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

Peter Selb (corresponding author)
Department of Politics & Public Administration
University of Konstanz
78457 Konstanz, Germany
phone: +49 7531 88 2321
email: peter.selb@uni.kn

Simon Munzert
Department of Politics & Public Administration
Graduate School of Decision Sciences
University of Konstanz

ABSTRACT

This paper 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 empirically assess 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.

INTRODUCTION

Election forecasting as an academic discipline is still in its infancy in Germany.[1] Only since the 2002 Bundestag election have the trial heat polls that regularly whip the passions of journalists and pundits during campaigns faced competition from scholars of elections employing alternative forecasting strategies.[2] The predominant academic approach uses regression analysis to predict future election returns from measures of government popularity, economic conditions and other fundamental variables – factors that are known well ahead of an election. Despite their considerable lead time, such regression-based forecasts have often proved more accurate even than polls conducted right before the election.[3] Most notably, Gschwend and Norpoth's 'chancellor model' pretty much exactly predicted the government parties' vote shares six months in advance of the 2002 and 2005 elections,[4] whereas the latter election was widely considered a 'debacle' for the polling companies.[5] No wonder then that current academic forecasts of the 2013 Bundestag election exclusively rely on different varieties of this same frugal yet successful empirical strategy,[6] while the rich body of data available from pre-election polls on voting intentions is largely dismissed by the scientific community.

We propose an alternative approach to election forecasting that capitalizes on the abundance of survey data collected during an election campaign. We agree that, over and above the problems inherent in any method of forecasting, trial heat polls suffer from specific methodological flaws, such as coverage, nonresponse, measurement and adjustment errors.[7] In so far as such errors occur systematically, however, a regression analysis based on historical polling and election data may help to get things straight under specific conditions.[8] To the extent that errors occur at random, on the other hand, a Bayesian-flavoured approach may be utilized to combine data from various polls in order to cancel out these errors, and to make forecasts more efficient.[9]

In the following section, we will lay out our approach in some more detail. Section 3 describes the polling data available and specifies which polls we eventually use for forecasting. Section 4 provides our forecast of the party vote shares at the 2013 Bundestag election that was held on 22 September. The forecast was published one month ahead of the election on Social Science Research Network (SSRN).[10] A final section that evaluates our forecast in retrospect was added after the election.

OUR APPROACH

Our approach consists of five steps: assessing the poll-vote link based on historical data, choosing an appropriate time window, extrapolating to the upcoming election using current polls, combining forecasts from various polls, and resampling the estimates to get uncertainty measures.

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]

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

The rationale motivating this step is that the relationship between polls and votes is not necessarily one-to-one. Pre-election polls usually ask respondents which party they would vote for if the election were to be held next Sunday. Answers to this hypothetical item are not to be confused with actual voting behaviour, and their distribution does not have to correspond to the election result. Some respondents may change or only then make up their minds between the time of the poll and the election date; others may then abstain from the election; yet other groups of future voters may fail to respond to the so-called 'Sonntagsfrage',[12] or to the survey altogether. All these phenomena will bias polls to the extent that the underlying causal processes, survey and electoral participation, opinion formation etc., are linked to partisan preferences. Provided the same phenomena recur from poll to poll and from election to election, however, a simple regression model depicted in Equation (1) will help detecting and eventually adjusting estimates of party vote shares for those regularities. For example, some authors observe that parties that perform poorly in early polls usually catch up until election day and vice versa.[13] Parameter estimates of α greater than zero and β less than one would indicate such 'regression-to-the-mean' phenomena.

Another possibility would be that polls systematically over- or underestimate the vote shares of certain parties, for instance, due to differential survey participation and misreporting propensities among party electorates.[14] Such 'industry effects' would be reflected in estimates of the party-level variance, σ_{ω}^{2}, substantively different from zero.[15] Relatedly, figures from some polling firms may consistently be more favourable toward one or the other political party.[16] Such 'house effects' would become apparent in a substantial estimate of the variance at the level of the party-institute interaction, σ_{ξ}^{2}.[17] Based on historical election and polling data, we can estimate the parameters $\alpha, \beta, \sigma_{\omega}^{2}, \sigma_{\xi}^{2}$ and σ_{ψ}^{2} with restricted maximum likelihood (REML). By way of Empirical Bayes prediction,[18] we can also retrieve predictions of the group-level residuals ω_j and ξ_{jk}. The estimation is implemented in R using the lme4 package.[19]

Step 2. Choosing a time window

Given the often observed volatility of trial heat polls,[20] the parameters of the above model can hardly be expected to be constant over the course of an election campaign, nor can we assume our model to perform equally well all the times. A simple way to accommodate this complication is to specify multiple time windows, to fit Equation (1) to election and polling data from each window, and to eventually select the temporal specification that yields the best fit to historical data.[21] Evidently, the window needs to be sufficiently long to include enough polls to estimate the parameters. We choose a window length of 12 weeks for our empirical analysis, and proceed in monthly intervals, starting from the respective election date.

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:

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

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,

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

where

$$\hat{V}_{ijk} = \frac{\text{vote}_{ijk}(1 - \text{vote}_{ijk})}{n_i - 1} \,, \tag{4}$$

and *n* is the sample size of poll *i*.[22] The reasoning behind the data combination is straightforward: we gain strength by pooling estimates from various polls, and by giving higher weight to more precise poll-specific estimates.

Step 5: Resampling the estimates

Step 4 is replicated in order to assign measures of accuracy to our estimates. In particular, we randomly draw *M* subsamples *m* of size $N_m = (N-2)^2/(N-1)$ from our initial sample of *N* polls,[23] which yields party-specific distributions of M estimates, $vote_{jm}$. We take the mean of each distribution as our point forecast of party j's vote share at the upcoming election. We will also present (nonparametric) confidence intervals for the forecasts in terms of the distributions' centiles to depict uncertainty around these points.

DATA

To implement our strategy, we need both historical election and polling data to fit Equation (1), as well as current polls to extrapolate to the upcoming election using Equations (2) to (4). Fortunately, wahlrecht.de, an independent internet site on elections, electoral rules, and voting rights in Germany,[24] provides a full real-time collection of published vote intention surveys from prominent polling firms (IfD Allensbach, TNS Emnid, Forsa, Forschungsgruppe Wahlen, GMS, Infratest dimap), covering the five Bundestag election campaigns between 1998 and 2013. We accessed the data on 12 August 2013, six weeks before election date on 22 September. The observation period covers a total of 2,157 polls.

As to the temporal specification of our model, Figure 1 presents absolute differences between averaged model estimates and actual party vote shares for alternative twelve-week time windows, running from month 0 to 23 to election.[25] The dashed lines represent election-specific cross-party averages of those differences, while the solid line indicates cross-election averages. Unsurprisingly, our poll-based forecasts tend to converge to the actual election result relatively shortly before the election date.[26] However, the lowest average forecasting error is achieved with a time window specification of 8 to 10 months to election.[27] Looking at the individual elections, this is particularly true for the 2005 election, but it also tends to hold for 2009 and, to a minor extent, 2002. An ad hoc account of this recurring non-monotonic pattern is unbalanced campaigning where parties mobilize their electorates asynchronously in the last few months before the election.[28] Exploiting this empirical regularity in our forecast affords us a lead time of 8 months to the election. This is markedly longer than the two,[29] three[30] and five to six months[31] lead time of the alternative approaches, and thus defies the objection that trial heat polls converge to the actual election outcome, if ever, too shortly before the election date to make for a worthwhile forecast.[32]

FIGURE 1 ABOUT HERE

Figure 1. Average absolute party-specific forecasting errors by 12-week time window beginning at a given month to election. Dashed lines indicate election-specific, the solid line indicates overall averages. Colours online only.

FORECAST OF THE 2013 BUNDESTAG ELECTION

Table 1 gives REML parameter estimates for Equation (1) using historical polling data from 8 to 10 months before the election. The estimated intercept is indeed greater than zero, while the regression weight of the poll variable is markedly below unity, which indicates that parties that score particularly high in the polls tend to suffer losses on election day and vice versa. The estimate of the party-level variance is substantive, while it is practically zero in the case of the party-institute interaction. To illustrate, Table 2 presents empirical Bayes predictions of the party-level residuals, $\hat{\omega}_j$. These estimates vary considerably across parties, from -1.9% for the parties lumped in the 'Others' category to 2.5% for the SPD. While these party-level effects were introduced to account for the possibility that the polls might systematically over- or underestimate some parties' vote shares (see above), their empirical interpretation is somewhat tricky, since they also capture plain size differences between the parties. Thus, we would naturally expect to observe higher estimates for bigger parties (CDU/CSU, SPD) relative to smaller parties. However, the fact that the estimate is higher for the smaller of the two major parties (2.5% for the SPD as compared to less than 1% for the CDU/CSU) supports what we have already observed in Figure 1: polls generally tend to overestimate CDU/CSU, but not SPD, vote shares. Finally, as to potential 'house effects', the estimated party-institute variance σ_{ξ}^2 is essentially zero, which suggests that these are negligible here.

Coefficient	*Estimate (S.E.)*
Intercept α	0.0201 (0.0072)
Poll β	0.8742 (0.0216)
Party-level variance σ_{ω}^2	0.0002
Party-institute-level variance σ_{ξ}^2	0.0000
Residual variance σ_{ψ}^2	0.0004

Table 1. REML estimates of the model of party vote shares in past elections, see Equation (1). 123 polls conducted 8 to 10 months before the 1998 to 2009 elections are included.

Party	*Estimate*
CDU/CSU	0.0073
SPD	0.0247
B'90/Die Grünen	-0.0083
Die Linke	-0.0043
FDP	-0.0008
Others	-0.0187

Table 2. Empirical Bayes estimates of the party-level posterior means, $\hat{\omega}_j$.

Now for the more exciting part, the model estimates are used to generate out-of-sample forecasts from 42 polls conducted 8 to 10 months before the 2013 election based on Equation (2). Figure 2 contrasts forecasts of the party vote shares at the 2013 election based on raw results (upper panel) with our model-based forecasts (lower panel). A look at the median poll-specific forecasts represented by the vertical lines in Figure 2 shows that our model expectably discounts CDU/CSU and, to a lesser extent, B'90/Die Grünen and 'other' parties' vote shares in the polls, while the opposite holds true for SPD and FDP. Given what we have learned about the historical link between polls and votes, this finding does not come as a surprise. Second, the model adjusted estimates exhibit lower variance than the raw polling results, shown by narrower shapes of the distributions. This is because the model discounts deviant values to some extent. Table 3 summarizes the 42 poll-specific forecasts according to steps 3 and 4 above. Bootstrapped means and 80% confidence intervals (based on 1’000 replications for each party-specific combination) are reported.

According to our forecast, the CDU/CSU would clearly maintain the plurality of votes, followed by the SPD, B’90/Die Grünen, Die Linke and FDP, all of which would pass the 5% threshold. Although the ‘other’ parties are predicted to gain 6.5% of the vote, chances are low that one of these residual parties overcomes the hurdle, since these votes would be scattered among several parties. Assuming a proportional vote-seat translation, neither of the coalitions preferred by the two grand parties, a CDU/CSU-FDP coalition or an SPD-B’90/Die Grünen coalition would gain a majority in the Bundestag (46.9% for the current coalition, 44.8% for an SPD-B’90/Die Grünen alliance). In contrast, all remaining coalitions discussed in the media are forecasted to reach a parliamentary majority: CDU/CSU-SPD 71.3%, CDU/CSU-B’90/Die Grünen 55.5%, SPD-FDP- B’90/Die Grünen 50.8%, SPD-B’90/Die Grünen-Die Linke 53.1%.

FIGURE 2 ABOUT HERE

Figure 2. Forecasts of the party vote shares at the 2013 Bundestag election: raw averaged polling results and model-based estimates, see Equation (2). 42 polls conducted 8 to 10 months before the 2013 election. Vertical lines represent median values, short vertical ticks single polls.

	Forecast	80% confidence interval	Official result	Forecasting error
CDU/CSU	0.381	[0.362; 0.398]	0.415	-0.034
SPD	0.282	[0.262; 0.304]	0.257	0.025

B’90/Die Grünen	0.135	[0.125; 0.143]	0.084	0.051
Die Linke	0.077	[0.068; 0.086]	0.086	-0.009
FDP	0.054	[0.048; 0.061]	0.048	0.006
Others	0.065	[0.053; 0.075]	0.110	-0.045

Table 3. Combined forecast for the upcoming election. Reported are bootstrapped precision-weighted means and 80% confidence intervals of 42 poll forecasts. Official results and forecasting errors added after the election.

RETROSPECTIVE EVALUATION AND AFTERTHOUGHT

While our prediction that the then-governing coalition of CDU/CSU and FDP would be replaced after the election came true,[33] our point estimates of the party vote shares were pretty wide off the mark, both in absolute terms and, to a minor extent, relative to other forecasting methods employed.[34] The average absolute forecasting error was 2.8%, ranging from a modest though decisive 0.6% for the FDP to a staggering 5.1% for B’90/Die Grünen (see Table 3).

How comes? Clearly, a data combination strategy will work best if figures from individual polls scatter randomly around their true yet unknown values. On the other hand, our relatively simple attempt to correct current polling errors based on experience will work best if those errors re-occur from election to election. A detailed inspection of our data suggests that both conditions are questionable in the present case (see Figures 2 and 3). Our model expected the majority of the polls to overestimate the CDU/CSU vote share when they in fact underestimated it. Given that there was not much of a movement in the CDU/CSU polling trends over the course of the campaign, we conclude that the historical polling error for these parties—and which our model tried to capitalize on—did not repeat in this election. On the other hand, the dramatic overestimation of the B’90/Die Grünen vote share seems to be rather a result of polling window selection. In the polls we used—8 to 10 months before the election— the Greens were consistently attested support of 14 to 15 percent of the voters. Our model discounted this to a certain extent, but could not compensate entirely for the substantive loss of support for the Greens over the last months before the election, where they dropped to about 9 percent—still more than the actual result. These figures indicate that the choice of the time window could have been overly optimistic and the comfortable lead time, which set our model apart from all other forecasting efforts, dearly purchased with a significant loss of precision.

The major problem with German polls seems to be that, at each election, the polling firms tend to err in the same direction, but that this effect seemingly reverts between elections – a pattern that resembles a kind of learning from past mistakes, but with overshooting. For instance, the bulk of polls overestimated the CDU/CSU vote share in 1998, (slightly) underestimated the vote for the Union parties in 2002, then again overestimated their vote shares in 2005 and 2009, just to underestimate them gain in the current election of 2013. A similar though less pronounced, oscillating pattern can be observed for the SPD and the smaller parties. Incorporating such dynamics into a regression model, perhaps by way of party-specific autoregressive terms, would be possible on principle, though the small number of time periods available (just four previous elections in our empirical case!) would currently complicate the estimation of its parameters in connection with German Bundestag elections. Contexts with longer polling traditions would be more conducive to the implementation of such a modeling approach.[35] For the time being, however, we have to conclude that correcting the polls using historical data in the German Bundestag setting is a difficult venture.

FIGURE 3 ABOUT HERE

Figure 3. Distribution of poll-specific forecasting errors in 2’157 trial heat polls, by party and election. Dashed vertical lines represent median forecasting errors. Solid vertical lines indicate the reference points (i.e., actual election results).

ACKNOWLEDGMENTS

We are grateful to Tobias Witt for superb research assistance. Further, we would like to thank the anonymous reviewers for valuable comments and suggestions.

SUPPLEMENTAL DATA AND RESEARCH MATERIALS

Data and code needed to replicate the results reported in this article can be accessed at https://github.com/simonmunzert/gerpol-forecasting-2013-election-polls.

ABOUT THE AUTHORS

Peter Selb is Professor of Survey Research at the University of Konstanz. His research interests include survey methods as well as political behaviour and public opinion, with a particular emphasis on elections and political representation.

Simon Munzert is a senior research fellow at the Chair for Survey Research, University of Konstanz, Germany. His research interests include methods of election forecasting, public opinion measurement, and techniques for web scraping and Big Data management.

NOTES

[1] B. Jérôme, 'Editor's Introduction for Forecasting the 2013 German Elections', *PS: Political Science & Politics* 46/3 (2013), pp.477-8.

[2] See T. Gschwend and H. Norpoth, '"Wenn am nächsten Sonntag...": Ein Prognosemodell für Bundestagswahlen', in H.-D. Klingemann and M. Kaase (eds.) *Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 1998* (Wiesbaden: VS Verlag für Sozialwissenschaften, 2001), pp.473-99; T. Gschwend and H. Norpoth, 'Prognosemodell auf dem Prüfstand: Die Bundestagswahl 2005', *Politische Vierteljahresschrift* 46/4 (2005): pp.682-8; B. Jérôme, V. Jérôme-Speziari and M.S. Lewis-Beck, 'A Political-Economy Forecast for the 2013 German Elections: Who to Rule with Angela Merkel?' *PS: Political Science & Politics* 46/3 (2013), pp.479-80; M.A. Kayser and A. Leiniger, 'A Benchmarking Forecast and Post-Mortem of the 2013 Bundestag Election' *German Politics* (forthcoming); H. Norpoth and T. Gschwend, 'Mit Rot-Grün ins Schwarze getroffen: Prognosemodell besteht Feuertaufe', in J.W. Falter, O.W. Gabriel and B. Weßels (eds) *Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 2002* (Wiesbaden: VS Verlag für Sozialwissenschaften, 2005),

pp.371-87; H. Norpoth and T. Gschwend, ‘The Chancellor Model: Forecasting German Elections’, *International Journal of Forecasting* 26/1 (2010), pp.42-53; H. Norpoth and T. Gschwend, ‘Chancellor Model Picks Merkel in 2013 German Election’, *PS: Political Science & Politics* 46/3 (2013): pp.481-2.

[3] See A. Gelman and G. King, 'Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?', *British Journal of Political Science* 23/4 (1993), pp.409-51, for an ingenious account of why long-shot regression-based forecasts are so successful while trial heat polls are so volatile and often inaccurate over the course of a campaign. Also see T. Plischke and H. Rattinger, '"Zittrige Wählerhand" oder invalides Messinstrument? Zur Plausibilität von Wahlprojektionen am Beispiel der Bundestagswahl 2005', in O.W. Gabriel, B. Weßels and J.W. Falter (eds), *Wahlen und Wähler. Analysen aus Anlass der Bundestagswahl 2005* (Wiesbaden: VS Verlag für Sozialwissenschaften, 2009), pp.484-509, for an application to the 2005 Bundestag election.

[4] Gschwend and Norpoth, ‘Prognosemodell auf dem Prüfstand’; Norpoth and Gschwend, ‘Mit Rot-Grün ins Schwarze getroffen’.

[5] L.M. Schaffer and G. Schneider, ‘Die Prognosegüte von Wahlbörsen und Meinungsumfragen zur Bundestagswahl 2005‘, *Politische Vierteljahresschrift* 46 (2005), pp.674-81, T. Plischke and H. Rattinger, ‘”Zittrige Wählerhand” oder invalides Messinstrument’. It should be mentioned that many polling firms were quite successful in forecasting the 2009 election, while the ‘chancellor model’ performed markedly worse than in 2002 and 2005, presumably due to the preceding grand coalition which partly thwarted the model’s logic (H. Norpoth and T. Gschwend, ‘The Chancellor Model’). Also see M. Klein, ‘Die "Zauberformel". Über das erfolgreiche Scheitern des Prognosemodells von Gschwend und Norpoth bei der Bundestagswahl 2005’. *Politische Vierteljahresschrift* 46, pp.689-91, for a critique of what he considers ad hoc adaptations of the 2002 chancellor model to the 2005 election scenario.

[6] Jérôme et al., ‘A Political-Economy Forecast’; Kayser and Leininger, ‘A Benchmarking Forecast’; ‚Norpoth and Gschwend, ‘Chancellor Model Picks Merkel’.

[7] See P. Selb, M. Herrmann, S. Munzert, T. Schübel and S. Shikano, ‘Forecasting Runoff Elections Using Candidate Evaluations from First Round Exit Polls’, *International Journal of Forecasting* 29/4 (2013): pp.541-7.

[8] See R. D. Fisher, R. Ford, W. Jennings, M. Pickup and C. Wlezien, ‘From Polls to Votes to Seats: Forecasting the 2010 British General Election’, *Electoral Studies* 30/2 (2011): pp.250-7; C. Wlezien and R.S. Erikson, ‘The Timeline of Presidential Election Campaigns’, *The Journal of Politics* 64/4 (2002): pp.969-93.

[9] See L.B. Brown and H.W. Chappell, Jr., ‘Forecasting Presidential Elections Using History and Polls’, *International Journal of Forecasting* 15 (1999): pp.127-35; S. Jackman, ‘Pooling the Polls over an Election Campaign’, *Australian Journal of Political Science* 40/4 (2005), pp.499-517; K. Lock and

A. Gelman, 'Bayesian Combination of State Polls and Election Forecasts', *Political Analysis* 18/3 (2010): 337-48.

[10] http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2313845 (accessed 20 Jul. 15 2015).

[11] That is, a normal distribution with mean zero and fixed variances σ^2 to be estimated from the data. At this stage, we ignore the statistical complications arising from the fact that party vote shares are bounded outcomes that sum to unity (see J. Aitchison, *The Statistical Analysis of Compositional Data* (London: Chapman & Hall, Ltd.)). Also note that our interest is in election-day outcomes, not varying party support over the course of the campaign. This distinguishes our model from recent space-state approaches to modelling the dynamics of electoral support (e.g., Fisher et al., 'From Polls to Votes to Seats'; Jackman, 'Pooling the Polls').

[12] 'Sunday question' – elections are traditionally held on Sunday in Germany.

[13] Fisher et al., 'From Polls to Votes to Seats'; Wlezien and Erikson, 'Timeline of Presidential Election Campaigns'.

[14] R.Tourangeau, R.M. Groves and C.D. Redline, ' Sensitive Topics and Reluctant Respondents: Demonstrating a Link between Nonresponse Bias and Measurement Error', *Public Opinion Quarterly* 74/3 (2010): pp.413-32.

[15] Also see Fisher et al., 'From Polls to Votes to Seats'. In our empirical application, we lump both marginal parties in previous elections and new parties in the current election into a category 'others' in order to obtain a random effect estimate for those parties.

[16] Jackman, 'Pooling the Polls'; R.R. Lau, An Analysis of the Accuracy of "Trial Heat" Polls During the 1992 Presidential Election, *Public Opinion Quarterly* 58/1 (1994), pp.2-20.

[17] We did not include an institute-level random effect in our model, since, due to the compositional nature of the dependent variable, there is no reason to suspect that certain polling firms simultaneously over- or underestimate the vote shares of all the parties.

[18] See B. Efron and C. Morris, 'Limiting the Risk of Bayes and Empirical Bayes Estimators-Part II: The Empirical Bayes Case', *Journal of the American Statistical Association* 67/337 (1972), pp.130-9; P. Brown and C. Payne, 'Election night forecasting', *Journal of the Royal Statistical Society, Series A* 138 (1975), pp.463-98.

[19] D. Bates, M. Maechler, B. Bolker and S Walker, 'lme4: Linear mixe-effects models using Eigen and S4' (2014), *R package version 1.1-7.*

[20] Gelman and King, ''Why Are American Presidential Election Campaign Polls So Variable'; also see J. Groß, *Die Prognose von Wahlergebnissen: Ansätze und empirische Leistungsfähigkeit (*Wiesbaden: VS Verlag für Sozialwissenschaften, 2014); Plischke and Rattinger, '"Zittrige Wählerhand" oder invalides Messinstrument'.

[21] An alternative approach would be to let some model parameters vary by period, or to specify a heteroscedastic version of Equation (1) in which the idiosyncratic error variance could be modeled as a function of time to election (Brown and Chappell Jr, 'Forecasting Presidential Elections').

[22] This is just a crude approximation of the variance of a multinomial proportion, but as is demonstrated in S. Fitzpatrick and A. Scott, 'Quick Simultaneous Confidence Intervals for Multinomial Proportions', *Journal of the American Statistical Association* 82 (1987): pp.875-8, it yields reasonable estimates of coverage probabilities.

[23] J.G. Kovar, J.N.K. Rao and C.F.J. Wu , 'Bootstrap and Other Methods to Measure Errors in Survey Estimates', *Canadian Journal of Statistics* 16 (1988): pp.25-45.

[24] M. Cantow, M. Fehndrich A. Schneider and W. Zicht, 'Sonntagsfrage Bundestagswahl. Wenn am nächsten Sonntag Bundestagswahl wäre…', available at http://www.wahlrecht.de/umfragen/index.htm (accessed 20 Jul. 2015).

[25] Polls are getting too infrequent to fit our model to time windows further than 23 months away from the election. Therefore, Figure 1 is limited to windows starting from 0 to 23 months to election.

[26] Also see Groß, 'Die Prognose von Wahlergebnissen'.

[27] We also tested how the window length affects the errors by varying it between one and four months. As expected, an increased width smoothens the forecasting error curve while a shorter width results in more pronounced changes. However, our finding that forecasting error is minimized in a time window of 8 to 10 months to election does not change under any of the specifications.

[28] Gelman and King, ''Why Are American Presidential Election Campaign Polls So Variable'.

[29] Kayser and Leininger, 'A Benchmarking Forecast'.

[30] Jérôme et al., 'A Political-Economy Forecast'.

[31] Norpoth and Gschwend, 'Chancellor Model Picks Merkel'.

[32] M.S. Lewis-Beck, 'Election Forecasting: Principles and Practice', *The British Journal of Politics & International Relations* 7/2 (2005), pp.145-64.

[33] The FDP failed to overcome the 5% threshold and thus, for the first time since 1949, did not make it into the Bundestag.

[34] For a comparative assessment of alternative forecasting methods, see Kayser and Leininger, 'A Benchmarking Forecast' and A. Graefe, 'German Election Forecasting: Comparing and Combining Methods for 2013', *German Politics* 24/2 (2015), pp.195-204.

[35] For instance, the scenario of UK general elections could provide a promising case for our model (neglecting the fact that forecasted vote shares alone are not fully informative about the distribution of Westminster seats), as polling data are available since the 1970s (see http://ukpollingreport.co.uk/; accessed 20 Jul. 2015) and, eyeballing the time series, overshooting patterns can be observed as well.