Election Forecasting using Postratification Methodology on Dalia Research's Europulse Data*

Moritz Hemmerlein & Alexander Sacharow

04 August, 2017

Summary or TL;DR

- Unweighted Dalia Europulse data faces similar problems as other online polling approaches; pre-stratification is not sufficient to create a reasonable election poll
- Applying post-stratification leads to more accurate election forcasting, in particular if reported past vote decisions are used for stratification
- Post-stratification purely on demographic information from the census data did not lead to significant improvements of accuracy, indicating that Dalia respondents are not representative for their demographic cluster
- The post-stratification with exit-poll data could be further fine-tuned with disaggregated exit poll data and more Europulse waves

*The R scripts and large parts of the raw data are available on our GitHub repository. For the remaining raw data please contact the authors, corresponding address: m.hemmerlein@mpp.hertie-school.org.

Introduction

Online polling for election forecasting is still in its infancy and major polling institutions remain hesitant to use it. However, potential sources of error of online polls are not fundamentally different from traditional polls: The polls need to ensure representativeness and take into account likely-voter behaviour. Representativeness is inparticular difficult as internet usuage is still not universal, inparticular among elderly, and most online approaches rely on self-selection of participants and, consequently, ruling out random sampling .

Dalia tries to counter these problems by pre-stratification of the sample along age groups and gender clusters. We used this prestratified data and tested different post-stratification approaches to improve the accuracy of the election poll further. The different post-stratifications are based on data from exit polls of Bundestagswahl 2013 and demographic data from the Census 2011. To measure the accuracy of the Dalia poll we used the moving average of the leading German eleciton polling institutes.

Methodology

To improve the accuarcy of the Dalia election poll, we used post-stratification methodology. Post-stratification is based on weights which are computed from the different shares of clusters of characteristics in population and in the sample. For example, if women above 60 from a rural settlement make of 4% of the population but only 2% in the sample, the responses from individuals of this cluster should be weighted with 2. Post-stratification can make polls highly representative, but often they have practical limitations: The number of cluster grows exponentially with the number of characteristics. This leads often to empty clusters or cluster with a low number of observations ($n < 30$). In the first case, ad-hoc adjustments have to make post-stratification possible (combing clusters). In the second case, the standard errors will be large, making the point estimator for the election result inaccurate. Nevertheless, we were able to employ post-stratification with three dimensional clusters (e.g. Age, Gender and self-reported vote at last election) and combing age clusters (in particular the higher age groups had to be merged to avoid empthy clusters)

We also considered two other approaches for the adjustment of the polls: raking and model-based post-stratification. Raking in comparison to post-stratification relies solely on marginal instead of combined probabilities, making it less accurate as post-stratificaiton but often more applicable as marginal probilities (one-dimensional distribution of a characteristic) are often available in public data. Model-based post-stratification on the other hand might also be a promising approch for small scale samples ($n = 1000$). Instead of using the actual responses in the clusteres, they are estimated with the help of an logistic regression. This reduces the problem of empty clusters, but the methodology is relatively complex and at the point of our project the related R packages were

Moritz Hemmerlein
Alexander Sacharow

not yet released.

Data

The data used for post-stratification and weighting comes from the German official election statistics (Der Bundeswahlleiter 2017). It contains combined information of turnout and voting decision by gender and age. Age is clustered in groups of 18-25, 26-35, 36-45, 46-60 and 60 plus. This data is the closest estimate of the actual voting population available and is a common measure to account for likely voter bias.

The big shortcomming of the data is that it is available in no more than two-dimensions, which means that the data is never clustered by more than two demographic variables. Potentially, other and more detailled exit poll data could be used to improve the post-stratification and further enhance accuracy and representativeness for actual voter.

As benchmark data we used the rolling average as computed by the *Süddeutsche Zeitung*. The raw data and the computing method can be found on GitHub.

Results

Following, we display our results for the December and the March data respectively. The first figure shows the election forecast yielded by the unweighted raw data and benchmarks it with the weighted results. As the graph shows, the forecasts differ significantly. The raw data has a strong bias towards the smaller parties and non-voters. Weighting this data yields more reasonable results with improvements for the CDU/CSU and the SPD while the smaller parties loose. Moreover, the expectation of non-voters gets more accurate.

The table below summarises the numerical results and displays the the root-mean-squared error of the estimates with the benchmark data (polling institutes rolling average). The weighting by gender and age using exit poll data yields a RMSE of 2.17 and 2.05 for December and March respectively.

Graphically the results are shown in the graph below. As one can see, the weighted estimates (points in the graph) move smoothly with the moving average forecast that is used as a benchmark.

Shortcomings and Potential for Improvement

- more detailed stratification
- other demographic information
- improving the questionaire to identify likely voters

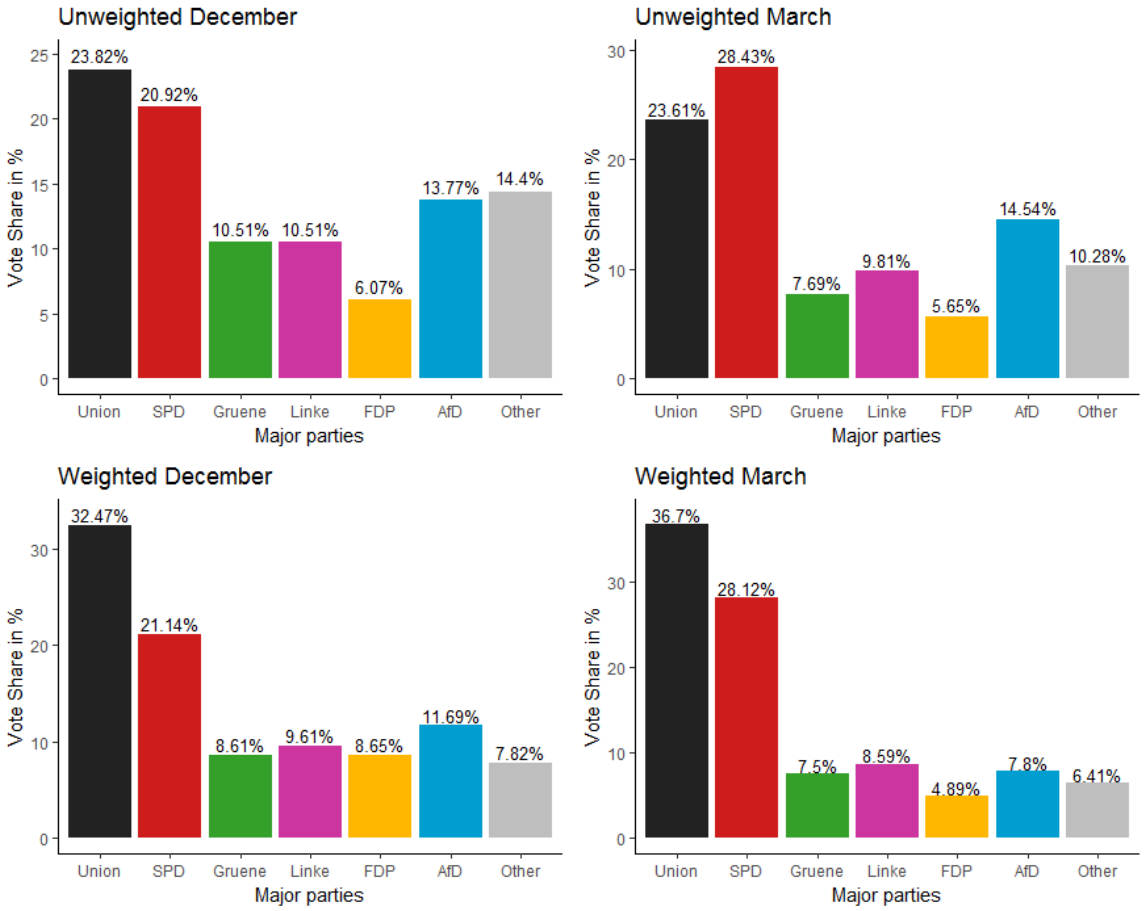

Figure 1: Weighted and unweighted polls

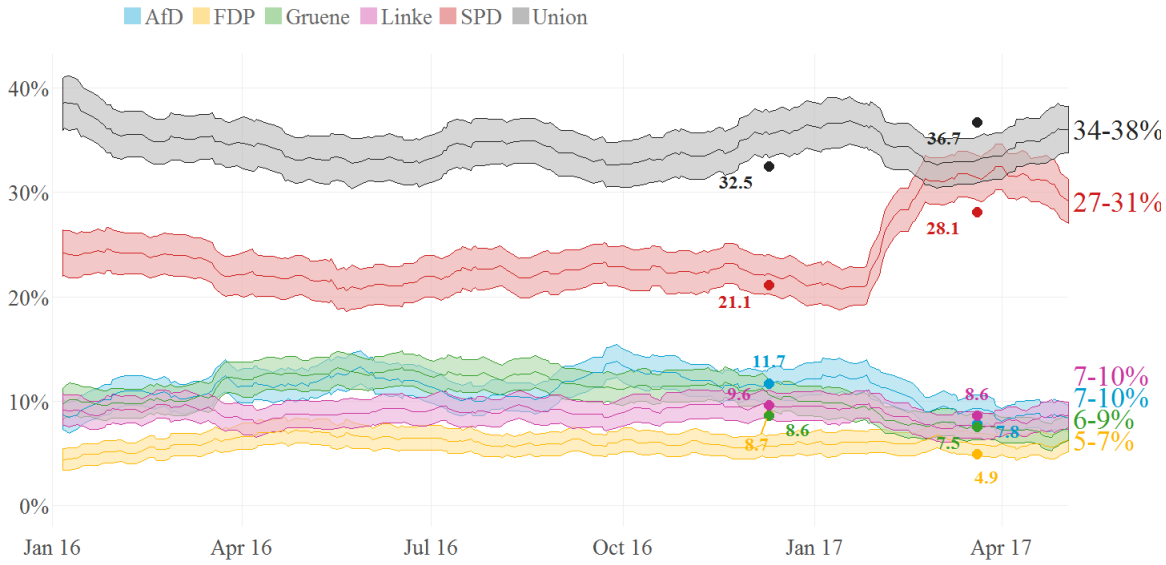

Figure 2: Weighted polls benchmarked to overall poll's rolling average

Moritz Hemmerlein
Alexander Sacharow

Table 1: Raw, weighted and benchmark estimates

December Method	Union	SPD	FDP	Gruene	Linke	AfD	Other	RMSE
SZ Rolling Average	35.50	22.20	5.70	10.60	9.70	11.60	4.70	
GAV Exit Polls	32.50	21.10	8.70	8.60	9.60	11.70	7.80	2.17
GAR Census Data	24.60	21.60	5.50	10.60	9.50	14.00	14.20	5.55
Dalia Unweighted	23.80	20.90	6.10	10.50	10.50	13.80	14.40	5.83
March Method	Union	SPD	FDP	Gruene	Linke	AfD	Other	RMSE
SZ Rolling Average	33.10	31.40	5.80	7.60	7.70	9.30	5.10	
GAV Exit Polls	36.70	28.10	4.90	7.50	8.60	7.80	6.40	2.05
GAR Census Data	23.10	30.00	5.80	8.40	9.00	13.80	10.00	4.61
Dalia Unweighted	23.60	28.40	5.60	7.70	9.80	14.50	10.30	4.75

Notes: The benchmark is the rolling average of all German major polls as computed by *Süddeutsche Zeitung*. GAV stands for strata combined of the variables gender, age and self-reported vote at the last election. GAR indicates strata of gender, age and religion. The last row of each month represents the raw results without weighting as collected by Dalia Research.

Der Bundeswahlleiter. 2017. "Ergebnisse Der Representativen Wahlstatistik." https://www.bundeswahlleiter.de/bundestagswahlen/2013/ergebnisse/repraesentative-wahlstatistik.html.