

# Citizen Forecasts of the 2021 German Election

Andreas E. Murr, *University of Warwick, United Kingdom*

Michael S. Lewis-Beck, *University of Iowa, USA*

There are various scientific approaches to election forecasting: poll aggregation, structural models, electronic markets, and citizen forecasting. With respect to the German case, the first two approaches—polls and models—perhaps have been the most popular. However, relatively little work has been done deploying citizen forecasting (CF), the approach described in this article. In principle, CF differs considerably from other methods and appears, on its face, quite simple. Before an election, citizens are asked in a national survey who they think will win. As the percentage of expectations for party X increases, the likelihood of an X win is judged to be higher. The method has been applied regularly with success in other established democracies, such as the United Kingdom and the United States.

We extend the application to Germany using responses to the expectation question in *Politbarometer* surveys since 1980. The forecasting targets are the vote shares of the following parties: Christian Democratic Union/Christian Social Union (CDU/CSU), Social Democratic Party (SPD), Free Democratic Party (FDP), Green Party, Left Party, and Others. Preliminary analyses of these data appear promising because they generate little prediction error, in- or out-of-sample. This track record allows guarded optimism in terms of the model's ability to accurately forecast the upcoming September 2021 election. We review the literature and the data and measures, followed by model building and relevant performance tests. CF, a hitherto neglected forecasting strategy for the German case, seems to shed considerable light on the sometimes dimly seen workings of this complex coalitional system.

## LITERATURE

Lewis-Beck and Skalaban (1989), who introduced CF as a distinct approach to the prediction of election outcomes, initially applied the method to the national probability samples of the American National Election Study. Dating from 1956, those surveys regularly asked respondents, in advance of the upcoming presidential election, “Who do you think will be elected president in November?” They found that across those eight surveys, voters correctly predicted the winner 69% of the time (Lewis-Beck and Skalaban 1989, 148). Such a fundamental result, illustrating the potential of voter expectations as a forecasting device, was replicated and elaborated on in subsequent American studies (Graefe 2014; Lewis-Beck and Tien 1999; Murr 2015). Furthermore, the approach has spread to Europe, particularly the United Kingdom (Lewis-Beck and

Stegmaier 2011; Murr 2011, 2016). Indeed, a recent paper drawing on 449 surveys from British elections (i.e., 1950 to 2017) demonstrates that forecasting models based on vote expectations clearly outperform those based on the more common method: vote-intention polls (Murr, Stegmaier, Lewis-Beck 2021).

How much attention has been given to CF in the case of German elections? Some, but not much. Ganser and Riordan (2015) discussed CF of vote shares in an *ex-post* study in which vote-intention questions actually performed better. Lehrer, Juhl, and Gschwend (2019) conducted a “wisdom of the crowds” study of the 2017 national election and found that CF does well in predicting the Alternative für Deutschland (AfD) vote—but, again, the effort was *ex-post*. Graefe (2015) reviewed four different forecasting methods applied to the 2013 German election—polls, prediction markets, expert judgment, and quantitative models. He found that polls, on average, were more accurate than other methods. Furthermore, Graefe (2016) conducted a CF of the 2013 election, asking a nonrepresentative sample of citizens to predict 14 election outcomes (e.g., Which candidate will be the next chancellor?, Which parties will exceed the 5% threshold?, and Which coalition will form?); he reported that they got 12 of the 14 correct. However, this also was an *ex-post* exercise. Thus, although CF seems to be garnering attention in the world of German election forecasting, the path forward remains open. For the first time (as far as we know), we offer an *ex-ante* German national election forecast based on 2021 voter expectations. This article discusses the database for that effort next.

## DATA AND MEASURES

The data we use are from the established *Politbarometer*, going back to 1980 (Forschungsgruppe Wahlen 2020). Our replication archive is available on Harvard Dataverse (Murr and Lewis-Beck 2021). The salient sampling characteristics of these surveys are discussed in detail in the online appendix. For our purposes, the most relevant feature is that they are national probability samples. In forecasting, lead time is a *sine qua non*. That is, the prediction must be made before the contest—it is hoped well before—so that the forecast not be judged as trivial (e.g., a day-before-the-election forecast) (Lewis-Beck 2005). The lead time we focus on is two months, which provides a forecast that stands at a significant distance from the election event. Moreover, it has the added advantage of being available for every election in the series. The median sample size of surveys with lead time of two months is

substantial: 1,518. (For a robustness check, we also experimented with one-, three-, and five-month leads, which had surveys for every election except one or two. See the discussion in the online appendix.)

Ideally, we would have multiple expectation questions to accommodate different objects of the forecast. In addition to

*CF, a hitherto neglected forecasting strategy for the German case, seems to shed considerable light on the sometimes dimly seen workings of this complex coalitional system.*

“Who will lead the government?,” we might ask, “What is the vote share you expect for party X?” or “Who will be chancellor?” or “Will there be a single-party government or a coalition government?” Instead, we have a generic question: “Who will win?” Certainly, “winning” can have different meanings for different voters. In an innovative paper, Stiers, Daoust, and Blais (2018) explored voters’ beliefs about the meaning of their party winning. Voters of the party that received the most votes were almost unanimous in perceiving a win; however, even parties that received fewer votes sometimes stated that their party had won—perhaps because they made gains over the previous contest (Stiers, Daoust, and Blais 2018). In the same way, German voters for a smaller party might declare a “win” if they expected their party to be part of a ruling coalition. Furthermore, even if a major German party such as the SPD joins the ruling coalition as a junior partner, this may be perceived as the party “losing” the election.

In addition to asking about winning, our question also is open-ended: “What do you think personally: Who will win the general election?” (our translation). That said, we know from the cited UK and US studies that this question predicts well and can be coded to account for different meanings that different people might have. We find that the answers generally fall into three clear categories: single-party, candidate, and coalition responses. For example, few voters would expect that a small party alone would win, but they might reasonably predict that it would be part of a winning coalition. For the major parties—the CDU/CSU and the SPD—we counted the single-party mentions; for the minor parties—the FDP, the

*Certainly, “winning” can have different meanings for different voters.*

Green Party, and the Left Party—we counted both the single-party and the coalition mentions (the remaining responses form the reference category.) In this way, we arrived at a forecasting model for a specific party.

Because AfD is a new party, emerging in 2013, we could not fit a regression equation to it alone; however, we included it in a generic vote-share equation for Other Parties. We know that the rise of the AfD could upset our predictions—especially because after 2017, the differentiation between major and minor parties became less plausible. Indeed, according to current polling, there may be only one major party left.

#### THE MODEL

In terms of general theory, CF holds that as collective voter expectations increase for a party, the party becomes more likely to win (i.e., govern in whole or in part). Of course, expectations naturally are adaptive, changing with the political winds. For the German case, a strong wind concerns

whether a “grand coalition” is made or is in the making. In fact, when the incumbent consists of a grand coalition such as the CDU/CSU and the SPD, it greatly changes the meaning of “winning” and therefore the nature of expectations. The specification, then, for the prediction equation to be estimated (by ordinary least squares) is Party Vote Share =  $f$  (vote expectation, grand-coalition status). Regression models for each of the five parties, calculated from available data for the 10 elections (i.e., 1983 to 2017), are listed in table 1. (We excluded the 1980 election because the FDP later switched its electoral strategy away from the SPD and toward the CDU/CSU.)

Using this vote-share equation, we can predict the vote share of each party. In terms of predicting the 2021 election, this means also reporting forecasts of which party will have the largest vote share, as well as which coalition scenarios are likely (e.g., which ones reach +50%). Overall, the regression results in table 1 encourage us in our task. First, we examine the goodness-of-fit statistics. The R-squared statistics are good to very good, with a median value of 0.74. The within-sample error is low, with a median of 0.02. The out-of-sample error, based on a “jackknife” cross-validating procedure and excluding each case in turn, is surprisingly low, with a median root mean squared error (RMSE) also of 0.02. Note that the out-of-sample RMSE improves even slightly if we normalize the predictions; this ensures that the sum of all predicted values, including Others, equals 100%. These estimates, especially the out-of-sample error, suggest the forecasting potential of the model. One manifestation of this potential is the prediction of

the chancellor’s party. When we assumed that the highest predicted party vote share will be the chancellor’s party, in a series of out-of-sample tests for each race, we were able to correctly predict nine of the 10 chancellors (incorrect only in 2002).

The regression results indicate much about the functioning of the German party system and the campaign mechanics that allow for rather accurate forecasting. The intercepts for the CDU/CSU and the SPD are the same (0.31), which suggests that they have about the same rather substantial reservoir of votes. The intercepts for the Green Party and the FDP also are

Table 1

Regression Models of Vote Share on Expectations That Party Will Win

	CDU/CSU	SPD	FDP	Green	Left	Others
Intercept	0.31 (0.05)	0.31 (0.02)	0.06 (0.01)	0.06 (0.01)	0.04 (0.01)	–
Expectations	0.20 (0.09)	0.22 (0.07)	0.25 (0.10)	0.60 (0.26)	8.09 (2.10)	–
Grand Coalition	–0.09 (0.03)	–0.10 (0.03)	0.05 (0.01)	0.04 (0.01)	0.04 (0.01)	–
Number of Observations	10	10	10	10	8	–
R-Squared (In-Sample)	0.64	0.86	0.74	0.64	0.88	–
RMSE (In-Sample)	0.04	0.03	0.02	0.01	0.01	–
RMSE (Out-of-Sample)	0.04	0.05	0.02	0.02	0.02	–
RMSE (Out-of-Sample, Normalized)	0.03	0.04	0.02	0.02	0.02	0.06
Chancellors Correctly Predicted (Out-of-Sample)	9 out of 10 (2002 was incorrectly predicted)					

Note: To normalize predictions, we divided by the sum of predicted values across parties if the sum was larger than 1 (e.g., 1998 and 2005). The prediction for Others equals 1 minus the sum of the (normalized) predictions for the remaining parties.

the same (0.06), emphasizing their “underdog” rivalry. Furthermore, the major parties lose about equally (minus 9 or 10 points) when under a grand coalition—a fact that makes sense and confirms some campaign lore. In addition, the gradient of the expectations slope remains roughly equivalent (between 0.20 and 0.25) across the CDU/CSU, the SPD, and the FDP. In fact, taking into account the standard errors of the coefficients, there is no evidence to suggest that the gradient of expectations differs among the CDU/CSU, the SPD, the FDP, and even the Green Party. The only exception is the gradient

including which party will be the largest and which coalitions will receive more than 50% of the vote. According to the *Politbarometer* newsletter of June 10, 2021, when 1,232 respondents were asked, “Who will win the general election?,” their answers were distributed as follows: CDU/CSU/Laschet = 64%; SPD/Scholz = 3%; Green Party/Baerbock = 9%; Other = 9%; and Don’t Know = 15%.

Overwhelmingly, respondents expect the CDU/CSU to “win.” With the caveat that the reported survey results include candidate mentions and exclude coalition men-

*The regression results indicate much about the functioning of the German party system and the campaign mechanics that allow for rather accurate forecasting.*

for the Left Party, although this may be a result of the smaller set of elections.

Broadly speaking, the model’s performance appears laudable, especially given its parsimony, with only two independent variables. Its contours can be further appreciated by examining the scatterplots for vote share regressed on party expectation, shown in figure 1. Observe how closely the points adhere to the line. Indeed, the linearity of the fits generally seems exemplary, with no acceleration in an upward curve as the expectation of winning passes the 50% mark. This condition indicates the ongoing tight competition, at least between the two main parties.

**THE CITIZEN FORECAST CALCULATION FOR SEPTEMBER 2021**

As of June 2021, one *Politbarometer* survey with vote expectations was available, enabling us to forecast with a lead time of about three and a half months. We can plug the survey results into the estimated regression equations of table 1 to predict the vote shares and derive predictions for other outcomes,

tions, we translated them into the following vote-share predictions:

$$\text{Vote Share (CDU/CSU)} = 0.31 + 0.20 * 0.64 - 0.09 = 0.34$$

This prediction does not cross the 50% threshold needed to govern. Which coalition would achieve that necessary majority? According to respondents’ expectations, the SPD is predicted to receive the following vote share:

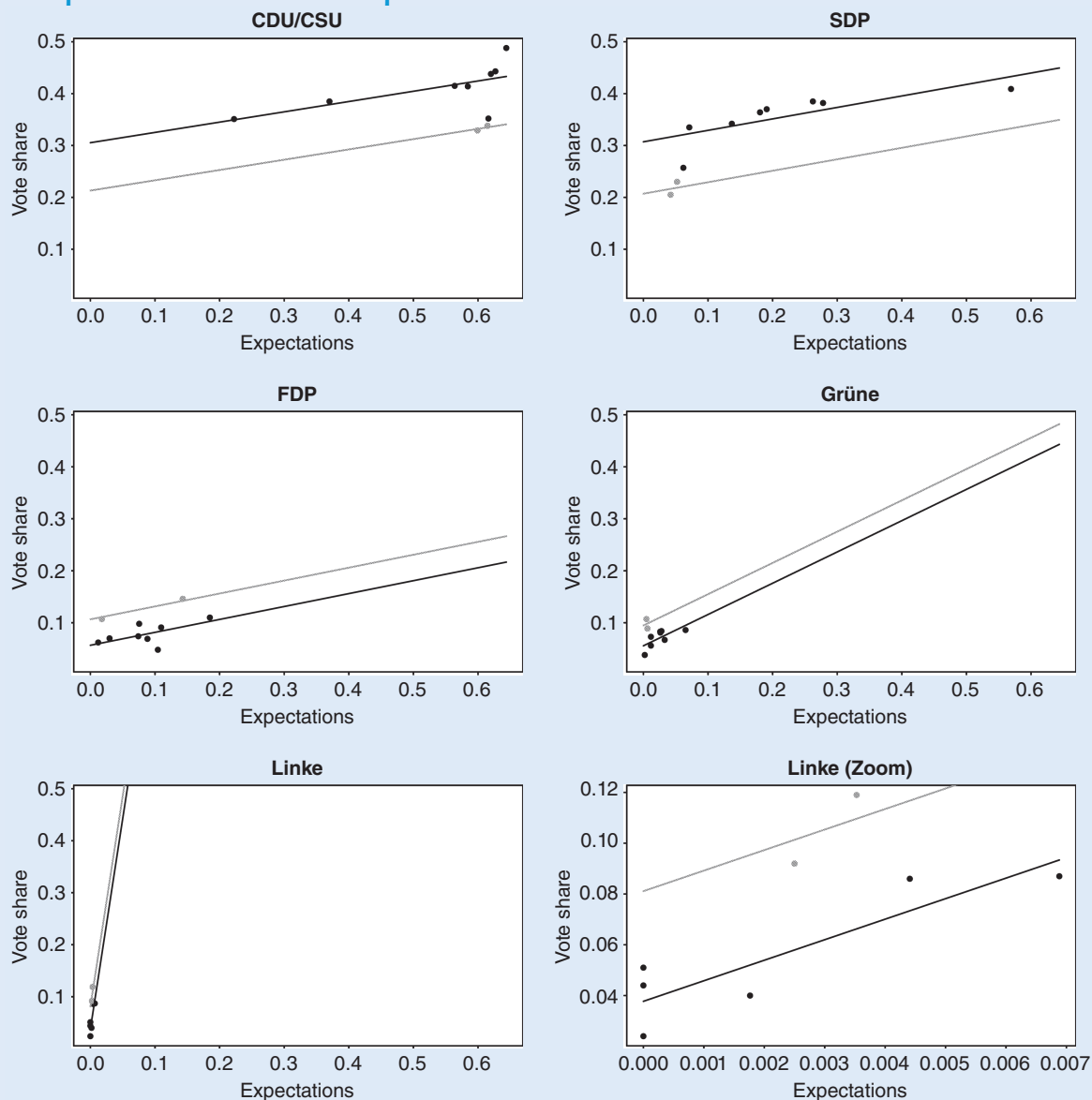
$$\text{Vote Share (SPD)} = 0.31 + 0.22 * 0.03 - 0.10 = 0.21$$

Adding the SPD share to the CDU/CSU share yields a comfortable majority coalition, as follows:

$$0.34 + 0.21 = 0.55$$

However, the CDU/CSU may reject the possibility of again joining with the SPD. If so, according to the currently available expectations data, a coalition with the Green Party would narrowly miss the threshold, as follows:

Figure 1  
Scatterplot of Vote Share on Expectations



Gray dots and lines indicate elections preceded by grand coalitions; black dots and lines indicate elections without a preceding grand coalition. (As a diagnostic tool, we also examine jackknifed predicted versus actual values. See the discussion in the online appendix.)

$$\text{Vote Share (Green Party)} = 0.06 + 0.60 * 0.09 + 0.04 = 0.15$$

Adding the Green Party share to the CDU/CSU share yields no majority, as follows:

$$0.34 + 0.15 = 0.49$$

What can we conclude? Of course, these point estimates are not without error. With respect to coalition formation, the CDU/CSU/SPD combination seems to be a far safer bet statistically (e.g., the absolute value of 55% clearly exceeds

the 49% alternative). However, the CDU/CSU/Green Party combination does not seem to be out of the question, either statistically (i.e., RMSE of 0.02) or politically (as an emergent strategy). Finally, what about the utility of the CF approach? In the German case, it enables us to clearly identify the leading party in the electorate and to identify the most likely coalition combinations.

#### SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1049096521000925>.

---

---

## DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the *PS: Political Science & Politics* Dataverse: doi:10.7910/DVN/WVTI2K. ■

---

## REFERENCES

- Forschungsgruppe Wahlen, Mannheim. 2020. "Politbarometer 1977–2019 (Partielle Kumulation)." GESIS Datenarchiv, Köln. ZA2391 Datenfile Version 12.o.o. <https://doi.org/10.4232/1.13631>.
- Ganser, Christian, and Patrick Riordan. 2015. "Vote Expectations at the Next Level: Trying to Predict Vote Shares in the 2013 German Federal Election by Polling Expectations." *Electoral Studies* 40:115–26.
- Graefe, Andreas. 2014. "Accuracy of Vote Expectation Surveys in Forecasting Elections." *Public Opinion Quarterly* 78 (S1): 204–32.
- Graefe, Andreas. 2015. "German Election Forecasting: Comparing and Combining Methods for 2013." *German Politics* 24 (2): 195–204.
- Graefe, Andreas. 2016. "Forecasting Proportional Representation Elections from Nonrepresentative Expectation Surveys." *Electoral Studies* 42:222–88.
- Lehrer, Roni, Sebastian Juhl, and Thomas Gschwend. 2019. "The 'Wisdom of Crowds' Design for Sensitive Survey Questions." *Electoral Studies* 57:99–109.
- Lewis-Beck, Michael S. 2005. "Election Forecasting: Principles and Practice." *British Journal of Politics and International Studies* 7:145–64.
- Lewis-Beck, Michael S., and Andrew Skalaban. 1989. "Citizen Forecasting: Can Voters See into the Future?" *British Journal of Political Science* 19 (1): 146–53.
- Lewis-Beck, Michael S., and Mary Stegmaier. 2011. "Citizen Forecasting: Can UK Voters See the Future?" *Electoral Studies* 30 (2): 264–68.
- Lewis-Beck, Michael S., and Charles Tien. 1999. "Voters as Forecasters: A Micromodel of Election Prediction." *International Journal of Forecasting* 15 (2): 175–84.
- Murr, Andreas E. 2011. "'Wisdom of Crowds'? A Decentralized Election Forecasting Model That Uses Citizens' Local Expectations." *Electoral Studies* 30 (4): 771–83.
- Murr, Andreas E. 2015. "The Wisdom of Crowds: Applying Condorcet's Jury Theorem to Forecasting US Presidential Elections." *International Journal of Forecasting* 31 (3): 916–29.
- Murr, Andreas E. 2016. "The Wisdom of Crowds: What Do Citizens Forecast for the 2015 British General Election?" *Electoral Studies* 41 (1): 283–88.
- Murr, Andreas E., and Michael S. Lewis-Beck. 2021. "Replication Archive for: Citizen Forecasts of the 2021 German Election." Harvard Dataverse. doi:10.7910/DVN/WVTI2K.
- Murr, Andreas E., Mary Stegmaier, and Michael S. Lewis-Beck. 2021. "Vote Expectations Versus Vote Intentions: Rival Forecasting Strategies." *British Journal of Political Science* 51 (1): 60–67.
- Stiers, Dieter, Jean-François Daoust, and André Blais. 2018. "What Makes People Believe That Their Party Won the Election?" *Electoral Studies* 55: 21–29.