AStA Advances in Statistical Analysis manuscript No.
(will be inserted by the editor)

KOALA: A new paradigm for election coverage

An opinion poll based "now-cast" of probabilities of events in multi-party electoral systems

Alexander Bauer · Andreas Bender · André Klima · Helmut Küchenhoff

Received: date / Accepted: date

Abstract
Common election poll reporting is often misleading as sample uncertainty is addressed insufficiently or not covered at all. Furthermore, main interest usually lies beyond the simple party shares. For a more comprehensive opinion poll and election coverage, we propose shifting the focus towards the reporting of survey-based probabilities for specific events of interest. We present such an approach for multi-party electoral systems, focusing on probabilities of coalition majorities. A Monte Carlo approach based on a Bayesian Multinomial-Dirichlet model is used for estimation. Probabilities are estimated, assuming the election was held today ("now-cast"), not accounting for potential shifts in the electorate until election day ("fore-cast"). Since our method is based on the posterior distribution of party shares, the approach can be used to answer a variety of questions related to the outcome of an election. We also introduce visualization techniques that facilitate a more adequate depiction of relevant quantities as well as respective uncertainties. The benefits of our approach are discussed by application to the German federal elections in 2013 and 2017. An open source implementation of our methods is freely available in the `R` package `coalitions`.

A. Bauer
Statistical Consulting Unit StaBLab, Department of Statistics, LMU Munich, Germany
Tel.: +49-89-2180-3197
Fax: +49-89-2180-5308
E-mail: alexander.bauer@stat.uni-muenchen.de
ORCID: 0000-0003-3495-5131

A. Bender
Statistical Consulting Unit StaBLab, Department of Statistics, LMU Munich, Germany
ORCID: 0000-0001-5628-8611

A. Klima
Statistical Consulting Unit StaBLab, Department of Statistics, LMU Munich, Germany

H. Küchenhoff
Statistical Consulting Unit StaBLab, Department of Statistics, LMU Munich, Germany

Keywords Election analysis · Opinion polls · Election reporting · Multinomial-Dirichlet · Bayes

1 Introduction

In multi-party democracies, approval of the government's and the opposition parties' work is usually measured by public opinion polls continuously conducted and published by various polling agencies. Reported quantities usually include the share of respondents that would vote for the respective political parties *if the election was held today* (party shares), the number of overall respondents and – often less prominent – information about sample uncertainty.

One party often does not obtain enough votes for a governance majority on its own, if the voting system based on proportional allocation of seats in parliament. Thus, multiple parties form a so-called *coalition* to jointly obtain the necessary majority of seats in parliament. Media usually reports the results of opinion polls by focusing on the reported party shares while ignoring sample uncertainty. This is misleading, especially if shares are used to infer the possibility of a majority for a specific coalition. For example, in the prelude to the 2017 German federal election, a coalition was oftentimes stated to "lose" its majority just because the reported joint voter share dropped below 50% from one opinion poll to the next (e.g., FAZ.net, 2017). Such interpretations are clearly inadequate as sample uncertainty (and often redistribution of votes) is not taken into account. This becomes especially problematic, when one or more parties are close to the country specific threshold of votes that has to be passed in order to enter the parliament. This was the case in the 2013 German federal election, where the reported share of the Free Democratic Party (FDP) was slightly above the 5% threshold but failed to enter the parliament on election night (cf. Section 1.2).

Beyond ensuring proper reporting of sample uncertainties, in our opinion, the focus of election poll reporting should in general be shifted away from the reported party shares. Instead, election coverage should focus on the most relevant question, i.e., how *probable* is a specific event or outcome, given the current political mood. As probabilities combine both, the reported shares and sample uncertainty in one number, they allow more precise as well as more adequate statements about specific events. Before an election, events such as the following are usually of interest:

- "Will a party obtain enough votes to enter the parliament (pass the threshold)?"
- "Will a party obtain the most (second most, third most, etc.) votes?"
- "Will a specific coalition obtain enough votes (joint majority) to form a governing coalition?"

In this article, we present our approach for election and coalition analysis (in German: **Koal**itions-**A**nalyse (KOALA)) that estimates probabilities for any such events, referred to as *p*robability *o*f *e*vent (POE) in the following. In Section 3, we will illustrate that the POE brings more value to opinion poll based election coverage. It is important to note that we quantify the contemporary political mood and the resulting event probabilities ("now-cast"), not taking into consideration potential shifts until election day ("fore-cast"). Approaches for predicting future election outcomes based on past information can, e.g., be found in Graefe (2017) or Norpoth and Gschwend (2010). A special focus is put on multi-party proportional representation electoral systems and the estimation of probabilities for (joint) majorities. POEs are estimated by Monte Carlo simulations of election outcomes from the Bayesian posterior distribution of party shares conditional on current observed opinion poll data. Prior to the German general elections 2013 and 2017, results based on (an earlier iteration of) our approach already entered media reporting (cf. ZEIT ONLINE, 2013; Gelitz, 2017).

All methods discussed in this article are implemented in `R` (R Core Team, 2017) and are available in the open-source package `coalitions` (Bender and Bauer, 2018). A `shiny`-based (Chang et al., 2017) website `koala.stat.uni-muenchen.de` visualizes estimated coalition probabilities and is used to communicate the results for German federal and state elections to the general public. The process of fetching new polls, updating the website and sending out Twitter messages based on the newest results is automated and allows for an immediate transfer of the estimated POEs to media outlets as well as the general public.

1.1 Data basis

As data base for our calculations, we use opinion polls conducted by established polling agencies that quantify the electoral mood in a limited time frame (*if an election was held today*). For each of the two elections discussed in Section 3, we base the discussion on opinion polls published by major German polling agencies (i.e., Allensbach, Emnid, Forsa, Forschungsgruppe Wahlen, GMS, Infratest dimap and INSA), starting one year before each election. Opinion poll data from these polling agencies is collected by and made publicly available on `www.wahlrecht.de`. Application of our approach to other countries requires systematic access to respective polling data.

1.2 Motivating example

In the last opinion poll conducted before the German federal election 2013 (Forsa, 2013), it was of special interest whether the conservative "Union" – i.e., the union of the parties CDU (Christian Democratic Union) and CSU (Christian Social Union in Bavaria) – and the liberal FDP would together once again obtain enough votes to form the governing coalition (cf. Table 1).

Table 1 Reported party shares in the Forsa opinion poll for the German federal election, published September 20th, 2013 with $n = 1995$ respondents.

Union	SPD	Greens	FDP	The Left	Pirates	AfD	Others
40%	26%	10%	5%	9%	2%	4%	4%

The German election system mandates a 5% vote share threshold for parties to enter the parliament. Votes for parties below this threshold and without at least three successful direct candidates are redistributed (proportionally) to parties above it. Table 2 depicts the resulting redistributed party shares given the poll in Table 1. It illustrates, that Union-FDP with its reported joint 45% voter share before redistribution, would obtain 50% of parliament seats after redistribution. Thus, ignoring uncertainty it could be concluded that a majority for this coalition is possible, if party shares would increase slightly for one of the two parties.

Table 2 Redistributed party shares based on the Forsa opinion poll for the German federal election, published September 20th, 2013 with $n = 1995$ respondents (cf. Table 1). Party shares marked with "–" indicate parties that would not pass the 5% threshold (before redistribution).

Union	SPD	Greens	FDP	The Left	Pirates	AfD	Others
44.44%	28.89%	11.11%	5.56%	10.00%	–	–	–

However, such a consideration completely ignores sample uncertainty and the probabilistic nature of the outcome. If the poll in Table 1 is representative for the electoral mood, one would expect that the FDP enters the parliament (passes the 5% threshold) with a probability of about 50%. Thus, the (posterior) distribution of the joint voter share is bimodal and also depends on whether the other "small" parties close to the 5% threshold enter the parliament. The example also illustrates that discussion of reported party shares can become very complex, due to sample uncertainty and the multitude of different outcomes this uncertainty entails. We therefore argue that probability based reporting of opinion poll results can answer the actual question of interest ("Will a coalition of Union-FDP obtain enough votes to obtain a majority of seats in the parliament?") more directly, while adequately taking into account the inherent uncertainty.

The remainder of the article is structured as follows: Section 2 introduces the Bayesian method used to estimate POEs as well as some details on the aggregation of multiple opinion polls and the correction of rounding errors. Section 3 illustrates the application of the approach to opinion polls in advance of the 2013 and 2017 German federal elections. A summary and discussion are presented in Section 4.

2 Methods

2.1 Estimating event probabilities from reported party shares

To estimate the POE conditional on opinion poll results we use the Bayesian framework to construct the *posterior* distribution of the party shares based on distribution of the reported shares and an assumption about their *prior* distribution.

Let $X_1, \ldots, X_P$ be the reported opinion poll count of respondents that would elect party $p,\ p = 1, \ldots, P$ (vote count). For example, in Table 1 the reported vote count for the *Union* is given by $X_1 = .40 \cdot 1995 = 798$. We assume that $(X_1, \ldots, X_P)^{\mathrm{T}}$ follows a Multinomial distribution

$$\boldsymbol{X} = (X_1, \ldots, X_P)^{\mathrm{T}} \sim Multinomial(n, \theta_1, \ldots, \theta_P), \tag{1}$$

where n is the sample size of the opinion poll and $\theta_p,\ p = 1, \ldots, P$ indicates the probability of party p being selected. Further assuming a simple random sample, i.e ignoring possible bias, θ_p represents the true percentage of voters for party p in the general population. Given one (pooled) survey, the distribution of the observed vote counts $\mathbf{x} = (x_1, \ldots, x_P)^{\mathrm{T}}$ is denoted by $f(\mathbf{x}|\boldsymbol{\theta})$.

For the prior distribution of the true party shares $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_P)^{\mathrm{T}}$ we chose an uninformative prior distribution (Jeffrey's prior; see Gelman et al. (2013))

$$\begin{aligned} &\boldsymbol{\theta} = (\theta_1, \ldots, \theta_P)^{\mathrm{T}} \sim Dirichlet(\alpha_1, \ldots, \alpha_P), \\ \text{with} \qquad &\alpha_1 = \ldots = \alpha_P = \frac{1}{2}, \end{aligned} \tag{2}$$

denoted by $p(\boldsymbol{\theta}|\boldsymbol{\alpha})$. As the Dirichlet distribution is a conjugate prior to the Multinomial distribution, the resulting posterior distribution (3) of parameters $\boldsymbol{\theta}|\mathbf{x}$

$$f(\boldsymbol{\theta}|\mathbf{x}) = \frac{f(\mathbf{x}, \boldsymbol{\theta})}{f(\mathbf{x})} = \frac{f(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}, \boldsymbol{\alpha})}{f(\mathbf{x})} \tag{3}$$

$$\propto f(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\alpha}) \tag{4}$$

$$\propto \prod_{p=1}^{P} \theta_p^{x_p} \cdot \prod_{p=1}^{P} \theta_p^{\alpha_p - 1} = \prod_{p=1}^{P} \theta_p^{x_p + \alpha_p - 1}, \tag{5}$$

is again a Dirichlet distribution with

$$\boldsymbol{\theta}|\mathbf{x} \sim Dirichlet(x_1 + 1/2, \ldots, x_P + 1/2). \tag{6}$$

Given the multivariate posterior (6) and using Monte Carlo simulations, POEs can be deduced for many types of events by simulating election results from (6) and calculating the percentage of simulations in which the event of interest occurred. This includes the probabilities for specific majorities derived from a complex, country-specific system of rules for the calculation of seats in the parliament (Sainte-Lague/Scheppers in Germany; Grofman and Lijphart (2003)). For example, given the Forsa poll introduced in Section 1.2, the coalition of Union-FDP obtained a majority of seats in 2 633 of 10 000 simulations, which equals an POE of 26% (see Section 3 for more details).

If it is known that estimates of specific party shares are biased for some opinion polls/agencies, this information could be included in the model by using an informative prior distribution. The prior parameters α_p would then be adjusted to have higher or lower values, respectively. However, such biases of polling agencies are hard to quantify as the true party share in the electorate is only known on election days. For our analyses, we therefore use the uninformative prior (2).

2.2 Aggregation of multiple polls (Pooling)

In the presence of multiple published opinion polls, pooling is used to aggregate multiple polls in order to reduce sample uncertainty. To ensure a reliable pooling regarding the current public opinion, we only use polls published within a certain period of time (e.g. 14 days) and only use the most recent survey published by each polling agency.

Considering a single poll i, the observed vote count X_{ip} for each of P parties follows a multinomial distribution with sample size n_i and underlying, unknown party shares θ_p in the population. Pooling over multiple such polls as independent random samples leads to another multinomial distribution for the summed number of votes $\sum_i X_{ip}$:

$$\left(\sum_i X_{i1}, \ldots, \sum_i X_{iP}\right)^{\mathrm{T}} \sim Multinomial\left(\sum_i n_i, \theta_1, \ldots, \theta_P\right). \quad (7)$$

Further analyses, however, showed that polls from different (German) polling agencies are correlated and the independence assumption does not hold. Therefore, we adjust the resulting multinomial distribution by using an *effective sample size* (Hanley et al., 2003), reflecting that the aggregation over multiple correlated polls does not contain information of a sample with $\sum_i n_i$ observations.

Quantification of pairwise correlation is done based on the variance of the party share difference between two polls for a specific party. The following holds for two independent random samples from poll A and B:

$$
\begin{aligned}
Var(X_{Ap} - X_{Bp}) &= Var(X_{Ap}) + Var(X_{Bp}) - 2 \cdot Cov(X_{Ap}, X_{Bp}) \\
\Leftrightarrow \quad Cov(X_{Ap}, X_{Bp}) &= \frac{1}{2} \cdot \left(Var(X_{Ap}) + Var(X_{Bp}) - Var(X_{Ap} - X_{Bp})\right).
\end{aligned}
\tag{8}
$$

We take $Var(X_{Ap})$ and $Var(X_{Bp})$ as the theoretical variances of the binomially distributed, reported vote count and estimate $Var(X_{Ap} - X_{Bp})$ based on the observed differences between the reported party shares. Having done so, it is possible to estimate the covariance $Cov(X_{Ap}, X_{Bp})$ and accordingly also the correlation. As the binomial variance is directly proportional to sample size, the effective sample size n_{eff} can be defined as the ratio between the estimated variance of the pooled sample and the theoretical variance of a sample of size one:

$$
n_{\text{eff}} = \frac{Var(\text{pooled})}{Var(\text{sample of size 1})}.
$$

In the case of two surveys,

$$
Var(\text{pooled}) = Var(X_{Ap}+X_{Bp}) = Var(X_{Ap})+Var(X_{Bp})+2\cdot Cov(X_{Ap}, X_{Bp})
$$

and $Var(\text{sample of size 1})$ the theoretical variance of the pooled share.
Considering the party-specific correlations between 20 surveys conducted by the two German polling agencies that provide updates most regularly, Emnid and Forsa, we on average end up with a medium high positive correlation, using mean party shares and sample sizes per institute for the theoretical variances. Comparisons of other agencies were not performed as too few published surveys that cover comparable time frames were available. For simplicity, we do not recalculate the correlation for each simulation, but rather set the correlation used in our calculations to 0.5, i.e., a medium positive correlation. For convenience, the calculation of n_{eff} is based on the party with most votes, as the specific party choice only marginally affects the results.

Considering for example two polls with 1 500 and 2 000 respondents, respectively, and a pooled share of 40% for the strongest party, the method leads to an effective sample size of $n_{\text{eff}} = 2\,341$. Thus, the method reduces sample uncertainty compared to using a single poll, while being quite conservative compared to the assumption of independence which would lead to an aggregate sample size of $1\,500 + 2\,000 = 3\,500$.

As noted above, in practice we use a time window of 14 days, i.e., all surveys published in the last 14 days are included in the calculation of the pooled sample. For some elections (e.g., state elections), opinion polls are updated very rarely. In such cases the time window and pooling procedure could be further modified, e.g., by including all surveys published within 14 days with full weight (using their reported sample size), and all surveys that were published between 15 and 28 days ago with halved weight (using the halved sample size).

2.3 Correction of rounding errors

Polling agencies usually only publish rounded party shares and raw data is not available. Therefore, we adjust the reported data by adding uniformly distributed random noise to the observed party shares $\tilde{\theta}_p$ in order to avoid potential biases caused by the use of rounded numbers:

$$\begin{aligned} &\tilde{\theta}_{p,adj} = \tilde{\theta}_p + r_{\gamma,p}, \\ &\text{with} \qquad r_{\gamma,p} \sim U[-\gamma, \gamma]. \end{aligned} \tag{9}$$

The correction coefficient γ is chosen according to rounding accuracy. E.g., for data rounded to 1% steps we use $\gamma = 0.5\%$. After random noise was added the adjusted shares are rescaled to ensure that all adjusted party shares $\tilde{\theta}_{p,adj}$ sum to 100%. Overall, instead of using rounded numbers and simulating n_s values from the resulting posterior, we perform n_s simulations where we first adjust the party shares using individually drawn $r_{\gamma,p}$ and then simulate one observation from each resulting posterior.

2.4 Potential sources of bias

Multiple sources of bias exist that can negatively affect the results. These sources are mainly based on (a) the assumpion that the underlying opinion polls estimate the true party shares without systematic bias, and (b) the assumption of a Multinomial distribution for the reported voter numbers.

Regarding systematic biases of specific polling agencies, one can differentiate between a *house bias* and an *industry bias* (see e.g. Pickup et al., 2011). A house bias quantifies systematic errors made by an individual polling agency, e.g. if an agency steadily overestimates the share of a specific party. An industry bias describes a bias of the whole polling industry, i.e. a systematic error that does not disappear when averaging over all polling agencies. Even if – as noted in Section 2.1 – such biases could be dealt with in our approach by using an appropriate informative prior, both types of bias are difficult to quantify as the true party shares in the electorate are only known on election days. Moreover, a direct comparison of the last pre-election polls with the final election outcome would not necessarily be appropriate as a relevant proportion of voters tends to make their decision only shortly before or even on election day. Analyzing the 2018 Bavarian state election, Küchenhoff et al. (2018) conducted a telephone survey in the two weeks running up to the election, i.e. the interval under study of most pre-election polls. In this survey, $\sim$ 19% of respondents stated to still be undecided (while $\sim$ 20% refused to answer). Accordingly, comparing election results with polls before the election would result in an adequate estimate of the error of *for-casts*, but not automatically of the error of *now-casts* on which we focus.

The assumed Multinomial distribution of voter counts can become inadequate based on the exact survey design of polling agencies, which comprises specific sampling designs and post-sampling weightings. This might lead to a higher or lower variance than achieved with the assumed Multinomial distribution and the sample size stated by the polling agency. Within our framework, this could be dealt with by adjusting the effective sample size. However, as polling agencies usually do not publish the unadjusted results, their survey design and weighting schemes, a proper quantification is again unfeasible. Polling agencies instead tend to publish sampling errors that underlie the reported party shares (see e.g. Forschunsgruppe Wahlen e.V., 2019). Since the sampling errors reported by major German polling agencies are rather similar to the uncertainty induced by our corresponding Multinomial distribution, we do not adjust the sample size of the individual polling agencies in our approach, but use the values reported by them.

3 Application

An earlier iteration of our method entered media reporting before the German federal elections 2013 and 2017 (cf. ZEIT ONLINE, 2013; Gelitz, 2017). We will discuss these two elections in order to elaborate the differences between standard media coverage of election polls – focused on the interpretation of the reported party shares – and our approach based on estimated POEs. Reported party shares as described in Section 1.1 were used as data basis. Polls from different agencies published within a time window of 14 days were aggregated (cf. Section 2.2). For the estimation of POEs $n_s = 10\,000$ simulations were performed.

3.1 German federal election 2013

In the legislative period from 2009 to 2013, the German government was formed by a coalition of the conservatives (Union) and the liberals (FDP). Before the election on September 22nd, 2013, the question whether the coalition could sustain its majority was therefore of main interest. The FDP played a key role, as the coalition could only be formed if the party had reached the minimum share of 5% of the votes. Figure 1 summarizes the reported party shares for the one year period prior to the election.

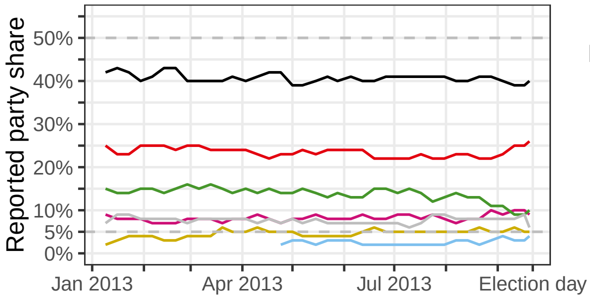

Fig. 1 Reported party shares based on Forsa opinion polls from October 2012 until election day on September 22nd, 2013. Shares for the AfD were only explicitly reported starting in April 2013. Before that time, the party is contained in "Others".

POE: FDP passing the 5% threshold

The poll-based prospect of FDP to successfully pass into parliament is visualized in Figure 2. As can be seen, the reported party share clearly exceeded the necessary hurdle of 5% only over short periods of time with maximum values of 6% (top left pane in Fig. 2). Similarly, the now-cast POE for the party to pass the threshold rarely rose over 50% (bottom left pane). In the last Forsa poll before election day, a party share of 5% was reported, stating that the event of FDP successfully passing into parliament was highly uncertain.

Comparing party shares and POEs, Figure 2 shows that relatively small changes in the reported party share can dramatically influence corresponding POE values, depending on the base level of the party share and – in this example – its closeness to the 5% threshold. In this regard, probabilities make it easier to deduce *relevant* information from opinion polls as they incorporate both the closeness of the reported shares to the relevant threshold as well as sample uncertainty. For example, party shares of 4% and 6% correspond to very definite POEs of near 0% and 100%, respectively and the reporting of such POEs leads to a much clearer perception of the current public opinion compared to the reported party share and survey sample size only.

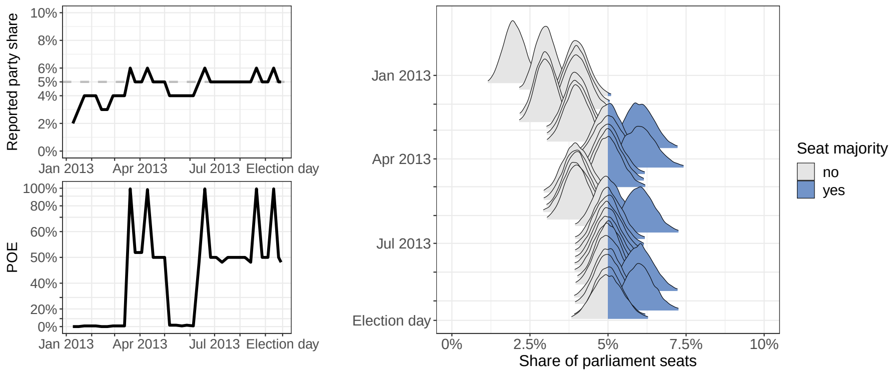

Fig. 2 Prospect of FDP to pass the 5% threshold before the German federal election in September 2013 based on Forsa opinion polls. Top left: Reported party share before redistribution. Bottom Left: Now-cast of the POE that FDP will pass the 5% threshold, based on 10 000 simulations. Right: Densities of the 10 000 simulated FDP shares. Areas under the density colored blue indicate the simulations in which the FDP passes the 5% threshold.

For the visualization of POEs we suggest to plot the distribution of simulated shares via density plots and to highlight the area associated with simulations where the event of interest occurred (see also Fig. 3). This has the advantage that POEs are communicated clearly (and intuitively) while the distribution of simulated shares additionally highlights the underlying uncertainty and the range of possible outcomes. Another advantage is that such visualizations can easily be extended to depict the development of POEs over time using so-called *ridgeline plots* (Wilke, 2017). In Figure 2 this plot type and the development of the POEs is compared to the observed FDP party shares usually reported in media.

To focus on the most relevant changes in the POEs, we propose the use of a skewed axis as shown in the bottom left of Figure 2. On this axis the range of values around 50% is stretched and the range of values near 0% and 100% is compressed. In this way, we put less weight on changes where an event is *still highly (im)probable* and emphasize more relevant changes after which an event gets more or less probable than 50%, respectively. Also, consistently using another axis for the estimated probabilities prevents confusion of POEs and voter shares.

POE: Union-FDP coalition majority

Figure 3 shows the simulated parliament seat shares for the coalition Union-FDP, based on the reported party shares in Table 1. The estimated density is clearly bimodal as the reported FDP share before redistribution equals exactly 5% and therefore FDP only enters the parliament in about 50% of the simulations. In this case, the estimated POE was 26%, thus a majority was observed in about one quarter of the simulations.

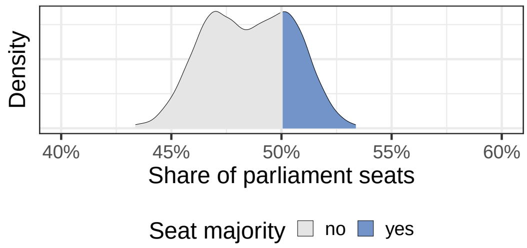

Fig. 3 Density of 10 000 simulated parliament seat shares for the coalition Union-FDP before the German federal election in September 2013 based on the Forsa opinion poll in Table 1. The area under the density colored blue indicates simulations with a Union-FDP majority, resulting in a POE of about 26%.

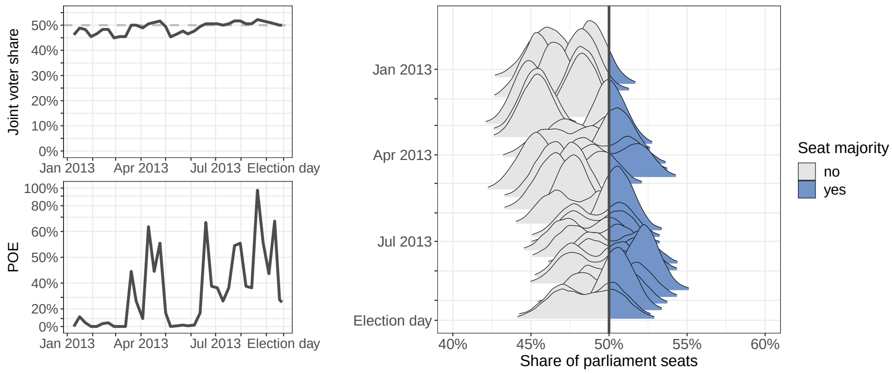

Fig. 4 Prospect of the coalition Union-FDP to obtain a government majority before the German federal election in September 2013 based on Forsa opinion polls. Top left: Reported joint voter shares after redistribution. Bottom left: Now-cast of the POE that the coalition will obtain a government majority, based on 10 000 simulations. Right: Densities of the 10 000 simulated parliament seat shares. Areas under the density colored blue indicate the simulations in which the coalition obtains a parliament seat majority.

Comparing the redistributed party shares and the POEs in Figure 4, it is again evident that even small changes in the joint redistributed voter share can make an immense difference regarding the POEs. Especially in the months preceding the election this leads to heavily fluctuating POEs based on the Forsa opinion polls. Furthermore, the development of the joint voter shares and the corresponding POEs nicely highlights another advantage of using such probabilities. The POEs do not only take into account the voter shares and sample uncertainty, but also implicitly cover the uncertainty regarding whether FDP passess the 5% threshold or not. Between the middle of June and the middle of July the POE drops from nearly 70% to under 30%, even though the joint voter share of Union-FDP only changes marginally. Taking into account the development of reported FDP party shares in Figure 2, it becomes clear that this drop is caused by a growing uncertainty of FDP passing into parliament. As the FDP share drops to 5% at the end of June also the POE for the Union-FDP seat majority declines heavily. Accordingly, the densities in the ridgeline plot in Figure 4 are unimodal or bimodal if the FDP share is clearly above/below or close to the 5% threshold, respectively.

3.2 German federal election 2017

After the German federal election in 2013, a "grand coalition" between Union and the social democratic SPD formed the government from 2013 to 2017. For the following election on September 24th, 2017, the goal of both Union and SPD was to obtain enough votes to form a coalition outside the grand coalition. Therefore, multiple potential coalitions were of interest before the election. In the following paragraphs, we will focus on the most prominently discussed coalitions, i.e., the Union-led coalition Union-FDP, and the SPD-led coalition of SPD, the Left party (Die LINKE) and the Green party, which – based on the joint voter share – was the strongest alternative to a Union-led government and was not clearly denied by the potential member parties until several weeks before election day. A question of major interest was also whether the right-wing party AfD, which slightly missed the 5% threshold in 2013, but gained support before the 2017 election, would become the third strongest party in parliament after Union and SPD. The pooled party shares before the 2017 election are shown in Figure 5.

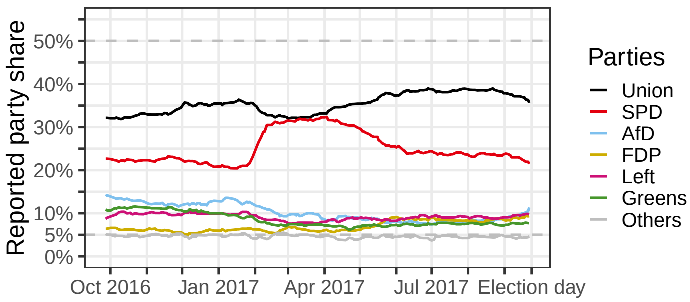

Fig. 5 Development of the pooled party shares from October 2016 until election day on September 24, 2017, based on a pooling time window of 14 days.

POE: Union-FDP coalition majority

Compared to the German federal election in 2013, the situation for a coalition between Union and FDP before the election in 2017 was quite different as FDP party shares were clearly above the 5% threshold (see Fig. 5) most of the time. However, as the share of Union was lower than in 2013, the joint redistributed voter share was mostly below 50%. As can be seen in Figure 6, the coalition had a joint, redistributed share of about 40% in October 2016 and reached its maximum share of nearly 49.8% about one month before election day.

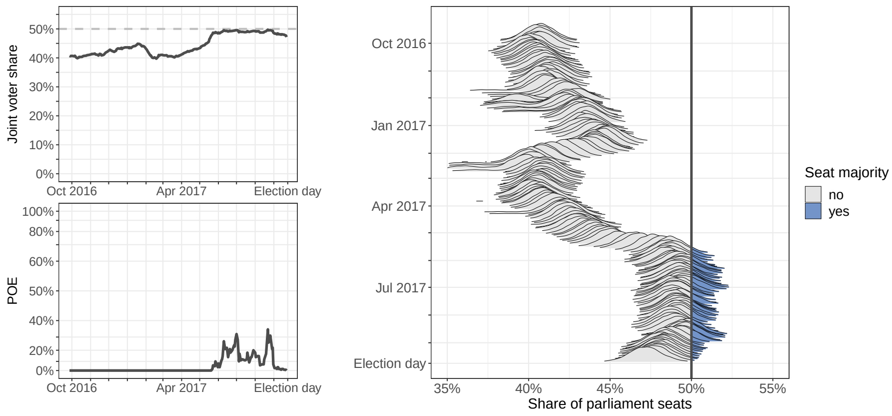

Fig. 6 Prospect of the coalition Union-FDP to obtain a government majority before the German federal election in September 2017 based on pooled opinion polls. Top left: Reported joint voter shares after redistribution. Bottom left: Now-cast of the POE that the coalition will obtain a government majority, based on 10 000 simulations. Right: Densities of the 10 000 simulated parliament seat shares. Areas under the density colored blue indicate the simulations in which the coalition obtains a parliament seat majority.

By comparison, the ridgeline plot in Figure 6 shows that joint voter shares below 48% correspond to very small POEs of $< 1\%$, based on pooled effective sample sizes of around 3 000. On the other hand, based on comparable sample sizes, shares of 49% and 49.5% corresponded to probabilities of around 14% and 25%, respectively. Overall, one month before election day the coalition had a good prospect reaching a seat majority based on a redistributed share of 49.8% and a POE of nearly 40%. However, until two days before the election the pooled share and the POE dropped to 47.4% and 0.4%, respectively, making a success of the two parties highly improbable.

POE: SPD-Left-Greens coalition majority
Regarding the party share development of the SPD, the year before the general election in 2017 was shaped by an unusually fast increase, starting at the end of January 2017, when Martin Schulz was elected to be the SPD chancellor candidate and a subsequent, steady decline from April 2017 on (see Fig. 5). Accordingly, the coalition between SPD, the Left and the Greens had their best joint poll results between February and May 2017 as is shown in Figure 7. The maximum share was reached in April with a redistributed voter share

of $\sim$ 50%, which corresponded to a POE of obtaining the parliament seat majority of $\sim$ 48%. Starting in April, the POE again dropped to negligibly small values. Shortly before election day, the joint voter share reached a value of around 41%, leading to POEs of practically zero. The ridgeline plot in Figure 7 again nicely visualizes the uncertainty underlying the event of interest. This is not only limited to parties forming the potential coalition, but also includes information about all other causes of uncertainty in the data. In November and December of 2016, for example, the seat share distribution is clearly bimodal as in a relevant share of simulations the FDP does not pass the 5% threshold and thus more votes are redistributed to other parties (including the SPD) in these cases.

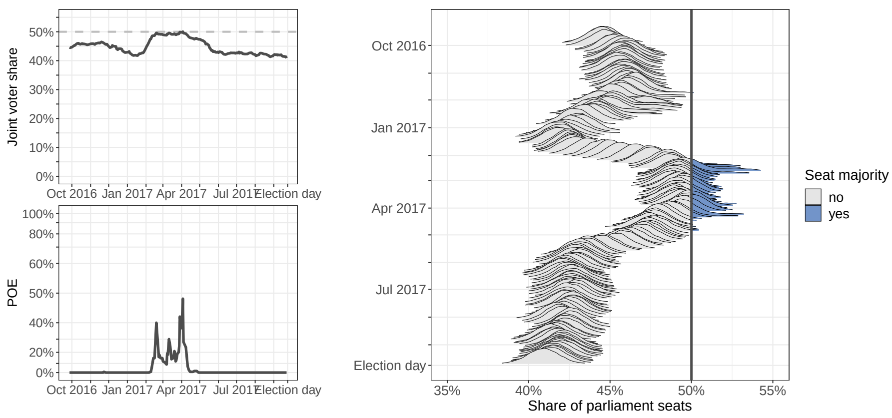

Fig. 7 Prospect of the coalition SPD-Left-Greens to obtain a government majority before the German federal election in September 2017 based on pooled opinion polls. Top left: Reported joint voter shares after redistribution. Bottom left: Now-cast of the POE that the coalition will obtain a government majority, based on 10 000 simulations. Right: Densities of the 10 000 simulated parliament seat shares. Areas under the density colored blue indicate the simulations in which the coalition obtains a parliament seat majority.

POE: AfD becoming third strongest party

Prior to the 2017 election, special interest was on the question which party would become the third largest party in parliament. With reported party shares of over 8%, the right-wing AfD had a very good prospect to become a member of the German parliament for the first time (see Fig. 8) and was polling close to other smaller parties. Using our KOALA approach, estimating the POE that the AfD becomes the third largest party in parliament is straightforward, adequately summarizing this event probability that simultaneously depends on all reported party shares.

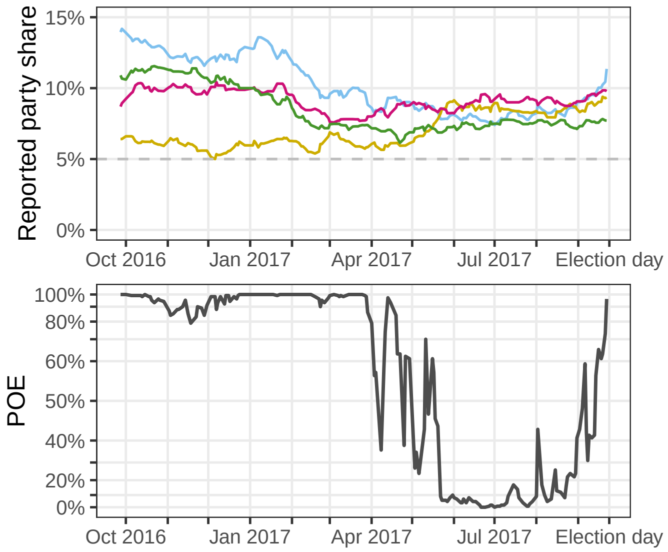

Fig. 8 Development of the prospect that AfD becomes the third largest party in parliament before the German federal election in September 2017 based on pooled opinion polls. Top: Reported voter shares before redistribution. Bottom: Now-cast of the POE that AfD will become the third largest party, based on 10 000 simulations.

In the year before the general election in 2017, reported AfD party shares underwent strong fluctuations. In January 2017 the party had a 3.9% lead over the Left and the Greens (corresponding to an estimated POE of becoming the third largest party in parliament of 100%). Subsequently, the AfD share dropped 1.9 percentage points behind the Left in June (corresponding to a 1.2% probability) and rose back to a 1.7 percentage point lead (in voter shares) over the Left and the FDP shortly before election day (corresponding to a POE of 96.8%).

4 Discussion

In this article we introduce a Bayesian approach to now-cast probabilities of election outcome related events (POEs) based on publicly available opinion poll data Sample uncertainty is reduced by combining polls from multiple polling agencies, while taking into account the correlation between polls. Rounding errors of reported party shares are also accounted for. The estimated POEs are easy to communicate to the general public and provide a new paradigm for election coverage and reporting of opinion polls. More specifically, the focus on event probabilities allows to capture changes in the current political mood and their effect on events of interest more intuitively and comprehensively while taking into account the potentially complex range of possible outcomes due to uncertainty inherent in the reported party shares. The value of POE based reporting was illustrated by application to the 2013 and 2017 German federal elections. Various visualization techniques were used to make the result accessible to the general public. POEs are continuously updated and made available on a dedicated website. The methods for pre-processing and calculation of POEs are available in the open-source `R` package `coalitions` that allows for a straightforward application of the method to any multi-party electoral system.

Our approach is based on results of opinion polls conducted by different polling agencies. Consequently, problems with well known sources of bias induced by non response, incorrect answers, non coverage etc. can occur. All institutes perform some correction methods to reduce these biases, using weighting or related procedures, but they do not make their procedures completely transparent. Currently we perform no additional assessment or correction for potential biases of individual polling agencies. We also do not perform forecasts or use any other information outside of reported party shares (and sample size).

Our long-term goal is to make probability based reporting of opinion poll based election coverage available to the general public. One limitation in this regard is the availability of properly structured data and its accessibility through an application programming interface (API). The creation of such a data base would greatly enhance the development of our and other methods for the analysis of opinion polls. Future iterations of our algorithms could also focus on enhancing the computation speed of the Monte Carlo based calculation of POEs, making updates available to the general public even more quickly, ideally in real time when new opinion polls are published.

References

Bender, A. and Bauer, A. (2018). coalitions: Coalition probabilities in multi-party democracies. *The Journal of Open Source Software*, **3**(23). doi: 10.21105/joss.00606. URL `http://joss.theoj.org/papers/10.21105/joss.00606`.

Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. (2017). *shiny: Web Application Framework for R.* URL `https://CRAN.R-project.org/package=shiny`. R package version 1.0.5.

FAZ.net (2017). Umfrage zur Bundestagswahl: Schwarz-Gelb verliert die Mehrheit. `http://archive.is/SuXVt`. Accessed 15 February 2018.

Forsa (2013). Wenn am nächsten Sonntag Bundestagswahl wäre... `http://archive.is/f9vse`. Accessed 15 February 2018.

Forschungsgruppe Wahlen e.V. (2019). *Methodik der Politbarometer-Untersuchungen.* As of January 2019.

Gelitz, C. (2017). Können die aktuellen Umfragen noch falschliegen? `http://archive.is/JydHd`. Accessed 15 February 2018.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian data analysis*, volume 3. CRC press Boca Raton, FL.

Graefe, A. (2017). The pollyvote's long-term forecast for the 2017 german federal election. *PS: Political Science & Politics*, **50**(3), 693–696.

Grofman, B. and Lijphart, A. (2003). *Electoral Laws and Their Political Consequences*. Algora Publishing. ISBN 978-0-87586-267-5. Google-Books-ID: o1dqas0m8kIC.

Hanley, J. A., Negassa, A., and Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: an orientation. *American journal of epidemiology*, **157**(4), 364–375.

Küchenhoff, H., Thurner, P. W., Klima, A., Mauerer, I., Knieper, T., Haupt, H., Mang, S., Schnurbus, J., Walter-Rogg, M., and Heinrich, T. (2018). *Universitätsstudie zur Bayernwahl USBW 18 (München – Passau – Regensburg). Erste Ergebnisse – Oktober 2018.*

Norpoth, H. and Gschwend, T. (2010). The chancellor model: Forecasting german elections. *International Journal of Forecasting*, **26**(1), 42–53.

Pickup, M., Matthews, J. S., Jennings, W., Ford, R., and Fisher, S. D. (2011). Why did the polls overestimate Liberal Democrat support? Sources of polling error in the 2010 British general election. *Journal of Elections, Public Opinion and Parties*, **21**(2), 179–209.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.R-project.org/`.

Wilke, C. O. (2017). *ggridges: Ridgeline Plots in 'ggplot2'*. URL `https://CRAN.R-project.org/package=ggridges`. R package version 0.4.1.

ZEIT ONLINE (2013). Serie: Wahlistik. `http://archive.is/1SU1I`. Accessed 15 February 2018.