

Forecasting Elections in Multiparty Systems: A Bayesian Approach Combining Polls and Fundamentals

Lukas F. Stoetzer¹, Marcel Neunhoeffer², Thomas Gschwend², Simon Munzert³ and Sebastian Sternberg²

¹ Department of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zürich, Switzerland. Email: lukas.stoetzer@uzh.ch

² Department of Political Science, University of Mannheim, A 5, 6, 68131 Mannheim, Germany. Email: marcel.neunhoeffer@gess.uni-mannheim.de, sebastian.sternberg@gess.uni-mannheim.de

³ Hertie School of Governance, Friedrichstr. 180, 10117 Berlin, Germany. Email: munzert@hertie-school.org

Abstract

We offer a dynamic Bayesian forecasting model for multiparty elections. It combines data from published pre-election public opinion polls with information from fundamentals-based forecasting models. The model takes care of the multiparty nature of the setting and allows making statements about the probability of other quantities of interest, such as the probability of a plurality of votes for a party or the majority for certain coalitions in parliament. We present results from two *ex ante* forecasts of elections that took place in 2017 and are able to show that the model outperforms fundamentals-based forecasting models in terms of accuracy and the calibration of uncertainty. Provided that historical and current polling data are available, the model can be applied to any multiparty setting.

Keywords: forecasting, public opinion polls; multiparty elections, German Federal Elections 2017, New Zealand General Election 2017

The majority of countries worldwide feature an electoral system for their main legislative body that is proportional or mixed in nature.¹ Most of these systems—as well as a considerable number of plurality-based systems, such as India, Canada, or the United Kingdom—generate multiparty politics, i.e., more than two parties being regularly represented in parliament. With rising interest in data-based coverage of campaign dynamics, there is a demand for models that deliver forecasts in such multiparty systems. However, existing dynamic forecasting models that have been developed for the US (Silver 2012; Erikson and Wlezien 2013; Linzer 2013)—a setting that is hallmarked by two-party races with long historical records—do not easily translate to multiparty systems.

In this letter, we present a general dynamic model to forecast party vote shares and other related quantities of interest in multiparty elections. To that end, we suggest to combine data from pre-election public opinion polls with information from fundamentals-based forecasting models. We extend the existing backward random-walk approach that has been proposed to forecast US presidential elections (Strauss 2007; Linzer 2013), to account for the compositional nature of party support in multiparty systems. In addition, we develop a Bayesian approach that combines

Authors' note: Replication materials are available online as a dataverse repository (Neunhoeffer *et al.* 2018) <https://doi.org/10.7910/DVN/MLYNX0>. We thank Marc Debus, Helmut Norpoth, Kai-Uwe Schnapp, and Steffen Zittlau 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 for Germany. We also thank Peter Ellis for making available data for New Zealand through the R package *nzelect*.

¹ According to the Institute for Democracy and Electoral Assistance (IDEA) Electoral System Family Database, 116 countries (53% worldwide) belong in one of these categories; see <https://www.idea.int/data-tools/data/electoral-system-design> (retrieved at June 14, 2018).

Political Analysis (2019)
vol. 27:255–262
DOI: 10.1017/pan.2018.49

Published
8 November 2018

Corresponding author
Lukas F. Stoetzer

Edited by
Jeff Gill

© The Author(s) 2018. Published
by Cambridge University Press
on behalf of the Society for
Political Methodology.

information from fundamentals with polls in a fully integrated model. The model, moreover, allows predicting a number of quantities of interest important to multiparty elections, such as the probability with which coalitions of parties might secure a majority of seats and the likelihood that a party can overcome an electoral threshold. These methodological innovations contribute to an emerging literature on synthetic forecasting models (Lewis-Beck and Dassonneville 2015; Lewis-Beck, Nadeau, and Bélanger 2016).

We present results of real-time step-ahead forecasts of two multiparty elections in September 2017: the German (in the main text) as well as the New Zealand election (in the Online Appendix C). In both instances, our dynamic *ex ante* forecast outperformed the fundamental forecast models.

1 A Dynamic Bayesian Measurement Model for Multiparty Elections

Our modeling strategy to forecast vote shares in multiparty elections comprises two components.² The first is a *fundamentals-based model* that provides a forecast for each (relevant) party's vote share long before the election campaign starts. For this purpose, previous contributions employ party-level predictors based on regularities of elections to forecast the election outcome (see, e.g., Norpoth and Gschwend 2010; Magalhães, Aguiar-Conraria, and Lewis-Beck 2012; Jérôme, Jérôme-Speziari, and Lewis-Beck 2013). Most of those regression-based fundamentals models can be defined by the election results of all parties in past elections \mathbf{V} , a matrix of predictors \mathbf{X} , and a vector of parameters θ , which link the predictors to the election results. Two distributions of those models are of central interest for Bayesian forecasting. First, the posterior density of the parameter estimates that infers from the relationship between predictors and election results. It is proportional to the likelihood derived from the statistical model and the priors for the parameters: $P(\theta|\mathbf{X}, \mathbf{V}) \propto P(\mathbf{X}, \mathbf{V}|\theta)P(\theta)$. Second, the posterior predictive distribution that provides a forecast for the upcoming election results \mathbf{v}_E given the predictors \mathbf{x}_E . This distribution takes "coefficient uncertainty" (Lauderdale and Linzer 2015, p.967) into account by integrating over the posterior distribution of the parameters from the predictive distribution for the upcoming election: $P(\mathbf{v}_E|\mathbf{x}_E, \mathbf{X}, \mathbf{V}) = \int_{\theta} P(\mathbf{v}_E|\theta, \mathbf{x}_E)P(\theta|\mathbf{V}, \mathbf{X})d\theta$. Unlike common implementations of fundamentals models, we integrate our fundamentals model into a dynamic Bayesian measurement model, which we describe below. For brevity, we discuss our application-specific Dirichlet regression model in detail in Online Appendix A.1.

In the following, we focus on the second component, which provides the core contribution of our approach: A *dynamic Bayesian measurement model* that estimates the current level of party support based on pre-election polls published during the election campaign, and that combines it with the forecasts from the fundamentals-based model. To that end, we draw on poll results published by different polling companies.

Let y_{ctp} be the reported vote share of party p ($= 1, \dots, P$) at time t ($= 1, \dots, T$) as published by polling company c ($= 1, \dots, C$). Let t represent the days of the campaign, whereby $t = 1$ corresponds to the first day and $t = T$ to Election Day. Each poll has a sample size of N_{ct} . We conceptualize each published poll result $\mathbf{y}_{ct} = (y_{ct1}, \dots, y_{ctP})$ as a P -dimensional random variable that is generated by a multinomial process, where $\boldsymbol{\pi}_{ct}^* = (\pi_{ct1}^*, \dots, \pi_{ctP}^*)$ is a vector of expected support at day t in company c 's poll:

$$\mathbf{y}_{ct} \sim \text{Multinomial}(\boldsymbol{\pi}_{ct}^*, N_{ct}). \quad (1)$$

The vote shares of each poll sum to 100 percent. To account for this, and to map the proportions into a vector of unbounded, real-valued quantities, we employ a log-ratio transformation (Aitchison 1986). Each entry of the expected support share vectors $\boldsymbol{\pi}_{ct}^*$ at time

² Replication materials are available online as a dataverse repository (Neunhoeffer et al. 2018) <https://doi.org/10.7910/DVN/MLYNX0>.

t for company c is divided by the expected party support for the last party P , π_{ctP}^* , before taking the log:³

$$\boldsymbol{\pi}_{ct} = \text{alr}(\boldsymbol{\pi}_{ct}^*) = \left(\log \left(\frac{\pi_{ct1}^*}{\pi_{ctP}^*} \right), \dots, \log \left(\frac{\pi_{ct(P-1)}^*}{\pi_{ctP}^*} \right) \right) = (\pi_{ct1}, \dots, \pi_{ct(P-1)}). \tag{2}$$

In a next step, using classical reliability theory, we decompose the transformed vector of expected support shares for each party into a latent party support vector $\boldsymbol{\alpha}_t = (\alpha_{t1}, \dots, \alpha_{tP})$, the so-called “true” support of each party among voters, as well as a vector of house effects $\boldsymbol{\delta}_c = (\delta_{c1}, \dots, \delta_{cP})$ that might systematically bias the published vote shares of each company (Jackman 2005), such that $\boldsymbol{\pi}_{ct} = \boldsymbol{\alpha}_t + \boldsymbol{\delta}_c$.⁴

In the context of dynamic forecasting of vote shares, i.e., updating existing forecasts with incoming information, a core quantity of interest is the evolution of party support over time. We model the level of support $\boldsymbol{\alpha}_t$ as a backward random walk,⁵ starting at Election Day and moving backwards in time to the start of the campaign, i.e.,

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t+1} + \boldsymbol{\omega}_t, \boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}). \tag{3}$$

This allows us to estimate party support levels for each day even if no new poll is released. Furthermore, this process assumes that the (log-ratio of the) party support level today depends on the respective level of the following day and a random error term. The variance of this random error, the so-called *evolution variance* \mathbf{W} (West and Harrison 1997), describes the rate of change between any two consecutive days. We constrain \mathbf{W} to be constant over time, but explicitly allow the latent states to covary:

$$\mathbf{W} = \begin{bmatrix} w_1^2 & \dots & w_{1,(P-1)} \\ \vdots & \ddots & \vdots \\ w_{1,(P-1)} & \dots & w_{(P-1)}^2 \end{bmatrix}. \tag{4}$$

The key advantage of deploying the random walk backwards (Strauss 2007; Linzer 2013) rather than forwards (see, e.g., Walther 2015) is that it allows integrating party-level forecasts from fundamentals-based models in the dynamic polls model.⁶ The backward random-walk process in the joint posterior distribution isolates the prior on Election Day. This, in turn, allows us to integrate the forecast from the fundamentals model into the dynamic polls model by setting the latent state of party support on Election Day equal to the log-ratio-transformed posterior predictive distribution from the fundamentals model:

$$P(\boldsymbol{\alpha}_T | \mathbf{x}_E) = \text{alr} \left(\int_{\boldsymbol{\theta}} P(\mathbf{v}_E | \boldsymbol{\theta}, \mathbf{x}_E) P(\boldsymbol{\theta} | \mathbf{V}, \mathbf{X}) d\boldsymbol{\theta} \right). \tag{5}$$

3 The common baseline reduces the dimension of the resulting vector $\boldsymbol{\pi}_{ct}$ 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 on the meaningful scale of party vote shares.

4 In order to identify the house effects, we assume that for each company c , the bias across all parties sums to zero, i.e., $\sum_p \delta_{cp} = 0$. The same is assumed for the sum of all biases across companies, i.e., $\sum_c \delta_c = 0$.

5 Strauss (2007) and Linzer (2013) use the term “reverse random walk.” We prefer “backward” to highlight the feature that the walk starts from a prior set on the Election Day and then moves backwards in time.

6 For forward random-walk processes, the prior is defined on the starting point of the time series. The evolution process between a forward and a backward random walk is mathematically equivalent, which can be seen when rearranging Equation (3) by subtracting $\boldsymbol{\omega}_t$ from the right-hand side. We thank an anonymous reviewer for drawing our attention to this.

With this integration, we can write a joint posterior distribution of both the fundamentals model and the dynamic polls model, which is proportional to the product of the two likelihoods,⁷ the posterior predictive distribution, the backward random walk and priors for the parameters from two model components. Integrating and estimating all parts of the model in a joint specification has a set of advantages. Most importantly, it automatically considers the complete uncertainty from the fundamentals model, which is relevant when aiming at accurately weighting the polls and the fundamentals for a synthetic forecast.

$$\begin{aligned}
 P(\boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{W}, \boldsymbol{\theta} | \mathbf{Y}, \mathbf{V}, \mathbf{X}) \propto & \underbrace{P(\mathbf{Y} | \boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{W})}_{\substack{\text{dynamic polls} \\ \text{likelihood}}} \underbrace{P(\mathbf{V} | \boldsymbol{\theta}, \mathbf{X})}_{\substack{\text{fundamentals} \\ \text{likelihood}}} \underbrace{P(\boldsymbol{\alpha}_T | \mathbf{x}_E)}_{\substack{\text{posterior predictive} \\ \text{distribution}}} \\
 & \prod_{t=1}^{T-1} \underbrace{P(\boldsymbol{\alpha}_t | \boldsymbol{\alpha}_{t+1}, \mathbf{W})}_{\substack{\text{backward random} \\ \text{walk}}} \underbrace{P(\boldsymbol{\theta})}_{\substack{\text{fundamentals} \\ \text{prior}}} \underbrace{P(\mathbf{W})}_{\substack{\text{evolution} \\ \text{variance}}} \underbrace{P(\boldsymbol{\delta})}_{\substack{\text{house} \\ \text{effects}}} \quad (6)
 \end{aligned}$$

The complete specification of the dynamic Bayesian forecasting model further requires priors for the evolution variance and the house effects. For this purpose, we decompose the covariance matrix into a $(P - 1) \times (P - 1)$ correlation matrix $\boldsymbol{\Omega}$ and a diagonal matrix with standard deviations on the main diagonal: $\mathbf{W} = \text{diag}(\boldsymbol{\sigma}) \boldsymbol{\Omega} \text{diag}(\boldsymbol{\sigma})$ (Lewandowski, Kurowicka, and Joe 2009). The prior on the correlation matrix is a weakly informative LKJ prior with $\nu = 50$, reflecting our prior belief that moderate correlations between the parties are possible⁸. The priors on the standard deviations $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{P-1})$ are independent weakly informative half-normal priors with $\sigma_j \sim \mathcal{N}_+(0, 0.1)$. This reflects our prior belief of a modest random walk for the log-ratio-transformed vote shares/polls.⁹

The priors for the house effects are defined such that, on average, we expect no house effects (i.e., $\delta_{cD} \sim \mathcal{N}(0, 1)$). The priors we chose in our application for the fundamentals Dirichlet regression model, $P(\boldsymbol{\theta})$, are discussed in Online Appendix A.1. We estimate the model by sampling from the posterior distribution via Markov Chain Monte Carlo algorithms, employing the No-U-Turn sampler (Hoffman and Gelman 2014), the default HMC variant in Stan. The code is written in Stan and implemented in rstan 2.17.3 (Stan Development Team 2018).¹⁰

2 Application to the German Federal Election 2017

To demonstrate the virtues of our dynamic Bayesian measurement model, we set it out to an out-of-sample test and applied it to the 2017 German general election.¹¹ We use polling data from the major German polling companies.¹² Figure 1 provides our final forecasts published as of 2

7 The likelihood of the dynamic polls model is given by $P(\mathbf{Y} | \boldsymbol{\alpha}, \boldsymbol{\delta}, \mathbf{W}) = \prod_{c,t} P(\mathbf{y}_{c,t} | \boldsymbol{\alpha}_{t+1}, \boldsymbol{\delta}_c, \mathbf{W})$. In this, \mathbf{Y} is a stacked matrix of all observed polls, $\mathbf{y}_{c,t}$ at c, t , the latent support, $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{T-1}]'$, and the house effects, $\boldsymbol{\delta} = [\boldsymbol{\delta}_1, \dots, \boldsymbol{\delta}_C]'$.

8 We present the estimated correlation matrix in Online Appendix B.3.4.

9 To illustrate, imagine a three party race where party₁ is currently at 0.5, party₂ at 0.3, and party₃ at 0.2. Applying the log-ratio transformation (taking party 3 as baseline - $\log(\frac{\text{party}_i}{\text{party}_3})$) yields the values 0.92, 0.41, and 0. Now a 0.1 increase in the log-ratio value of party₁ to 1.02 would yield the following approximate vote shares: party₁ = 0.52, party₂ = 0.29, and party₃ = 0.19. This would imply a change of vote share of about two percentage points from one day to the next for party₁—a rather unusually big jump.

10 The model code as well as convergence diagnostics are documented in a web application made with Stan Development Team (2017) and available at <https://femps.shinyapps.io/combined-model/>. The estimation was performed with six parallel chains of 1,000 iterations each. 500 iterations were discarded as a warm-up period. In total, 3,000 samples were generated. Standard diagnostic tools indicate satisfying convergence of the chains.

11 The daily forecasts of an older version of our model were made available on <http://zweitstimme.org>, a platform launched for that purpose.

12 In particular, these are *Institut für Demoskopie Allensbach*, *Forschungsgruppe Wahlen*, *forsa*, *Emnid*, *GMS*, *Infratest dimap*, and *INSA*.

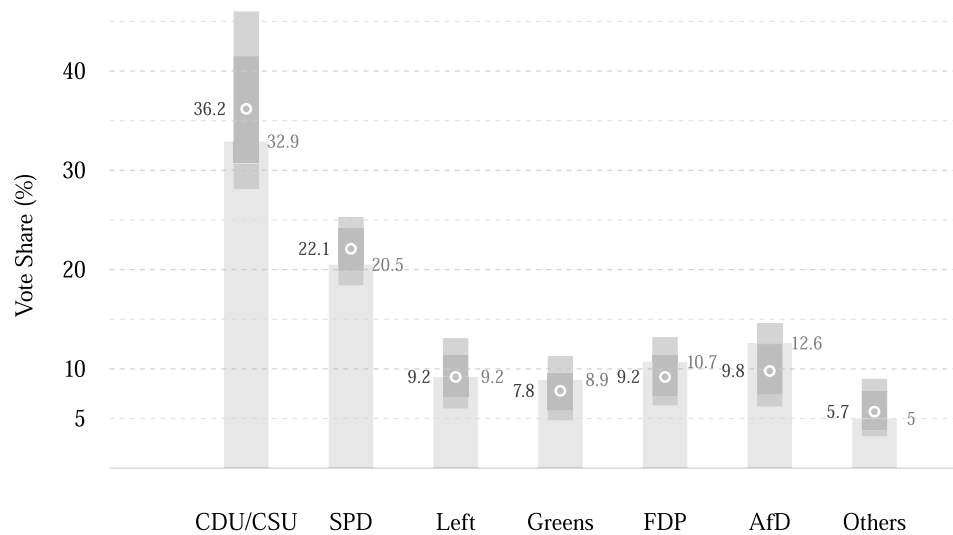


Figure 1. Forecast of the 2017 election 2 days prior to the election. Point estimates along with $\frac{5}{6}$ ($\approx 83\%$) (dark gray) credible intervals and 95% (light gray) credible intervals, the light gray histogram bars represent the election results.

days before the election, along with the respective $\frac{5}{6}$ ($\approx 83\%$) credible intervals.¹³ Accordingly, we predicted that the CDU/CSU would reach 36.2% [30.7%; 41.5%], the SPD 22.1% [19.8%; 24.2%], the Left Party 9.2% [7.2%; 11.4%], the Greens 7.8% [5.9%; 9.6%], the FDP 9.2% [7.2%; 11.4%], the AfD 9.8% [7.4%; 12.5%], and Others 5.7% [3.8%; 7.8%]. As Figure 1 shows, those final forecasts are reasonably close to the final results: Six out of seven $\frac{5}{6}$ -credible intervals include the final outcome and our final forecast has an RMSE of 1.88, which is a comparably small error in multiparty forecasting scenarios.¹⁴

How did the forecast develop over the campaign? Figure 2 highlights a central feature of our model. Early in the campaign, the fundamentals-based model's forecast (dashed horizontal line) still has a substantial impact on the election-day forecast, with the predictive distributions being centered around it. Closer to Election Day, the polls become more informative and can pull the forecast away from the fundamentals-based model. In instances where the final election result (horizontal line) deviates from the fundamentals-based model, this strongly improves the predictive performance (see, e.g., for SPD or AfD), in cases where they coincide, the forecast does not change much (see, e.g., Left Party). In other words, the dynamic component filters new information from the polls, which reflect short-term dynamics unaccounted for by the fundamentals-based component.

Our model is at least as good as the predictions from a pure fundamentals-based model, but closer to the election it gains in predictive accuracy. We show this in more detail with an application to previous elections in Online Appendix B.3.1. In comparison to monthly averages of the poll results (dark points in Figure 2), our model does better in cases where the fundamentals forecast draws our prediction away from the polls and toward the final result, which can happen

¹³ We chose $\frac{5}{6}$ intervals as they are well-suited to communicate uncertainty to a general audience. The observed vote share of each party has a 5-in-6 chance to fall into this interval. This equals the chance of rolling anything other than a six on a fair die.

¹⁴ When evaluating our model based on past elections, it turned out that the forecasts were overconfident, which is due to the fact that the poll-based model exclusively relies on current, not historical polling data and respective errors. To integrate our expectation of deviations between published polls and actual election results, we use a strategy similar to the one employed by Hanretty, Lauderdale, and Vivyan (2016) and add an additional error term to the forecast of each party vote share on Election Day based on how much the polls were off from the actual election results in past elections. The so constructed 95%-credible intervals provide a coverage rate of 94% for the observed results of the last four German elections.

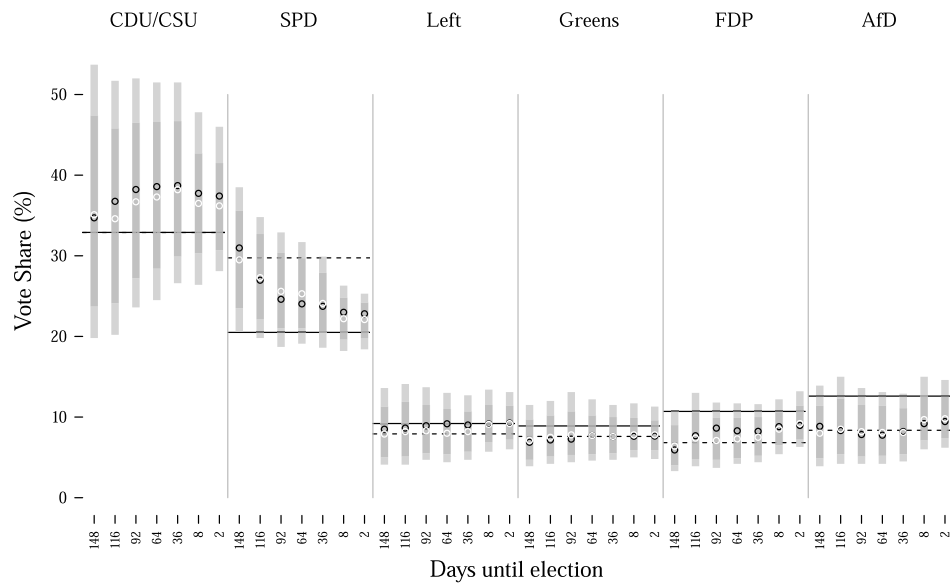


Figure 2. Development of the dynamic Bayesian forecasting model’s vote share predictions over time for the Federal election 2017, starting 148 days until the final prediction 2 days before the election. The light points show the mean prediction; dark gray bars depict the 50% credible intervals and light gray bars the 95% credible intervals. Each party’s observed vote share is indicated by the solid horizontal line. The mean forecast of the fundamentals Dirichlet regression model is marked by the dashed horizontal line. The dark points plot the monthly poll averages.

particularly early in the campaign (see, e.g., for the CDU/CSU 116–36 days before the election). In the case of a strong decline in public support for the SPD, our model adopted more conservatively to the new level, which in this instance led to weaker predictions 116–36 days before the election. A more detailed comparison of the RMSE of the polls compared to our forecast is given in the Appendix B.3.2.

Another feature of the implemented Bayesian setup is that deriving other quantities of interest is straightforward. This is particularly useful in multiparty settings, where relative strengths of parties have important implications for government formation. Drawing on the MCMC simulations, we derived such quantities of interest and correctly predicted that (1) seven parties¹⁵ entered parliament (with a probability of 95%), (2) the AfD had by far the highest probability for becoming the third strongest parliamentary group (43%)¹⁶, and (3), most importantly, that merely two of the plausible coalition options had a reasonable chance to gain a parliamentary majority: the Grand coalition (CDU/CSU–SPD) (>99%) as well as the “Jamaica coalition” of CDU/CSU, Greens, and FDP (96%).

3 Discussion

We developed a dynamic Bayesian forecasting model for multiparty elections. For the first time, we implemented a backward random-walk strategy in this context. This approach allows us to, first, integrate predictions from a fundamentals-based model as priors on Election Day, and second, to mix it with the information we gain through pooling the polls during the election campaign.

To illustrate the generalizability of the approach, we successfully deployed the same model for another step-ahead forecast in the context of the New Zealand general election 2017, the results

¹⁵ Counting CDU and CSU as two parties.

¹⁶ The other probabilities for the small parties to become third strongest parliamentary group are 25% for the FDP, 26% for the Left and 6% for the Greens.

of which are presented in Online Appendix C. Our model correctly predicted the strong increase in Labour vote share and the resulting post-electoral bargaining options.

Finally, a caveat: We believe that while the described model provides an attractive and generic framework for dynamic forecasting of electoral outcomes in many multiparty settings, it should not be applied blindly in other contexts. First, although we found some striking similarities in the performance of fundamentals predictors in both the German and the New Zealand setting, the variable choice for the fundamentals component is likely to be context-specific, which is why we did not consider the details of this modeling step in the main text. Second, our model leverages national-level polling results only. In systems that tend to show strong disproportionality between votes and seats, reliable estimates of parliamentary majorities often require district-level forecasts. If district-level polling data are at hand, it should pay off to integrate these into the dynamic component—similar to the original application (Linzer 2013) or other approaches (e.g., Hanretty, Lauderdale, and Vivyan 2016).

Supplementary material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2018.49>.

References

- Aitchison, John. 1986. The statistical analysis of compositional data. *Journal of the Royal Statistical Society* 44(2):139–177.
- Erikson, Robert S., and Christopher Wlezien. 2013. Forecasting with leading economic indicators and the polls 2012. *PS: Political Science and Politics* 46(1):38–39.
- Hanretty, Chris, Ben Lauderdale, and Nick Vivyan. 2016. Combining national and constituency polling for forecasting. *Electoral Studies* 41(March):239–243.
- Hoffman, Matthew D., and Andrew Gelman. 2014. The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15(1):1593–1623.
- Jackman, Simon. 2005. Pooling the polls over an election campaign. *Australian Journal of Political Science* 40(4):499–517.
- Jérôme, Bruno, Véronique Jérôme-Speziari, and Michael S. Lewis-Beck. 2013. A political-economy forecast for the 2013 German Elections: Who to rule with Angela Merkel? *PS: Political Science and Politics* 46(3):479–480.
- 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.
- Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe. 2009. Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100(9):1989–2001.
- Lewis-Beck, Michael S., Richard Nadeau, and Éric Bélanger. 2016. The British general election: Synthetic forecasts. *Electoral Studies* 41:264–268.
- Lewis-Beck, Michael S., and Ruth Dassonneville. 2015. Comparative election forecasting: Further insights from synthetic models. *Electoral Studies* 39:275–283.
- Linzer, Drew A. 2013. Dynamic Bayesian Forecasting of Presidential Elections in the States. *Journal of the American Statistical Association* 108(501):124–134.
- Magalhães, Pedro C., Luís Aguiar-Conraria, and Michael S. Lewis-Beck. 2012. Forecasting Spanish elections. *International Journal of Forecasting* 28(4):769–776.
- Neunhoffer, Marcel, Lukas F. Stoetzer, Thomas Gschwend, Simon Munzert, and Sebastian Sternberg. 2018. Replication data for: Forecasting elections in multi-party systems: A Bayesian approach combining polls and fundamentals. <https://doi.org/10.7910/DVN/MLYNX0>, Harvard Dataverse, V1.
- Norpoth, Helmut, and Thomas Gschwend. 2010. The chancellor model: Forecasting German elections. *International Journal of Forecasting* 26(1):42–53.
- Silver, Nate. 2012. Methodology. Five Thirty Eight Blog. <https://fivethirtyeight.blogs.nytimes.com/methodology/>.
- Stan Development Team. 2017. shinystan: Interactive Visual and Numerical Diagnostics and Posterior Analysis for Bayesian Models. R package version 2.4.0. <http://mc-stan.org/>.
- Stan Development Team. 2018. RStan: The R interface to Stan. R package version 2.17.3. <http://mc-stan.org/>.

- Strauss, Aaron. 2007. Florida or Ohio? Forecasting presidential state outcomes using reverse random walks. 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.