

Combining Forecasts for the 2021 German Federal Election: The PollyVote

Andreas Graefe, *Macromedia University of Applied Sciences, Munich*

The PollyVote project began in 2004 with the goal of demonstrating the value of combining forecasts for election forecasting. The PollyVote averages forecasts from different methods to generate a forecast of an election outcome. So far, the PollyVote method has been applied to seven major elections: five in the United States and two in Germany. On average, across the last 100 days prior to the five US presidential elections from 2004 to 2020, the combined PollyVote forecast was more accurate than any of its component forecasts (Graefe 2021a). Across the two German federal elections in 2013 and 2017, the combined PollyVote forecast ranked eighth of all 24 forecasts that were available for both elections (Graefe 2019).

However, the PollyVote does not only provide accurate forecasts; in doing so, it also maintains a track record of available forecasts. This enables observers to see how forecasting practice changes over time. For example, Graefe (2021a) showed that for US presidential-election forecasting, fundamentals-only models—which rely solely on structural economic data (e.g., GDP, unemployment, and inflation) or political data (e.g., fiscal spending and war fatalities)—have given way to models that also incorporate public-opinion data (e.g., trial-heat polls and job-approval ratings), such as those published by *FiveThirtyEight.com* and *The Economist*, which currently are considered state of the art. In addition, PollyVote's collection of historical forecasts allows for comparing the relative accuracy of different methods over time and in different contexts. For example, Graefe (2019) showed that the relative accuracy of different methods often varies substantially among elections. In other words, forecasts that worked particularly well in one election often performed poorly in the next election. Of course, this is a major reason why combining forecasts is such a good idea: the method prevents forecasters from choosing a forecast that ends up being way off.

This article continues the long-term tradition of combining forecasts for major American and German elections and presents the PollyVote forecast for the 2021 German federal election.

FORECAST COMPOSITION

For forecasting both German federal elections in 2013 and 2017, the PollyVote averaged forecasts within and across four component methods (i.e., poll aggregators, betting markets, expert judgment, and models) to calculate its combined forecast (Graefe 2019). Before the 2021 election, I revised that procedure in accordance with changes made to the PollyVote prior to the 2020 US presidential election (Graefe 2021a). First, I merged betting markets and expert judgment into a new component (i.e., expectations), to which I added the citizen forecasts generated by Murr and Lewis-Beck (2021). This reduced the number of component methods to three. Second, the PollyVote now distinguishes models based on their underlying information. In the German case, this resulted in two categories of model forecasts: fundamentals-only and fundamentals-plus. Table 1 lists the forecasts of the PollyVote and its components as of June 21, 2021. Forecasts from May 1 to June 21 are available at Harvard Dataverse (Graefe 2021b).

Polls

By far, polls are the most widely known method for forecasting elections, even though they ask respondents only for whom they would vote if the election was held *today*—or, in the German case, the coming Sunday. Therefore, polls do not provide forecasts; they merely capture public sentiment at the time of the survey, which may change—and often does—especially if the election is still far off. As a result, the predictive value of polls typically decreases the longer the time to Election Day. Another issue is that polls conducted around the same time by different survey institutes can vary significantly due to methodological differences. This is how poll aggregation helps because random errors tend to cancel out in the aggregate. However, poll aggregators cannot eliminate biases that are shared across pollsters—for example, due to systematic nonresponse (Gelman et al. 2016).

At the time of this writing, estimates from three poll aggregators were available. *Election.de* calculates simple unweighted averages of the latest polls from seven survey institutes. *Pollytix.de* calculates weighted averages of available polls, placing more weight on more recent polls and those with larger samples. *KOALA* (Bauer et al. 2021) also calculates weighted averages of the most recent polls based on their observed sample sizes and accounts for correlations among pollsters.

This is an updated version of the original article. For details please see the notice at <https://doi.org/10.1017/S1049096521001797>.

Table 1
Vote-Share Forecasts (in %) of the PollyVote and Its Components

	CDU/CSU	SPD	Greens	Left	FDP	AfD	Others
POLLYVOTE	29.2	16.4	18.7	7.5	11.2	10.5	6.4
(80% confidence interval around the forecast)	(23.9–34.6)	(12.4–20.3)	(14.5–22.9)	(4.5–10.6)	(7.8–14.7)	(7.2–13.9)	(3.3–9.5)
Change (in %-points) to 2017 election result	–3.7	–4.1	+9.8	1.7	+0.5	–2.1	+1.2
COMPONENT FORECASTS							
Polls	27.8	15.5	20.6	6.9	12.5	10.2	6.5
Election.de	28.1	15.6	20.6	7.0	12.0	10.3	6.4
KOALA (Bauer et al. 2021)	28.1	15.3	20.4	7.1	12.7	9.9	6.4
Pollytix.de	27.2	15.5	20.8	6.6	12.8	10.4	6.7
Expectations	28.9	16.6	19.4	8.0	10.8	10.1	6.1
Betting markets (Wahlfieber.de)	28.7	15.1	20.5	5.2	12.5	10.6	7.4
Citizen forecasts (Murr and Lewis-Beck 2021)	34.0	21.0	18.0	13.0	11.0	NA	NA
Expert judgment	27.4	16.3	22.7	7.0	10.6	10.6	5.5
Models	30.9	17.0	16.0	7.7	10.3	11.4	6.7
Fundamentals-Plus	30.0	15.7	18.5	6.9	11.5	10.7	6.7
Zweitstimme.org (Gschwend et al. 2021)	31.0	16.0	17.0	7.0	12.0	11.0	6.0
wer-gewinnt-die-wahl.de (Groß 2021)	29.0	15.4	20.0	6.7	11.0	10.4	7.5
Fundamentals-Only	31.8	18.3	13.5	8.6	9.2	12.0	6.6
Länder (Kayser and Leininger 2021)	29.9	19.9	13.6	8.6	9.2	12.0	6.6
Political history (Quinlan, Schnaudt, and Lewis-Beck 2021)	33.8	16.8	NA	NA	NA	NA	NA

Note: Table shows forecasts as of June 21, 2021. Daily updated forecasts are published at pollyvote.com.

Expectations

As discussed previously, I averaged forecasts from two formerly separate components (i.e., betting markets and expert judgment) and added citizen forecasts to create this new component. My reason for doing this is that betting markets, expert judgment, and citizen forecasts are similar with respect to their underlying information: people’s expectations of what will happen on Election Day. In the past, treating these

may be able to estimate whether changes and trends in poll numbers—such as the rise of the Green Party in the spring of 2021—are only temporary or are likely to sustain.

Prior research suggests that experts have value in election forecasting. An analysis of 452 individual expert forecasts made across the four US presidential elections from 2004 to 2016 found that 62% of experts’ forecasts correctly predicted the directional error of polls—even though the

PollyVote’s collection of historical forecasts allows for comparing the relative accuracy of different methods over time and in different contexts.

methods as separate components essentially overweighted expectations in the combined PollyVote. Therefore, the goal of this change was to assign equal weights to polls, expectations, and models, which is consistent with the evidence-based advice to combine forecasts that differ in their underlying information (Armstrong 2001).

Expert Judgment

People typically assume that experts can provide useful forecasts in their domain of expertise. Concerning election forecasting, it seems reasonable to assume that experts can assess how recent and future events may affect polls. Experts also

typical expert forecast incurred a 7% larger error than the RealClearPolitics poll average (Graefe 2018). As is the case with all forecasts, combining individual expert forecasts will reduce error; another study of US presidential elections found that the average forecasts of several experts was more accurate than polls for long-term forecasts (Jones and Cuzán 2013).

I emailed survey invitations to 37 members of the German Society for Electoral Studies who had participated in the 2017 surveys. Respondents were asked to provide their estimate for the vote shares of the parties listed in table 1 in the 2021 federal election. Table 1 shows simple averages of the individual

forecasts from 17 experts who responded in the second survey round from May 28 to May 31, 2021.

Betting Markets

In the German case, betting (or prediction) markets typically allow participants to buy and sell shares of several contracts, one each per political party. The price of each contract is thereby that party's vote-share forecast. For instance, if the

PollyVote estimates an 88% chance that the CDU/CDU, the party of outgoing Chancellor Angela Merkel, will remain the strongest party, with a vote share of 29.2%.

Green Party's contract trades at 20.6 cents, the market predicts the party to gain 20.6% of the vote. If participants believe the final vote share will be higher (lower), they would buy (sell) the Green Party at that price; if they think that the forecast is correct, they would do nothing. Therefore, participants should become active only if they think that they know better than the market. Furthermore, participants have an incentive to do so because they win (or lose) money depending on the accuracy of their forecasts.

Thus, the price mechanism of the market serves as a powerful mechanism to aggregate participants' expectations into a forecast—at least in theory. In practice, especially in recent years, markets provided less accurate forecasts than mainstream methods such as poll aggregators. For example, across the last 100 days prior to the five US presidential elections from 2004 to 2020, betting markets missed the final popular vote by 1.9 percentage points. In comparison, the corresponding error of poll aggregators was 21% lower at 1.5 points (Graefe 2021a). Across the last 46 days prior to the two German elections in 2013 and 2017, the average forecast error of betting markets was 2.1 points—24% higher than the corresponding error of poll aggregators, which was 1.7 points (Graefe 2019).

At the time of this writing, betting-market forecasts were available from *Wahlfieber.de*, which was part of the PollyVote in both 2013 and 2017. *Wahlfieber* is a market operated with play money instead of real money. That is, participants receive a certain endowment of play money that they can use to buy or sell shares of the parties. Individual participants' performance is measured through rankings, and the best performing participants typically win prizes. Although we might expect market manipulation to be higher in play-money markets compared to real-money markets (Graefe 2017), *Wahlfieber* was the most accurate betting market across the two elections in 2013 and 2017 (Graefe 2019).

Citizen Forecasts

Citizen forecasts are derived from survey respondents' answers to questions of who they think will win the election (in addition to for whom they would vote). As shown by Graefe (2014) in the case of US presidential elections, aggregate answers to that question can be translated into highly accurate

vote-share forecasts. Across the last 100 days prior to the seven presidential elections from 1988 to 2012, citizen forecasts were more accurate than polls, betting markets, models, and expert forecasts.

Given the method's predictive accuracy, it is unfortunate that the vote expectation question rarely is included in public-opinion surveys. In the German case, only *Politbarometer* has a historical record of expectation questions going back to 1980.

At the time of this writing, only one such survey was available for the 2021 election, which Murr and Lewis-Beck (2021) translate into vote-share forecasts for the Christian Democratic Union/Christian Social Union (CDU/CSU), the Social Democratic Party (SPD), the Green Party, the Left Party, and the Free Democratic Party (FDP).

Models

Models for forecasting party-vote shares in German federal elections rely on aggregate data. The PollyVote's model component forecast shown in table 1 lists forecasts from four models, two of which rely on structural data (fundamentals-only); the other two models also include polls (fundamentals-plus). PollyVote first averages the forecasts within and then across these categories to calculate its model-component forecast.

Fundamentals-Only

The two fundamentals-only models listed in table 1 forecast the election outcome solely based on fundamental data available long before the election.

The model by Kayser and Leininger (2021) predicts the parties' vote share in each state and aggregates the numbers to predict the federal-election outcome. Their model's predictor variables are national quarterly GDP growth, a party's vote share in the preceding federal and state elections, whether that party provides the chancellor, and the number of years that the chancellor has been in office.

The model by Quinlan, Schnaudt, and Lewis-Beck (2021), a newcomer in this election, relies on four variables to predict the vote shares of the CDU/CSU and the SPD. The four predictor variables measure the parties' vote shares in the previous election, whether the country is governed by a grand coalition at the time of the election, the number of prime ministers that a party holds across German states, and the dynamics of German reunification.

Fundamentals-Plus

The models by *Zweitstimme.org* (Gschwend et al. 2021) and *wer-gewinnt-die-wahl.de* (Groß 2021) are comparable to widely known US models (e.g., *FiveThirtyEight.com* and *The Economist*) in heavily relying on polling data and providing daily updated forecasts of the election outcome.

One major difference is that the German models do not incorporate data on the state of the economy but instead rely on party-competition information such as previous election results and which parties historically formed the government. In addition to predicting the parties' vote shares, which is what is included in the PollyVote, the models provide probabilistic forecasts of various electoral outcomes, such as the chance that a party receives a plurality of the vote or that certain coalitions are mathematically possible.

ELECTION FORECAST

Table 1 lists the vote-share forecasts from the PollyVote and its component for the six largest parties (i.e., CDU/CSU, SPD, Green Party, Left Party, FDP, and Alternative for Germany [AfD]), plus the vote of all remaining parties combined (i.e., Others). As of June 21, 2021, slightly more than three months before the election, the PollyVote predicted that all six parties currently represented in parliament again will pass the 5% electoral threshold. Whereas reentry in the Bundestag is practically certain for five of the six parties, the PollyVote estimates that there is a 14% chance that the Left Party could miss the 5% threshold.¹ Four of the six parties are expected to lose vote shares compared to the 2017 election outcome. In particular, the PollyVote expects voters to punish both parties that formed the grand coalition (i.e., the CDU/CSU and the SPD), which are predicted to lose approximately four percentage points each. Nevertheless, the PollyVote estimates an 88% chance that the CDU/CDU, the party of outgoing Chancellor Angela Merkel, will remain the strongest party, with a vote share of 29.2%. The most left-wing party (i.e., Left Party: -1.7 percentage points) and the most right-wing party (i.e., AfD: -2.1 points) are expected to lose approximately two points each. Only the FDP and the Green Party are predicted to gain votes relative to 2017. Whereas the FDP's gain is likely moderate (+0.5 point), the Green Party is expected to more than double its 2017 result, with a predicted vote share of 18.7% (+9.8 points).

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the *PS: Political Science & Politics* Dataverse: <https://doi.org/10.7910/DVN/KWTJJV>. ■

NOTE

1. Probability forecasts are estimated based on the PollyVote's historical forecast errors across the two German federal elections in 2013 and 2017.

REFERENCES

- Armstrong, J. Scott. 2001. "Combining Forecasts." In *Principles of Forecasting: A Handbook for Researchers and Practitioners*, ed. J. Scott Armstrong, 417–39. New York: Springer Publishing Company.
- Bauer, Alexander, Andreas Bender, Andre Klima, and Helmut Küchenhoff. 2021. "Mundus vult decipi, ergo decipiatur: Visual Communication of Uncertainty in Election Polls." *PS: Political Science & Politics*. In this issue.
- Gelman, Andrew, Sharad Goel, Douglas Rivers, and David Rothschild. 2016. "The Mythical Swing Voter." *Quarterly Journal of Political Science* 11 (1): 103–30. DOI: [10.1561/100.00015031](https://doi.org/10.1561/100.00015031).
- Graefe, Andreas. 2014. "Accuracy of Vote Expectation Surveys in Forecasting Elections." *Public Opinion Quarterly* 78 (S1): 204–32. DOI: [10.1093/poq/nfu008](https://doi.org/10.1093/poq/nfu008).
- Graefe, Andreas. 2017. "Political Markets." In *The SAGE Handbook of Electoral Behavior*, ed. Kai Arzheimer, Jocelyn Evans, and Michael S. Lewis-Beck, 861–82. London: SAGE Publications.
- Graefe, Andreas. 2018. "Predicting Elections: Experts, Polls, and Fundamentals." *Judgment and Decision Making* 13 (4): 334–44.
- Graefe, Andreas. 2019. "Accuracy of German Federal Election Forecasts, 2013 & 2017." *International Journal of Forecasting* 35 (3): 868–877. DOI: [10.1016/j.ijforecast.2019.01.004](https://doi.org/10.1016/j.ijforecast.2019.01.004).
- Graefe, Andreas. 2021a. "Embrace the Differences: Revisiting the PollyVote Method of Combining Forecasts for US Presidential Elections (2004 to 2020)." Social Science Research Network Working Paper. ssrn.com/abstract=3871059.
- Graefe, Andreas. 2021b. "Replication Archive for: Combining Forecasts for the 2021 German Federal Election: The PollyVote." Harvard Dataverse. <https://doi.org/10.7910/DVN/KWTJJV>.
- Groß, Marcus. 2021. "A Long-Short Term Event Memory State-Space Model for Multi-Party Elections." *Zenodo*. DOI: [10.5281/zenodo.3697270](https://doi.org/10.5281/zenodo.3697270).
- Gschwend, Thomas, Klara Müller, Simon Munzert, Marcel Neunhoeffer, and Lukas F. Stotzer. 2021. "The Zweitstimme Model: A Dynamic Forecast of the 2021 German Federal Election." *PS: Political Science & Politics*. DOI: [10.1017/S1049096521000913](https://doi.org/10.1017/S1049096521000913).
- Jones, Randall J., Jr., and Alfred G. Cuzán. 2013. "Expert Judgment in Forecasting American Presidential Elections: A Preliminary Evaluation." Chicago: Annual Meeting of the American Political Science Association. ssrn.com/abstract=2303052.
- Kayser, Mark, and Arndt Leininger. 2021. "A Länder-Based Forecast of the 2021 German Bundestag Election." *PS: Political Science & Politics*. In this issue.
- Murr, Andreas E., and Michael S. Lewis-Beck. 2021. "Citizen Forecasts of the 2021 German Election." *PS: Political Science & Politics*. In this issue.
- Quinlan, Stephen, Christian Schnaudt, and Michael S. Lewis-Beck. 2021. "Forecasting Bloc Support in German Federal Elections: A Political-History Model." *PS: Political Science & Politics*. In this issue.