Forecasting Elections in Multi-party elections: A backwards random-walk approach*

Lukas Stoetzer†
University of Zurich

Simon Munzert‡
Humboldt-University of Berlin

Thomas Gschwend§
University of Mannheim

Marcel Neunhoeffer¶
University of Mannheim

Sebastian Sternberg‖
University of Mannheim

ABSTRACT

We present results of an ex-ante forecast of party-specific vote shares at the German Federal Election 2017. To that end, we combine data from published trial heat polls with structural information. The model takes care of the multi-party nature of the setting and allows making statements about the probability of certain events, such as the plurality of votes for a party or the majority for coalition options in parliament. The forecasts of our model are continuously being updated on the platform zweitstimme.org. The value of our approach goes beyond the realms of academia: We equip journalists, political pundits, and ordinary citizens with information that can help make sense of the parties' latent support and ultimately make voting decisions better informed.

*NOTE: We thank Marc Debus, Helmut Norpoth and Kai-Uwe Schnapp for their helpful comments. Furthermore, we thank Jochen Groß, Marcel Noack, Gertrud Petrig (Institut für Demoskopie Allensbach) and Rainer Schnell for making available historic polling data. All data and computer code necessary to replicate the results in this analysis will be made publicly available on our web pages on publication.

†Corresponding Author: PostDoc, Department of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zürich, Switzerland (lukas.stoetzer@uzh.ch).

‡Postdoctoral research and teaching fellow, Faculty of Humanities and Social Sciences, Humboldt-University of Berlin, Universitätsstr. 3b, D-10099, Berlin, Germany (simon.munzert@hu-berlin.de).

§Professor, Department of Political Science, University of Mannheim, A 5, 6, D-68131, Mannheim, Germany (gschwend@uni-mannheim.de).

¶PhD Candidate, Department of Political Science, University of Mannheim, A 5, 6, D-68131, Mannheim, Germany (mneunhoe@mail.uni-mannheim.de).

‖PhD Candidate, Department of Political Science, University of Mannheim, A 5, 6, D-68131, Mannheim, Germany (ssternbe@mail.uni-mannheim.de).

1 Introduction

We present a dynamic Bayesian forecasting model to predict party support in a multiparty election contest. Our strategy is to systematically combine information from a new fundamentals-based forecasting model with results from national-level pre-election polls that are publicly released during the campaign. Multiparty races are fundamentally different from races, such as US presidential elections, where one merely has to predict who will win a two-party contest — although, granted, this is hard already. For instance, the concept to talk about who "wins" an election is inherently fuzzy because the largest party might not necessarily end-up winning the prize of holding the prime minister. In fact, the largest party might not even be part of new government. Governments in multiparty systems normally consist of several parties that agree to form a coalition in order to assure them a majority of seats in parliament. Single-party governments rarely exist. Thus in multiparty contests the focus is on different quantities of interest.

Our primary target is to forcast vote shares of seven parties for the upcoming 2017 German Federal election on September 24. Given the nature of a multiparty contest, we also predict which coalitions of parties might secure a majority of seats in the Bundestag. Coalition formation depends on which parties are represented in parliament. In order to communicate the corresponding forecast uncertainty, we present the probability that a particular coalition will get a majority of seats in parliament. We thereby also account for the fact that votes for some parties are wasted in the sense that these parties do not overcome a threshold to gain a seat[1].

The general modeling strategy goes back to Linzer (2013) and Strauss (2007). Our innovation is that we, first, allow for more than two parties or candidates, second, account for the compositional nature of party support in multiparty systems (Aitchison, 1986), and, third, that we communicate our estimation uncertainty also through $\frac{5}{6}(\approx 83\%)$ Bayesian Credible Intervals, equating a failure of the model with the intuitive probability of rolling a 6 with a fair six-sided dice. We develop a new fundamentals-based model and apply it to information that is available prior to the election campaign in Germany 2017. We also leverage pre-election polls from the 2017 Federal election campaign in order to forecast the outcome. Our integrated modeling strategy provides a formula for how to weight current party support levels as measured in polls with historical trends using a fundamental-based forecast. Thus, our forecast is always

[1] Note that other models are necessary in order to predict which coalition actually forms (Martin and Stevenson, 2001, e.g.,) given a distribution of seats.

a compromise between the current party support level estimated through a dynamic Bayesian measurement model and the predictions of our fundamental-based model. We allow the model to borrow strength across time through the use of random-walk priors because polling data is sparse and unreliable — particularly early in the campaign. Moreover, the model also filters away day-to-day variation in the polls due to sampling error and house effects.

To summarize, our model allows us to improve on typical fundamentals-based forecasting strategies by dynamically updating parties' current support level as polls get published until election day. Such a model that combines both, a fundamentals-based model and pre-election polls does improve both, accuracy and precision of our initial fundamentals-based forecast.

2 Forecasting German Elections

We are by no means the first trying to predict the outcome of German elections. Previous attempts to forecast German elections leverage, similar to our fundamentals-based model, historical information that is readily available (Gschwend and Norpoth, 2000, 2001; Kayser and Leininger, 2016, 2017; Norpoth and Gschwend, 2003). Other strategies have been to rely on pre-election polls and combine this information (Selb and Munzert, 2016; Walther, 2015). Fundamentals-based models are quite elegant because they use only a few data points but are able to predict some relevant forecasting target, most often the combined vote share of incumbent coalition parties, fairly well. In addition, they use to provide substantial lead time — in our case, 200 days. Conversely, the (seeming) advantage of dynamic models using pre-election polls is that they are able to pick up campaign dynamics until election day. The disadvantage of exclusively relying on models using pre-election polls is that they essentially require the polls to be correct. We have seen recently that this is not always the case.

Nevertheless, pre-election polling data represent a rich data source for tracking the evolution of voter preferences and party support during an election campaign and can also be leveraged to forecast election outcomes. Not only is the media interested in anticipating election outcomes or communicating campaign dynamics to the public but there is also renewed scholarly interest in explaining under what conditions certain campaign events have an impact on party support. Sampling variability is one source of error that accounts for a large share of the total variation in reported party support during an election campaign. The typical sample sizes of pre-election polls are not big enough to detect small changes in party support. There is a need for pooling this information to filter out the substantive information. Not every observed change is real. Furthermore, there are differences between polling organizations in survey design, question wording, sampling weights, and so forth that contribute to the larger total survey error (Schnell and Noack, 2014). For instance, the realized raw polling results might get weighted by procedures that does not permit to replicate the process when scholars work with the raw data. Moreover, different survey firms might treat undecided voters and, in particular, likely voters differently. While we do not distinguish statistical design effects from other potentially systematic and undocumented differences across survey firms, we will nevertheless account for a combination of them as so-called *House effects*. They might cause polling firms to produce estimates of party support that are potentially systematically favorable to particular parties and unfavorable to others.

3 A dynamic Bayesian forecasting model for multiparty elections

In this section, we present a modeling strategy to forecast vote shares in multiparty elections. We develop a dynamic Bayesian forecasting model that consists of two components. First, we develop a fundamental-based component that is able to provide a forecast for each party based on regularities we know about how voters behave and how they determine the outcome of an election. Second, we develop a dynamic Bayesian measurement model that estimates the current level of party support based on published information about voting intention in pre-election polls. The model is able to update the current level of support for every party if new polls get published. By using a backwards random walk approach (Linzer, 2013; Strauss, 2007) we are able to leverage both components and provide a forecast that is a compromise between both, information from polls and the fundamentals-based prediction. The exposition of the model is tailored to the application to the 2017 German Federal Election, but can generalized to other settings, too.

3.1 A fundamentals-based forecasting model for multiparty elections

Fundamentals-based models have clear advantages when predicting the outcome of elections. First, as polls tend to exhibit relatively large forecasting variance when election day is still far away, fundamentals models tend to be more reliable early in the campaign cycle. Second, they help put current election in a historical context, which is useful to build expectations about how special a particular election and its campaign really is. In contrast to many election observers who look merely how the current campaign plays out, fundamentals-based models allow us to learn from regularities across many elections and leverage them to forecast and explain the outcome of an upcoming election. In our case, we systematically leverage information on all 18 post-war federal elections in Germany since 1949.

We consider vote shares y_{pe} of party p $(= 1, \ldots, P)$ at election e $(= 1949, 1953, \ldots, 2013)$. The number of parties varies across elections. Until the 1976 election, we model shares of *CDU/CSU, SPD, FDP* and "*others*" (as combined share of all remaining competing parties). From 1980 on, we also model the vote shares of the *Greens* (later *Bündnis 90/Die Grünen*) and since 1990 also the vote share of the *Left Party* (originally *PDS*). Finally, the right-wing *AfD* is considered from 2013 onward. Our goal is to use the information from previous elections ($e \leq 2013$) to predict y_{p2017} for all seven parties. Therefore, we assume that the data-generating process of y_{pe} is distributed as follows:

$$y_{pe} \sim N(\mu_{pe}, \sigma^2), \tag{1}$$

while the systematic component of the model is a linear function of covariates

$$\mu_{pe} = \beta_e^0 + \sum_k \beta_e^k x_{pe}^k. \tag{2}$$

There is no consensus in the literature how to select predictors for fundamentals-based models and which predictors are the most relevant ones in the multiparty context of German elections. Let us briefly outline our strategy to overcome this "specification uncertainty" (Lauderdale and Linzer, 2015, p.966), and to address the bias-variance trade-off (Hastie, Tibshirani and Friedman, 2009) inherent in any such situation: First, we define a universe of conceivable covariates[2] and generate different sets of predictors using all possible combinations ($2^{10} - 1 = 1023$ in total). Next, we regressed y_{pe} on the respective set of predictors and inspected each R^2. Two observations stick out. On the one hand, the relative increase in R^2 for models with more than three predictors drops considerably. On the other hand, three specific predictors come-up consistently in the best performing models when compared across other models with the same

[2] Among other we have previous vote share, average vote share across the previous three elections (Norpoth and Gschwend, 2003, 2010), average vote share in polls 230-200 days before an election (Selb and Munzert, 2016), vote share in state elections during the last legislative term, indicator variables for certain types of parties (party of chancellor, large party, in government, in parliament) as well as context variables that vary only across elections (unemployment rate and rate-of-change in unemployment in the year prior to an election).

number of predictors. First, the vote share in the previous election (with '0' for new parties), second, the average party vote shares as published in all available polls $230 - 200$ days before the election, and third, a dummy variable to indicate the party of the chancellor. All three variables can also be motivated theoretically and allow, therefore, to test established theories about elections and voting behavior. To sum up, we have yet an empirical justification of the "Rule of Three" (Achen, 2002, pp.445-8). The following paragraphs give a theoretical justification for the selection of those three fundamentals in the German case.

Elections are not held in a political vacuum. It is well known that voters develop long-term stable attachments to political parties (Campbell et al., 1960). The distribution of such attachments in the aggregate allows us to form expectations about the outcome of a given election under normal circumstances (Converse, 1966). We operationalize such a normal-vote baseline as the party's vote share in the previous election (which we set to '0' if the party competes for the first time)[3]. Panel (a) in Figure 1 shows the relationship between previous and current party vote shares across elections. While this predictor clearly helps separate small from large parties and also help explain variation within these clusters, we can also see that our first predictor is not enough to explain the performance of some parties that gained or lost considerably.

Parties get support not only from their partisan base, but also from so-called undecided voters or even partisans of other parties. These voters might be motivated to support a different party by their preference for particular issues and/or candidates. In order to account for such short-term effects on a party's vote share, we leverage published information from pre-election polls about voters' intentions to vote (Groß, 2010; Schnell and Noack, 2014; Selb and Munzert, 2016). We operationalize the level of support for each party before the start of each election campaign as the average vote share in polls 230-200 days before an election.[4] Panel (b) in Figure 1 shows that our second predictor performs already quite well in predicting the actual vote shares on election day.

Our third predictor accounts for the fact that for every performance evaluation of the government, it is important which party holds the chancellorship. Credit and blame regarding the performance of the incumbent government most heavily registers with the support for the chancellor's party. The chancellor is the most visible politician in government. We therefore

[3] Kayser and Leininger (2017) use the same operationalization as predictor for their model while the Gschwend and Norpoth's "chancellor model" operationalize a party's normal-vote baseline as the average vote in the last three *Bundestag* elections (Norpoth and Gschwend, 2003, 2010, 2013).

[4] Selb and Munzert (2016) find generally a better forecasting performance with poll results before the campaign introduces noise to the system.

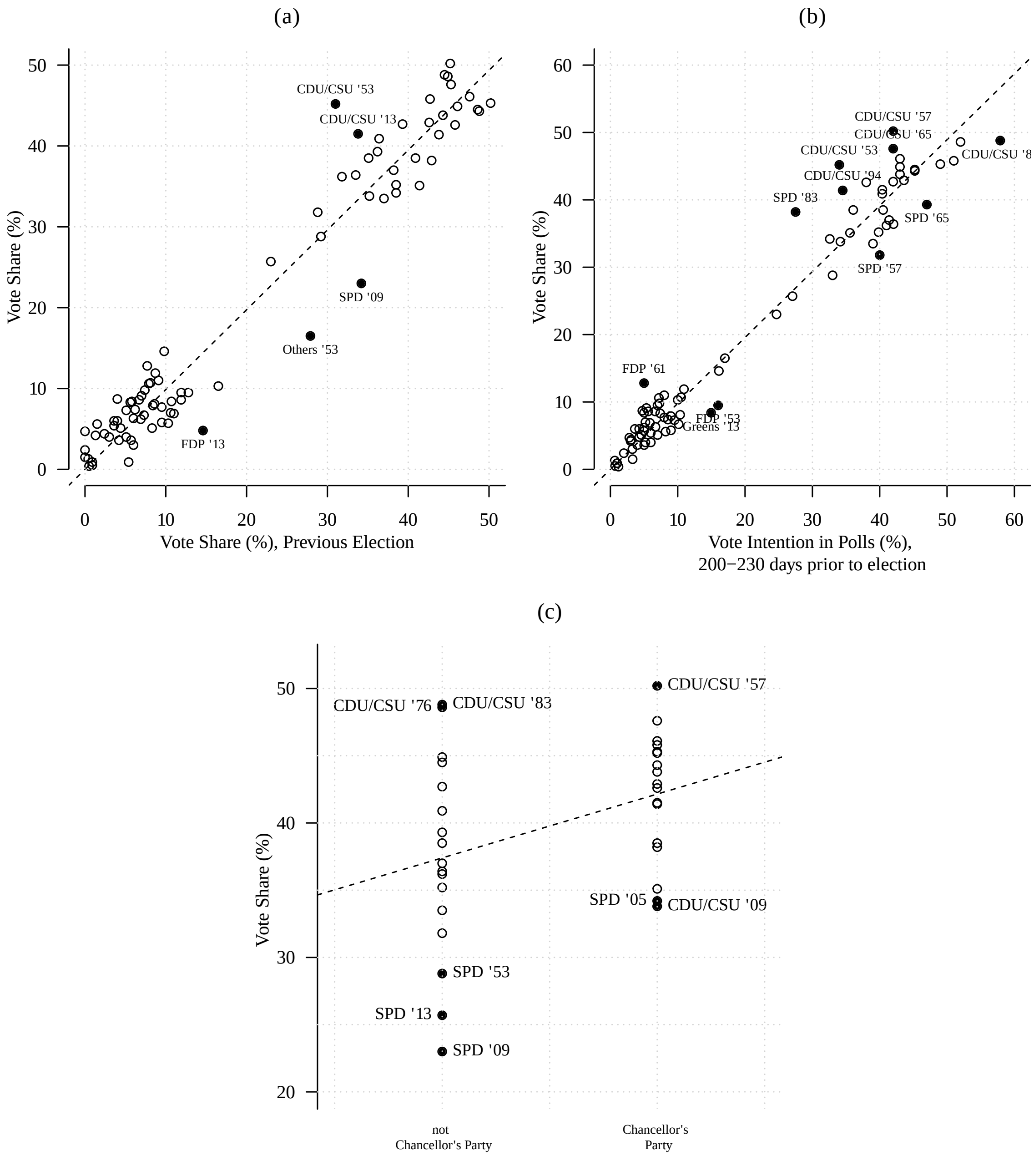

Figure 1: Relationship between predictors and vote share, 1953-2013.

construct an indicator variable scoring '1' for the party that holds the chancellorship[5]. Panel (c) in Figure 1 makes transparent that the party of the incumbent chancellor has on average a larger vote share than the respective other large party that does not hold the chancellorship (the small parties are not shown).

Our fundamentals-based forecasting model is based on a time frame of about 70 years. Rather than running an ordinary least squares regression assuming the same data-generating process across all those elections (as other fundamentals-based models usually do), we relax this assumption. It is well known that the German electorate is increasingly less partisan than it used to be (Arzheimer, 2006). If this dynamic process is supported by our data, we should expect the effect of long-term factors to decrease over time, and conversely, the effect of short-term factors to increase over time. In order to account for that, we allow the parameters of Equation 2 to vary according to a forward random walk across elections as follows:

$$\beta_e^k \sim N(\tilde{\beta_e^k}, \sigma_k^2), \tag{3}$$

$$\tag{4}$$

while we allow any parameter at election e to be a draw from a normal distribution with a mean that comprises of the sum of the previous parameter and a drift parameter of the random walk, i.e.:

$$\tilde{\beta_e^k} = \beta_{e-1}^k + \gamma_{\text{drift}}^k. \tag{5}$$

Figure 2 depicts the estimated β coefficients for the last 17 elections. The pattern confirms our expectations; while the importance of prior election results decreases over time, the polls get more predictive in foreseeing the final outcome. The Chancellor-party effect varies over time, but does not mirror a clear trend. For the 2017 elections, we extrapolate the observed trends for all coefficients, given the estimates of the drift parameter and the random-walk component.

The forecast of the fundamentals-based component will not be viewed in isolation. Rather, it will serve as an anchor to our dynamic measurement model, or more specifically, as an informative prior for a backwards random walk. The next section demonstrates how we combine

[5]The 1983 election is special case because the party of the chancellor right before the election, the CDU/CSU, was not considered the incumbent that is to blame for the current situation. The SPD just lost the chancellorship a few months earlier through a reshuffling of the government. Similar to the coding strategy of the chancellor model (Gschwend and Norpoth, 2000, 2001) we therefore consider the SPD as incumbent party of the chancellor for the 1983 election.

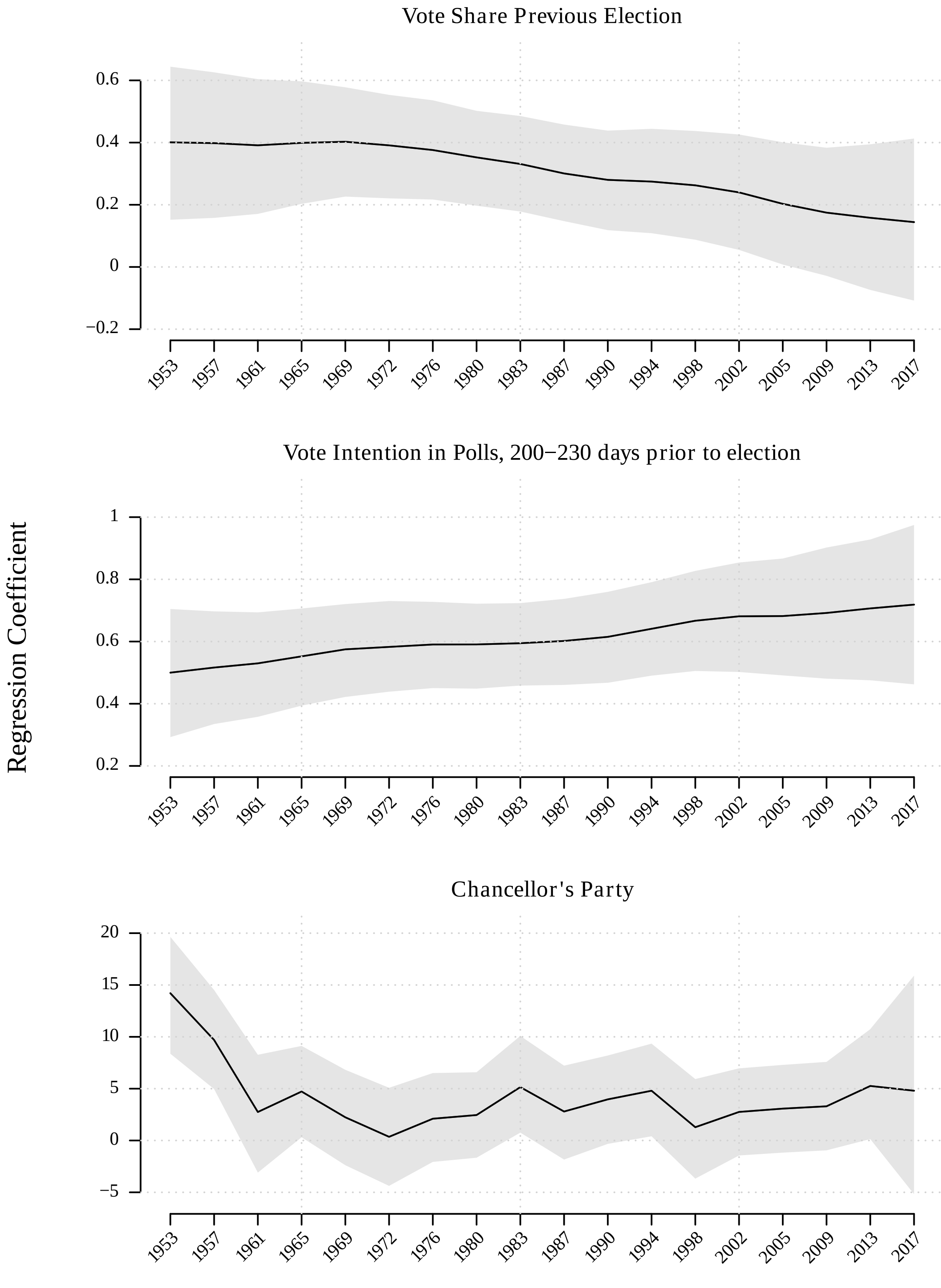

Figure 2: Coefficients for fundamentals-based forecasting model

information from pre-election polls with our fundamentals-based forecast.

3.2 How to combine pre-election polls with fundamentals-based forecasting model

Let y_p denote a forecast of the vote share received by party p $(= 1, \ldots, P)$ based upon our fundamentals-based model generated far in advance of election day with $\sum_p y_p = 1$. We will use y_p to specify a prior for each party's vote share in the upcoming election. The resulting prior is then updated using pre-election vote shares of parties gathered in the last seven months to the election.

More and more of pre-election polls will become available the longer the campaign progresses. Let t $(= 1, \ldots, T)$ represent the days of the campaign, whereby $t = 1$ corresponds to the first day for which we have polling data available and $t = T$ to election day. The model can be fitted any time prior to the election using all the polling information available up to that point. However, earlier polls need not to be included.

The current level of party support for each party is estimated using a combination of poll results and house effects while accounting for the size of those surveys. To that end, let y_{ptc} be the published vote share of party p at time t by polling company c $(= 1, \ldots, C)$. Every poll has a sampling size of N_{ct} assuming random sampling. Using classical reliability theory, we assume that the observed vote share for each party p will be at each point t in time a function of a latent party support vector $\alpha_t^* = (\alpha_{t_1}^*, \ldots, \alpha_{t_P}^*)$, the so-called "true" support of each party among voters, as well as a vector of house effects $\delta_c^* = (\delta_{c_1}^*, \ldots, \delta_{c_P}^*)$ that might systematically bias the published vote shares of the parties and a stochastic component. Those house effects occur because each polling company implements their own survey design using particular question wordings in a particular order, have their respective models to identify likely voters or further weighting mechanisms. These factors can bias each poll such that the reported results are systematically different from the true party support. Moreover, we assume those biases are likely to be present in other polls of that company as well and, thus, do not vary across time for each polling company. In order to identify the house effects, we assume that for each company c the bias across all parties sums to zero, i.e., $0 = \sum_p \delta_{c_p}^*$ as well as the sum of all the biases across companies, i.e. $0 = \sum_c \delta_c^*$.

We conceptualize each published poll result y_{tc} as a P-dimensional random variable that is generated by a multinomial process based on a sample size N_{ct} with an expected value of

$\alpha_t^* + \delta_c^*$. Thus,

$$y_{tc} \sim \text{Multinomial}(\alpha_t^* + \delta_c^*, N_{ct}). \tag{6}$$

All published vote shares for each poll sum up to 100 percent. To account for this, and to map such a vector of proportions into a vector of unbounded, real-valued quantities, we employ a log-ratio transformation (Aitchison, 1986) of the latent party support vector. Specifically, we employ a particular version of the log-ratio transformation, where each entry of the latent party support vector α_t^* at time t is divided by the latent party support for *other parties* α_{tP}^* at this point in time before taking the log, i.e.,

$$\alpha_t = \Big(log(\frac{\alpha_{t1}^*}{\alpha_{tP}^*}), \ldots, log(\frac{\alpha_{t(P-1)}^*}{\alpha_{tP}^*})\Big) = (\alpha_{t1}, \ldots, \alpha_{t(P-1)}) \tag{7}$$

Using a common baseline also reduces the dimension of the resulting vector α_t of log-ratios by 1 (i.e., $1 \leq p \leq P-1$). After transforming and modeling the $P-1$-dimensional vector of log-ratio transformed party support, the obtained results are transformed back and expressed in the meaningful scale of party vote shares.[6]

We conceptualize α_t as a backward random walk starting at election day and moving backwards in time to the start of the campaign, i.e.

$$\alpha_t = \alpha_{t+1} + \omega_t, \quad \omega_t \sim N(0, W) \tag{8}$$

This process assumes that the (log-ratio of the) party support level today depends on the respective level of the following day and an error term because we do not know how party support level evolves over time. That said, it is equally likely to go up or down. The direction in which party support levels have moved does not predict where they will move.

The variance of this evolution process, the so-called *evolution variance* W (West and Harrison, 1997), describes the rate of change between any two consecutive days. It tells how the random walk process evolves. Assuming such a process allows us to compute party support levels for each day even if no poll is released.

[6]The backward transformation is given by:

$$\alpha_t^* = \Big(\frac{exp(\alpha_{t1})}{1 + \sum_{p=1}^{P-1} exp(\alpha_{tp})}, \ldots, \frac{1}{1 + \sum_{p=1}^{P-1} exp(\alpha_{tp})}\Big) = (\alpha_{tp}^*, \ldots, \alpha_{tP}^*)$$

We assume W to be constant over time and independent across parties, i.e.

$$W = diag(\sigma^2) \tag{9}$$

where $\sigma^2 = (\sigma_1^2, \ldots, \sigma_{(P-1)}^2)$ is a vector of party-specific evolution variances.

The advantage of conceptualizing such a random-walk process backwards (Linzer, 2013; Strauss, 2007) rather than forwards (Jackman, 2005) is that it allows to integrate party-level forecasts from fundamentals-based models as specific priors on Election Day (T). Two aspects have to be considered: First, the forecasts should also lay on the unit interval. Second, as the latent support now is defined on the log-ratio scale, so should the forecasts from the fundamentals-model. To account for the two, we transform the forecast from the fundamentals-based model. For the constraint, the expectation and variance of the forecast y_p are redefined in terms of the shape parameters a_p and b_p of a beta distribution. [7]

$$\alpha_{tp}^* \sim Beta(a_p, b_p) \tag{10}$$

In order to map those expectations to log-ratio shares, we further transform the priors. To complete the model, we work with uninformative priors for the party specific evolution variances, the initial state α_1 and the house effects.

When evaluating our combined model based on past elections, we found that the estimated uncertainty of our predictions was too small. The actual election result of a party was included less often in the respective credibility intervals of the model forecasts than to be expected. This implies that the model is more certain about its forecasts than it should be given its performance predicting previous elections. Polls are fairly certain a few days before the election despite being off. Modeling party support in the aggregate based on vote intention published in pre-election polls is apparently not the same thing as actual vote choice for those parties on election day. Even the best surveys are not representative of the population eligible to vote. Supporters of some parties might be more likely to respond to surveys. Moreover, there might be notorious late-deciders or undecided voters that end-up supporting different parties disproportionally. Finally, rather deterministically the design of many surveys is different from what uncertainty based on simple random samples assumes. Survey methodologists discuss these different forms of biases

[7] We follow Jackman (2005, p.55). Given that the forecats are normally distributed with $y_p \sim N(\mu_{f_p}, \sigma_{f_p}^2)$ we transform those values to the beta shape paramters according to: $a_p = \left(\frac{1-\mu_{f_p}}{\sigma_{f_p}^2} - \frac{1}{\mu_{f_p}}\right)\mu_{f_p}^2$ and $b_p = a_p(\frac{1}{\mu_{f_p}-1})$.

that plague such polls in the context of “total survey error” (Schnell and Noack, 2014). In order to account for such biases, we use a strategy similar to the one employed by Hanretty, Lauderdale and Vivyan (2016) in the context of UK elections and add an additional error term for the forecast of each party on election day based on the how much polls were off from the actual election results. This leaves us with the question: Which is an appropriate variance? Historically, the polls were off on average (based on the root mean squared error of all polls three days before the election) by .18 on the log-ratio scale. Thus, for all $1 \leq p \leq P-1$, we add an error term to account for the final party support vector on the log-ratio scale $\widetilde{\alpha_{Tp}} = \alpha_{Tp} + s_p$, with $s_p \sim N(0, .18)$. By re-transforming $\widetilde{\alpha_{Tp}}$, we get our final forecast for each party.

4 Data, Estimation and Evaluation

To calibrate the fundamentals-based model, we leverage data on all 18 federal elections in Germany since 1949.[8] To identify vote intention 230-200 days ahead of the respective election, we rely on data initially collected by Groß (2010), later appended and made available by Schnell and Noack (2014).[9] For all polls published since 2009, we use data provided on the online platform wahlrecht.de. We exclude polls from firms that only publish rarely. This leaves us with polls from the following firms: Institute für Demoskopie Allensbach, Forschungsgruppe Wahlen, forsa, Emnid, GMS, Infratest dimap and INSA. These are also the polls we use to estimate our final dynamic Bayesian forecasting model for the elections from 2002 to 2017.

To estimate the model parameters and to obtain the predicted party vote shares from the fundamentals-based and the dynamic Bayesian forecasting model, we simulate their posterior distribution via a Markov-Chain-Monte-Carlo algorithm implemented in JAGS (Plummer, 2016)[10]. We use three MCMC chains with 100,000 iterations each[11]. Furthermore, we use a burn-in of 100,000 iterations, only saving each 100th iteration[12].

Figure 3 shows the relationship between observed and predicted vote shares of parties for past Federal elections. The vote shares were predicted out-of-sample. On average across all parties and past elections, the RMSE (root mean squared error) is 3 percentage points. In simpler

[8]The result of the 1949 election is used as an indicator of long-term party identification, but is not part of the training set.

[9]Furthermore, we filled gaps in the time series with data made available by the polling company Allensbach.

[10]We will provide all data and code necessary to replicate our analysis at our Github account https://github.com/zweitstimme/btw-2017.

[11]We use two MCMC chains for the models not used for the final forecast.

[12]The convergence of the Gibbs sampler is checked using the Gelman-Rubin diagnostic and visual inspection (Gelman and Rubin, 1992; Brooks and Gelman, 1998).

terms: the predicted vote shares of the fundamentals-based model are, on average, 3 percentage points off from the actual vote shares[13]. The right graph of Figure 3 shows the change in party vote shares. It is evident that the fundamentals-based model is quite off for some observations, in particular with respect to the large parties. For the winner of the Federal elections 2013, the CDU/CSU, the fundamentals-based model predicted only an increase in the vote share of 1.7%, whereas the actual increase was 7.7%[14]. Here the pre-election polls come into play. If we have a closer look at the observations which were poorly predicted by the fundamentals-based model, we notice some late ups and downs in party support. This is something the fundamentals-based model cannot capture. To leverage the additional information about trends in party support, we combine the pre-election polls with the fundamentals-based model in our final dynamic Bayesian forecasting model.

We use the polling results of the SPD during the 2013 Federal election campaign to illustrate how the anchoring process in the backwards random walk works. Figure 4 shows the results of the dynamic forecasting model for the SPD vote-share over time, starting 148 days before the election. Using this example we want to highlight two important features of the model. First of all, we see that the uncertainty about the SPD vote share considerably declines over time. 148 days before the election, the 95% credible interval reaches from about 22% to 29%, whereas it is only between 26% and 28% eight days before the election. Also, the final vote-share of the SPD in the election 2013 is included in all intervals.

Second, Figure 4 illustrates how the model's weight on the fundamental's based forecast and the pre-electoral party support as measured in polls changes over time. At the beginning of the electoral campaign, the model puts more weight on forecast of the fundamentals-based model, whereas it puts more and more "trust" into the polling trends when elections come temporally closer. Take the two graphs in the upper part of the figure, 148 and 64 days before the election. Here, the predicted vote share for the SPD slowly approaches towards the horizontal dashed line indicating the fundamentals-based forecast. In contrast, eight days and one day before the election the model diverges from the fundamental-based forecast and rather approximates the tendency of the public support expressed in the polls. This is reasonable because the fundamentals-based model initially provides much information about the final election outcome,

[13] The MAE (mean absolute error), which does not penalize outliers as much as the RMSE, is 2.2%.

[14] Noteworthy are also the outliers for the CDU/CSU and other parties at the Federal election 1953. One possible reason for this could be the consolidation of the party system in the early days of the BRD, and a related particularly strong vote-share increase for the CDU/CSU.

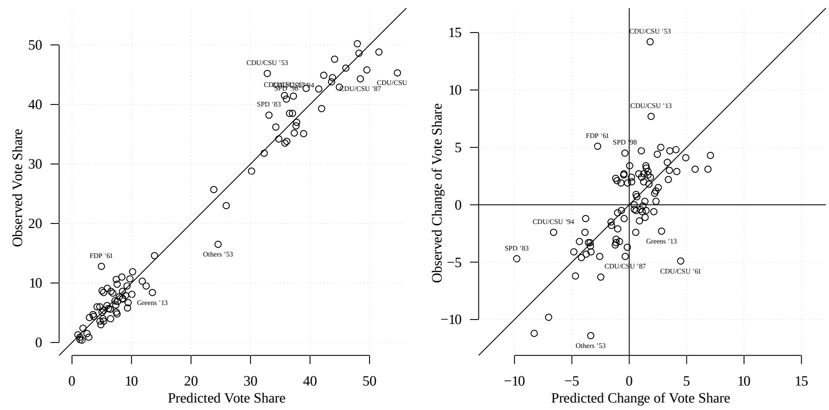

Figure 3: Out-of-sample prediction of vote shares for the fundamentals-based model for Federal elections, 1953-2013.

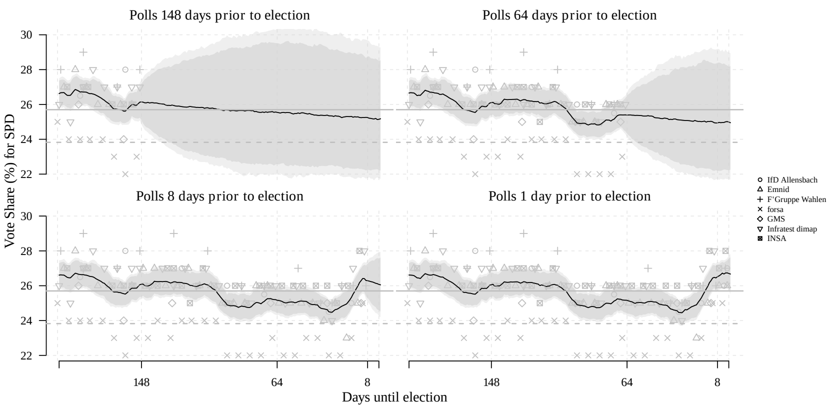

Figure 4: SPD vote share 2013 prediction based on the dynamic Bayesian forecasting model. The symbols represent the party supported reported in the respective polls. The solid line depicts the median latent SPD party support of the posterior distribution; the shadowed area depicts the 90% and 95% credible intervals. The observed 2013 SPD vote share is indicated by the solid horizontal line (25.7%), and the forecast of the fundamentals-based model is marked by the dashed horizontal line (23.8%).

whereas polls become more accurate over time and are thus more and more emphasized [15].

The dynamic feature of our model considerably improves its predictive performance for the last four elections. Table 1 compares the RMSE for the fundamentals-based model with our dynamic Bayesian forecasting model for different points in time of the campaign. The average error of the fundamentals-based model for the elections 2002–2013 is relatively small with only 2.54. Our dynamic Bayesian forecasting model provides a similar accuracy 148 to 36 days before the election, but strongly improves during the last eight days to, on average, 1.83. One day before the election the RMSE is at 1.98 again. This pattern holds for the other elections: the predictive performance of both the fundamentals-based and the dynamic Bayesian forecasting model is about the same until 36 days to the election, but then substantially increases for the dynamic model. The only exception here is the case of the CDU in the election 2005, where the polls systematically overestimated the party's support. A good example for the strength of our dynamic approach is the 2013 Federal election. The forecast of the fundamentals-based model was quite off with a RMSE of 3.47, however its misleading predictions could be bolstered by the dynamic poll component of the dynamic Bayesian forecasting model (with a RMSE of 1.33 and 1.69 one and eight days before the election, respectively).

Table 1: The RMSE of the out-of-sample predictions by data basis and election year. Pure fundamentals-based model and dynamic baysian forecasting model arranged by time to election.

Model	RMSE				
	2002-13	2002	2005	2009	2013
Pure Fundamentals-Based	2.54	1.87	2.32	1.91	3.47
Dynamic. 1 day prior to election	1.98	1.23	3.24	1.38	1.33
Dynamic. 8 days prior to election	1.83	1.08	2.86	1.06	1.69
Dynamic. 36 days prior to election	2.27	1.96	2.80	1.64	2.65
Dynamic. 64 days prior to election	2.22	1.56	2.81	1.48	2.67
Dynamic. 92 days prior to election	2.54	2.06	2.92	1.78	3.14
Dynamic. 120 days prior to election	2.80	2.60	3.27	2.18	3.02
Dynamic. 148 days prior to election	2.59	2.18	3.03	1.88	3.05

Obviously, an accurate electoral prediction is not only a function of the average expected deviation from the observed result, but also a correct uncertainty estimation. The 95% credible intervals provide a coverage rate of 94% for the observed results of the last four elections. In other words, our credible intervals contain the true values about as often as one would statistically expect.[16]

With Figure 5, we illustrate again how the dynamic poll component in our final model

[15]The effect of this trade-off is especially strong for smaller parties, because the fundamentals-based model yields to a more accurate prediction of them. More examples for the elections 2002–2012 can be found in the

improves the predictions over time. For this we use the last Federal election in 2013. Two advantages of our model become evident. First, the forecasts improve over time. The centre of the credible intervals is — at least for the majority of parties and different points in time — between the true results and the predictions of fundamentals-based model. This implies a reduction of the prediction error approaching the election day. For instance, 36 days before the election day the prediction for the Greens is close to the prediction of the fundamentals-based model, which quite overestimated their final vote share. However, the dynamic component of the model learned from the drastically decreasing popularity ratings of the Greens and pulled the forecasts away from the fundamentals-based forecast towards the true outcome. Second, the patterns reveal that the forecasts become more accurate over time, especially for smaller parties. Take for instance the case of the FDP 2013: 148 days before the election, everything between reaching a vote-share of 10% and not clearing 5% seemed possible. 36 days prior to the election our model predicted a vote share around 5%, close to the actual result of 4.8%. For this case, it is notable that the expected vote share remains relatively constant over time, but that the precision of our forecasts improves. Overall, we expect a similar pattern for our forecasts of the Federal election 2017.

5 Forecasting the German Federal Election 2017

In our final step, we use the dynamic Bayesian forecasting model to forecast the vote shares of seven parties of the upcoming 2017 German Federal election on September 24. Furthermore, we also predict which coalition of parties might secure a majority of seats in parliament, thereby presenting probabilities that a particular coalition will have enough seats to form a government. In our ex-ante prediction, we only use data which was available until June 16, 2017, 101 days before the election. To obtain the vote intention variable of the fundamentals-based model, we use 18 different polls which were published between February 6 and March 8, 2017. They thus contain the time period after the announcement of Martin Schulz's candidacy for chancellor, an important component of this electoral campaign's dynamic. We restrict our forecasts of party vote shares to these having at least a realistic chance to passes the five percent hurdle according to polls. These parties are the CDU/CSU (we forecast both parties' vote shares together), the

online archive of this article.

[16]The coverage rates at different points in time are: one day before the election 95%, eight days before the election 100%, 36 days before the election 95%, 64 days before the election 95%, 92 days before the election 95%, 120 days before the election 90%, and 148 days before the election 90%.

SPD, the Left Party, the Greens, the FDP and the AfD. The fundamentals-based model forecasts the following party vote shares, which we then use as prior for the dynamic Bayesian forecasting model: CDU/CSU 35.0% (SD = 4.4), SPD 27.5% (3.0), Left Party 8.4% (2.2), Greens 7.9% (2.1), FDP 6.4% (2.2), AfD 9.1% (2.3).

Figure 6 shows the forecast developments based on newly published polls. To get an intuition for the dynamic nature of our model, we present the forecasts beginning as early as one year before the election. The development clearly depicts the so-called "Schulz effect", which was an extremely positive shock for the SPD time series resulting in a negative shock in the time series of the other parties. In the meantime, this effect has considerably declined: the actual estimated latent CDU/CSU party support is even higher than before the announcement of Martin Schulz's candidacy for chancellor. The time-series further depitcs that the fundamentals-based model's forecast still has a substantial impact on the final forecast for the election day: all curves currently shrink towards the fundamentals-based forecast. This means we expect poorer results for the CDU/CSU and the FDP than the current pre-election polls suggest. However, we expect that the closer the election day approaches, our dynamic Bayesian forecasting model will diverge from the fundamentals-based forecast and will put more and more weight on the party-support expressed in polls.

Figure 7 provides our final forecasts of seven party vote shares for the upcoming 2017 German Federal election, 101 days before election day, along with the respective $\frac{5}{6}(\approx 83\%)$ credible intervals. We chose the interval as it has an intuitive equivalent in the non-political science world: Rolling a 6 (or any other number) with a six-sided dice has a probability of roughly 17%. This is equivalent to the probability of our intervals to not cover the true result. Accordingly, the CDU/CSU will reach Union 36.8% [29.7%; 44.4%], the SPD 26.5% [20.5%; 32.8%], the Left Party 8.6% [6.3%; 11.3%], the Greens 7.3% [5.1%; 9.8%], the FDP 8.4% [5.9%; 11.3%], the AfD 8% [5.6%; 10.7%], and Others 4.3% [3.2%; 5.5%].[17] It is evident that the uncertainty of our final forecasts is still substantial[18] We want to emphasize again that due to the dynamic nature of our model, these numbers contentiously change over time. We publish our updated forecasts on our webpage *zweitstimme.org* every time a new pre-election poll is published.

[17]The reported numbers represent the median, not the mean prediction of the posterior distribution. The sum of these numbers is thus slightly different from 100%.

[18]The uncertainty intervals presented in Figure 7. are wider than the intervals for the election day presented in Figure 6. This is because the latter does not take the shock correction into account.

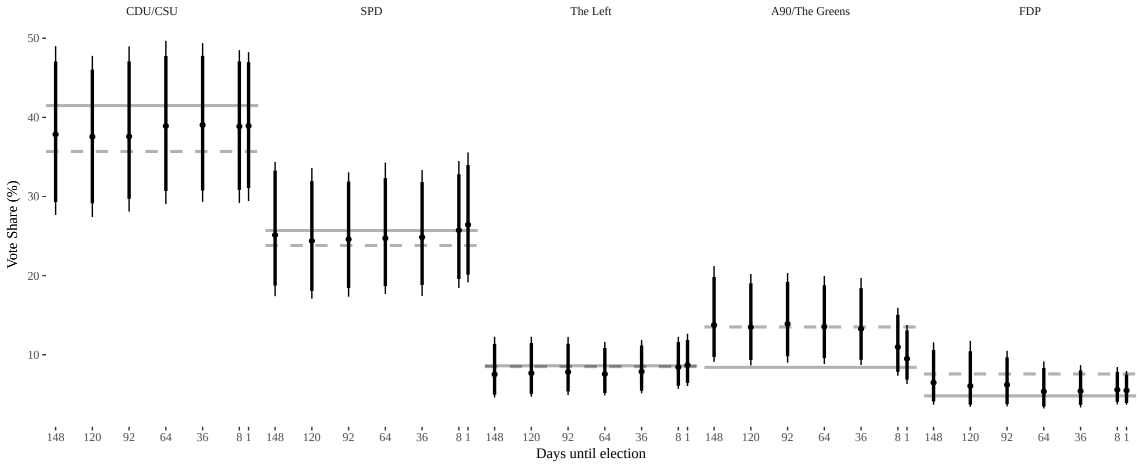

Figure 5: Development of the dynamic Bayesian forecasting model's vote share predictions over time for the Federal election 2013, starting 148 days until the final day before the election. The points show the median prediction; the thick and thin lines depict the 90% and 95% credible interval, respectively. Each party's observed vote-share is indicated by the solid horizontal line, and the forecast of the fundamentals-based model is marked by the dashed horizontal line.

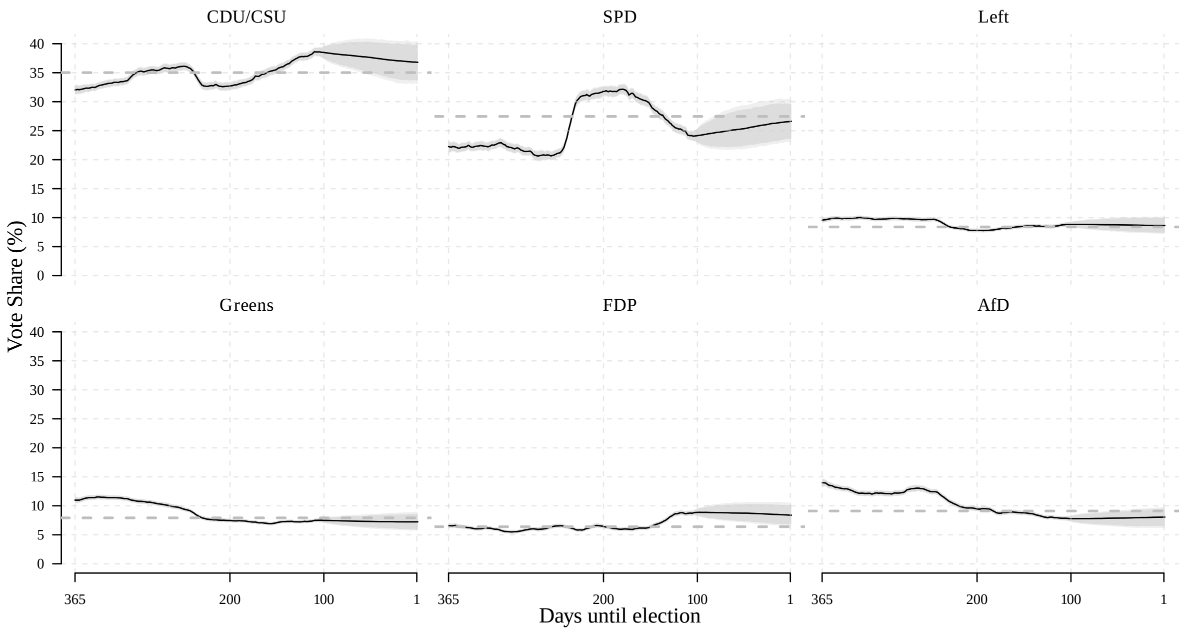

Figure 6: Dynamic Bayesian forecasting model predictions of each party's vote share for the Federal election 2017, starting 365 days before the election including polls up until 101 days prior to the election. The solid line depicts each party's median latent party support; the shadowed area depicts the $\frac{5}{6}$% and 95% credible intervals. The fundamentals-based model's forecast is indicated by the dashed horizontal line.

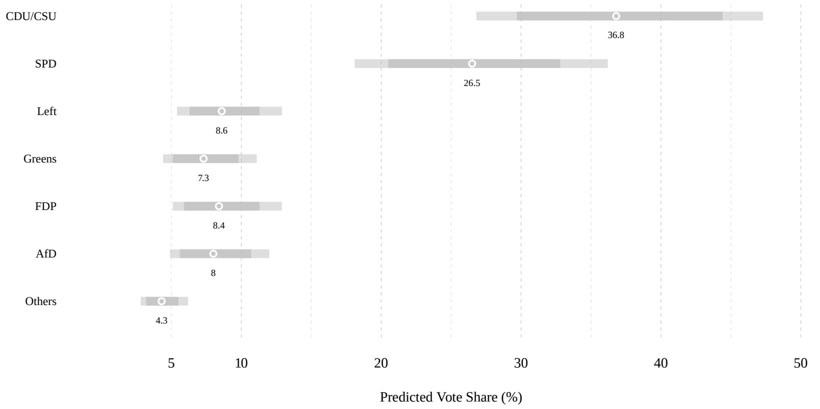

Figure 7: Forecast of the 2017 election 101 days prior to the election. Point estimates, 95% (light grey), and $\frac{5}{6}(\approx 83\%)$ (dark grey) credible intervals.

A great advantage of Bayesian models is how easy it is to calculate probabilities of other events based on the predicted outcomes. In our application, we use the results of the simulations to calculate the probability that a particular coalition will get a majority of seats in parliament. We also consider the likelihood of parties passing the five percent hurdle. Given the nature of the German multiparty context, these quantities of interests might be more important than just reporting individual party vote shares. We report some of these quantities in Table 2. The probability that the CDU/CSU will be the largest party in the Bundestag is 88%, whereas it is only 12% for the SPD. What is even more important is which coalition would have a majority of seats in the new Bundestag in absolute terms. Looking at our forecasted party vote shares along with our uncertainty estimates, unsurprisingly, it is almost certain that a grand coalition of CDU/CSU and SPD would have the necessary majority of seats. The probability of a so-called "Jamaica coalition" (black/green/yellow) of CDU/CSU, Greens, and FDP is currently at 77%. Yet, our results suggest that a coalition between the CDU/CSU and FDP is not completely unlikely with a probability of 30% [19].

With respect to the five percent hurdle, we do not see any risk for the smaller parties not clearing it. Our model further predicts that the FDP and the AfD have the highest chances of increasing their vote shares as compared to the last election. We further estimate that SPD and the Left Party are likely to improve their last election result in one out of two cases. Furthermore, in six out of seven simulations we see six different parties in the Bundestag. The AfD, the newest party on the German political landscape, is in one out of three simulations the third strongest

[19]Please note that all probabilities only refer to a certain coalition having an arithmetical majority. We consider whether a potential partner would pass the five percent threshold, but we do not take into account strategic considerations in the coalition formation process.

Table 2: Event Probabilities based on the simulated vote shares for the 2017 German Federal Election. The pie charts in the third column visualize the Event Pobabilities.

Event	**Probability (in %)**
Biggest Party	
CDU/CSU	88
SPD	12
Arithmetic Majority for Coalition	
CDU/CSU + SPD	100
CDU/CSU + FDP	30
CDU/CSU + A'90/The Greens	22
CDU/CSU + A'90/The Greens + FDP	77
SPD + A'90/The Greens	0
SPD + A'90/The Greens + Left	13
SPD + A'90/The Greens + FDP	12
Will not pass the 5% threshold	
FDP	2
AfD	3
A'90/The Greens	7
Gain in Vote Share as compared to 2013	
CDU/CSU (>41.5%)	18
SPD (>25.7%)	56
The Left (>8,6%)	48
A'90/The Greens (>8,4%)	24
FDP (>4.8%)	99
AfD (>4.7%)	98
More Events	
Six factions in the Bundestag	87
AfD third strongest faction	24

fraction in the Bundestag.

6 Next Steps

We developed a dynamic Bayesian forecasting model for multiparty elections. For the first time, we implemented a so-called *backwards random-walk approach* that works in this context. This approach allows us to systematically leverage information that is available long before the election campaign gets underway in a separate fundamentals-based model. We integrate the predictions from this fundamentals-based model as priors on election day to counterbalance the information we gain when pooling the polls during the election campaign.

There are at least three conceivable extensions on which we plan to work in future iterations. First, we know that there is often a lot of movement between parties that send certain coalition signals (Gschwend, Stoetzer and Zittlau, 2016; Gschwend, Meffert and Stoetzer, 2017). Consequently, the assumption that the party-specific error components are independent does not hold. We need to account for that by parameterizing the evolution variance of the random walk appropriately. Second, polls are typically less precise in forecasting election results compared to what we would expect based on our current formulation. Therefore, we would like to explicitly model the measurement error that is inherent in even the best models. The consequences of this is that we conceptualize polling results not simply as observations but as estimates which are to some degree uncertain. This additional uncertainty, the measurement error variance, will increase the evolution variance and will therefore make our current forecasts less precise. Third, we would like to conceptualize voter transitions to provide a better micro foundation of the variance of party support within election campaigns.

References

Achen, Christopher H. 2002. "Toward a new Political Methodology: Microfoundations and ART." *Annual Review of Political Science* 5(1):423–450.

Aitchison, John. 1986. "The Statistical Analysis of Compositional Data." *Journal of the Royal Statistical Society* 44(2):139–177.

Arzheimer, Kai. 2006. "'Dead men walking?' Party identification in Germany, 1977–2002." *Electoral Studies* 25(4):791–807.

Brooks, Stephen P B and Andrew G Gelman. 1998. "General methods for monitoring convergence of iterative simulations." *Journal of computational and graphical statistics* 7(4):434–455.

Campbell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Converse, Philip E. 1966. The Concept of Normal Vote. In *Elections and the Political Order*, ed. Angus Campbell, Philip E. Converse, Warren E. Miller and Donald E. Stokes. New York: Wiley pp. 9–39.

Gelman, Andrew and Donald B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 7(4):457–472.

Groß, Jochen. 2010. *Die Prognose von Wahlergebnissen. Ansätze und empirische Leistungsfähigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gschwend, Thomas and Helmut Norpoth. 2000. Soll und Haben: Die deutsche Wählerschaft rechnet mit den Parteien ab. In *50 Jahre Empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten*, ed. Markus Klein, Wolfgang Jagodzinski, Ekkehard Mochmann and Dieter Ohr. Wiesbaden: Westdeutscher Verlag pp. 389–409.

Gschwend, Thomas and Helmut Norpoth. 2001. Wenn am nächsten Sonntag...Ein Prognosemodell für Bundestagswahlen. In *Wahlen und Wähler: Analysen aus Anlass der Bundestagswahl 1998 [Elections and Voters. Analysis of the Federal Election 1998]*, ed. Hans-Dieter Klingemann and Max Kaase. Wiesbaden: Westdeutscher Verlag pp. 473–499.

Gschwend, Thomas, Lukas Stoetzer and Steffen Zittlau. 2016. "What drives rental votes? How coalitions signals facilitate strategic coalition voting." *Electoral Studies* 44:293–306.

Gschwend, Thomas, Michael F. Meffert and Lukas F. Stoetzer. 2017. "Weighting Parties and Coalitions: How Coalition Signals Influence Voting Behavior." *Journal of Politics* 79(2):642–655.

Hanretty, Chris, Ben Lauderdale and Nick Vivyan. 2016. "Combining national and constituency polling for forecasting." *Electoral Studies* 41(March):239–243.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *The Elements of Statistical Learning*.

Jackman, Simon. 2005. "Pooling the polls over an election campaign." *Australian Journal of Political Science* 40(4):499–517.

Kayser, Mark A. and Arndt Leininger. 2017. "A Länder-based Forecast of the 2017 German Bundestag Election." *PS: Political Science & Politics* (July):forthcoming.

Kayser, Mark Andreas and Arndt Leininger. 2016. "A Predictive Test of Voters' Economic Benchmarking: The 2013 German Bundestag Election." *German Politics* 25(1):106–130.

Lauderdale, Benjamin E. and Drew Linzer. 2015. "Under-performing, over-performing, or just performing? The limitations of fundamentals-based presidential election forecasting." *International Journal of Forecasting* 31(3):965–979.

Linzer, Drew A. 2013. "Dynamic Bayesian Forecasting of Presidential Elections in the States." *Journal of the American Statistical Association* 108(501):124–134.

Martin, Lanny W. and Randolph T. Stevenson. 2001. "Government formation in parliamentary democracies." *American Journal of Political Science* 45(1):33–50.

Norpoth, Helmut and Thomas Gschwend. 2003. "Against All Odds? The Red-Green Victory." *German Politics and Society* 21(1):15–34.

Norpoth, Helmut and Thomas Gschwend. 2010. "The chancellor model: Forecasting German elections." *International Journal of Forecasting* 26(1):42–53.

Norpoth, Helmut and Thomas Gschwend. 2013. "Chancellor Model Picks Merkel in 2013 German Election." *PS: Political Science & Politics* 46(03):481–482.

Plummer, Martn. 2016. "Jags - Just Another Gibbs Sampler.".
URL: *http://mcmc-jags.sourceforge.net/*

Schnell, Rainer and Marcel Noack. 2014. "The Accuracy of Pre-Election Polling of German General Elections." *Methods, Data, Analyses* 8(1):5–24.

Selb, Peter and Simon Munzert. 2016. "Forecasting the 2013 German Bundestag Election Using Many Polls and Historical Election Results." *German Politics* 25(1):73–83.

Strauss, Aaron. 2007. "Florida or Ohio? Forecasting presidential state outcomes using reverse random walks.".
URL: *http://www.mindlessphilosopher.net/princeton/strauss_ reverse random walk.pdf*

Walther, Daniel. 2015. "Picking the winner(s): Forecasting elections in multiparty systems." *Electoral Studies* 40:1–13.

West, Mike and Jeff Harrison. 1997. *Bayesian Forecasting and Dynamic Models.* 2nd ed. New York: Springer.