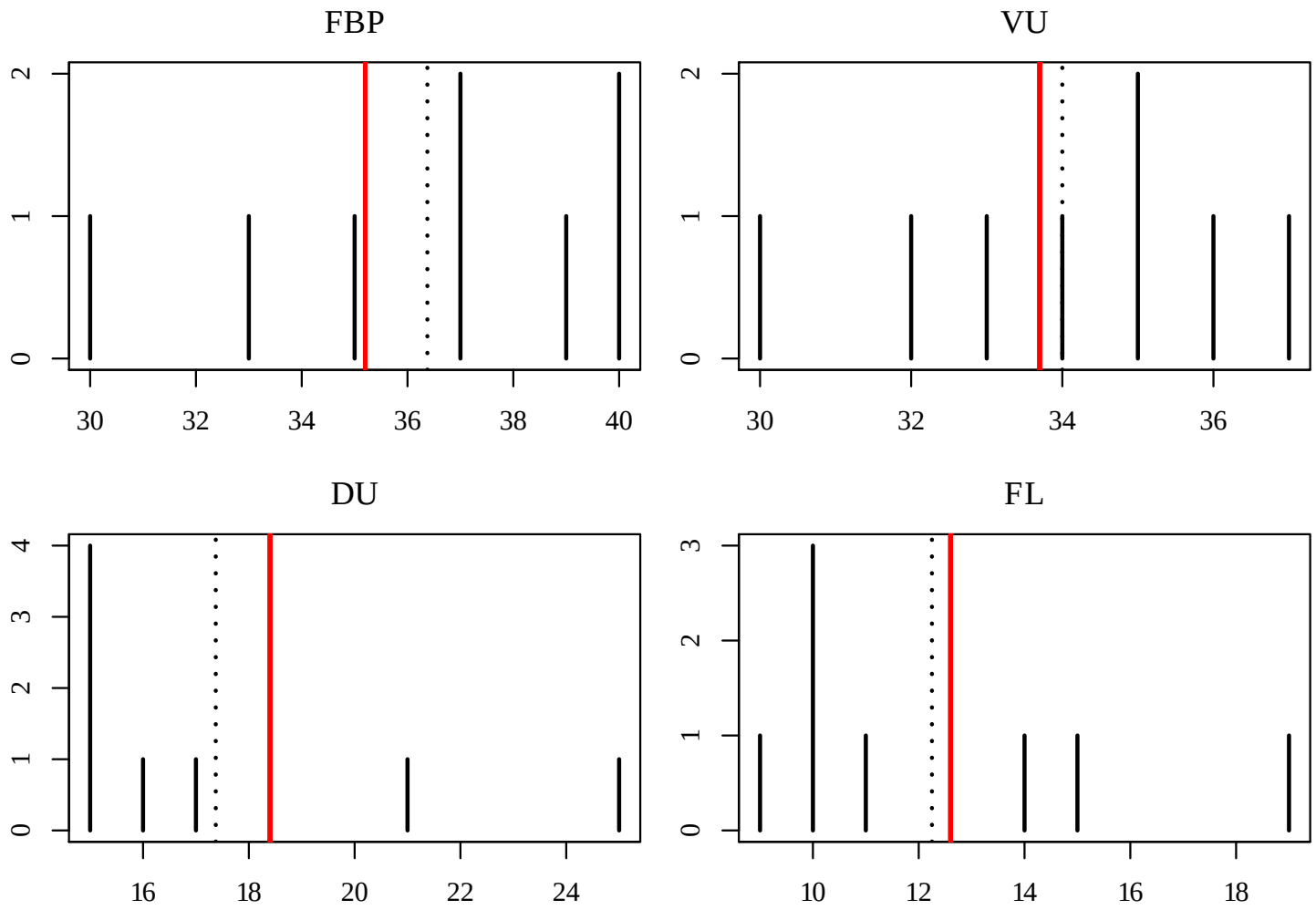
"If you have to forecast, forecast often." – Edgar R. Fiedler in *The Three Rs of Economic Forecasting—Irrational, Irrelevant and Irreverent*, June 1977.

Lists of potential applications

- https://en.wikipedia.org/wiki/National_electoral_calendar_2017
- https://en.wikipedia.org/wiki/Local_electoral_calendar_2017

- We will keep an eye on these cases throughout the semester
- Some of them might serve as good applications for your forecasting project

Election forecast of last week



FBP = "Fortschrittliche Bürgerpartei"
VU = "Vaterländische Union"
DU = "Die Unabhängigen"
FL = "Freie Liste"

Election forecast of last week

rank	respondent	mae	rmse	FBP	VU	DU	FL
1	Michael Chaitow	2.07	2.29	37	33	15	15
2	Rafael Schmuziger Goldzweig	2.18	2.22	37	32	21	10
3	Victoria Dykes	2.32	2.42	33	37	16	14
4	Moritz Hemmerlein	2.52	3.05	40	34	15	11
5	Nadina Iacob	2.52	2.79	39	35	17	9
6	Akira Sasaki	3.02	3.28	40	35	15	10
7	Jérémie Bonnemort	3.28	4.00	35	30	25	10
8	Hendrik Frank	4.32	4.61	30	36	15	19

FBP	VU	DU	FL
35.20	33.70	18.40	12.60

rank	note
1	There was very little information on party share in previous elections (except for the DU and FL). I used previous election results for these, and attempted to extrapolate based on news reporting for the rest.
2	Combining the general stability of the electoral results in the last decades and the recent political changes, I would expect the government coalition to remain majoritarely with the blacks and the reds (with a lower share of the votes if compared with the 2013 results). Following the trends of other European countries, I would expect a higher share of votes to the Independents, for its populist and anti-migration concerns. Considering the general cohesiveness of the electorate and the general support of the monarchy, I don«t expect to see a considerable change in the share of votes of the Free List.
3	I ended up just using a betting website's prognostics ("wahlfieber.at"), because I couldn't find anything else in the last two years! Literally nothing.

Election forecast of the week

- 12 February: Turkmenistani presidential election
- Make an educated guess about the outcome and document it here: https://goo.gl/forms/rHgeYUhvaWtcEyvm1 (also on Moodle)

Discussion

Question 1

Discuss the incentives and challenges involved in election forecasting for a scholar of political science!

Question 2

Imagine you had (practically) unlimited resources and your sole task was to forecast the outcome of the US presidential election one month ahead of time with as much precision as possible. Which strategy would you implement?

Question 3

There is a prominent view that election forecasts as well as public opinion polls published ahead of elections are normatively problematic because they can, by themselves, affect public opinion and ultimately change the outcome of an election. Discuss this view!

See you next week!