

**Forecasting proportional representation elections
from non-representative expectation surveys**

(Forthcoming with changes in Electoral Studies)

Andreas Graefe

LMU Munich & Macromedia University, Germany

graefe.andreas@gmail.com

Abstract. This study tests non-representative expectation surveys as a method for forecasting elections. For dichotomous forecasts of the 2013 German election (e.g., who will be chancellor, which parties will enter parliament), two non-representative citizen samples performed equally well than a benchmark group of experts. For vote-share forecasts, the sample of more knowledgeable and interested citizens performed similar to experts and quantitative models, and outperformed the less informed citizens. Furthermore, both citizen samples outperformed prediction markets but provided less accurate forecasts than representative polls. The results suggest that non-representative surveys can provide a useful low-cost forecasting method, in particular for small-scale elections, where it may not be feasible or cost-effective to use established methods such as representative polls or prediction markets.

Keywords: election forecasting; citizen forecasts; electoral expectations; expert judgment; forecast accuracy; wisdom of crowds; prediction markets

Introduction

Response rates in traditional phone surveys have decreased below 10% in recent years (Kohut, Keeter, Doherty, Dimock, & Christian, 2012). This trend not only undermines the assumption that respondents form a random and representative sample of the population but also increases the cost and time of conducting surveys. At the same time, the Internet makes it possible to quickly collect responses from non-representative samples at virtually no cost. For example, an opt-in poll on the Xbox gaming platform collected a total of 750,148 responses during the 45 days preceding the 2012 US presidential election (Wang, Rothschild, Goel, & Gelman, 2015).

Such non-representative samples can provide useful information about public opinion or election outcomes if one can account for systematic differences between the survey respondents and the target population (e.g., age, gender, education, ideology, party affiliation, etc.). For example, Wang et al. (2015) used multilevel regression and poststratification to align their highly non-representative sample with the target population. They then calculated forecasts by projecting the adjusted polling results to election day, using an approach similar to Erikson and Wlezien (2008). The resulting forecasts performed equally well than traditional representative polls. This approach of forecasting from non-representative polls is highly promising as a quick and cost-effective alternative to traditional methods. However, one limitation is that the approach requires good data, which may not always be easy to obtain. For example, the first step, poststratification, requires sufficient demographical data on both the survey respondents and the target population. The second step of translating raw polling results to election forecasts requires polling data on historical elections, which may be unavailable for small-scale, local elections.

The present study tests an alternative approach for forecasting elections from non-representative surveys, which does not require additional data. Rather than utilizing responses to the traditional vote intention question, forecasts are derived from responses to the vote expectation question, which asks respondents how they expect the election to turn out. The expectation question is usually kept simple by framing the election outcome as a selection problem. While the exact phrasing depends on the specifics of the particular electoral system, citizens are commonly asked to predict the candidate (or party) that will

lead the government after the election. For example, the question in the American National Election Studies (ANES) asks respondents which candidate they expect to be elected president or who will win the election in their home state. The question in the British General Election Studies asks which party will get the most MPs or, alternatively, which party will win. The question in the German Longitudinal Election Study asks which coalition of parties will form a government.

Although the use of the expectation question in pre-election surveys goes back before the emergence of intention polling (Hayes, 1936), scholars have only recently begun to study its value for predicting election outcomes in plurality elections in the UK and the US (Graefe, 2014, 2015a; Lewis-Beck & Stegmaier, 2011; Murr, 2011, 2015a, 2015b; Rothschild & Wolfers, 2012). For example, one study compared the accuracy of the expectation question to polls, prediction markets, quantitative models, and expert judgment for predicting election winners and vote shares in the seven US presidential elections from 1988 to 2012. Across the last 100 days preceding each election, responses to the expectation question correctly predicted the election winner with a hit rate of 92%, which was more accurate than the corresponding hit rate of polls (79% correct), prediction markets (79%), expert judgment (66%), and quantitative models (86%). When predicting vote shares, expectations were again most accurate. Compared to traditional polls, expectations reduced forecast error by 51%. Compared to prediction markets, the most accurate of the four methods, error was reduced by 6% (Graefe, 2014). Another study used ANES data from the 15 elections from 1952 to 2008 to analyze the relative accuracy of the expectation question and the intention question when both are asked in the same survey. The expectation question provided more accurate forecasts than the intention question when predicting election winners, vote shares, and probabilities of victory. Furthermore, the study showed that the expectation question also performs well with small and highly non-representative samples. Aggregated and statistically adjusted expectations from two subsamples that contained only Democratic or only Republican voters yielded more accurate forecasts than the complete sample of vote intentions (Rothschild & Wolfers, 2012)

In sum, prior research has shown that expectation surveys provide highly accurate forecasts in plurality elections. However, it is unclear whether the findings generalize to elections held in more

complex electoral systems. The present study thus tests the predictive value of expectations derived from a non-representative sample for forecasting an election in a multi-party system with proportional representation, namely the 2013 German Federal Election.

Barriers to Expectation Surveys in Electoral Systems with Proportional Representation

While the use of expectation surveys is straightforward in plurality elections, their implementation is more challenging in multi-party systems with proportional representation due to theoretical and methodological barriers discussed in this section.

Theoretical barriers

In plurality elections, citizens often face a rather simple prediction task: they have to pick the winner out of two possible options. While there have been small third-party movements in the US in the past, their candidates never had a realistic chance to win the election. Thus, even with more than two parties (or candidates) in the race, the task of predicting US presidential election winners is essentially reduced to a binary choice problem. This also holds for elections with stronger third parties such as in the UK. Although there is an established third party (the Liberal Democrats, which can count on up to 20% of the vote) and other smaller parties (which, in total, have received up to 12% of the vote), one of the two major parties (Labour or Conservative) has always controlled the government since World War II. Thus, when being asked to predict the winner, nearly all respondents will name either Labour or Conservative. Similar to the US, the prediction task is thus reduced to a binary choice problem and most respondents are able to accurately forecast the election winner (Lewis-Beck & Stegmaier, 2011; Murr, 2011).

In contrast, predicting election outcomes in multi-party systems with proportional representation is more complicated. The German federal election, which is the subject of the present study, provides an example. In the German electoral system, voters cast two votes. While the first vote determines the regional representative by plurality, the second vote determines the distribution of seats by party in the Bundestag. The system thus combines a plurality vote for regional representatives with elements of proportional representation (Ganser & Riordan, 2015).

Although there are two major parties, the Christian Democratic Union (CDU/CSU) and the Social Democrats (SPD), the 1957 election was the only time a party (i.e., CDU/CSU) gained an absolute majority in the history of the German Bundestag. In order to establish a governing majority of seats, parties thus usually need to form a coalition with one or more smaller parties that are represented in parliament. In order for a party to be represented in parliament, it has to exceed the electoral threshold by gaining at least 5% of the vote.

In such an electoral system, voters may engage in strategic voting. In the most basic form of strategic voting, which to some extent is also present in US and UK elections, voters may avoid wasting votes for parties that have no chance of entering parliament. However, voters may also choose to depart from their preferred party and give their vote to another party, for example, to help a potential coalition partner exceed the electoral threshold. The extent to which voters engage in strategic voting then depends on their expectations about how many votes the parties are likely to get and which coalitions are likely to be formed (Meffert, Huber, Gschwend, & Pappi, 2011). Although only a small share of voters might possess the political knowledge to use such strategies, predicting their behavior and its effect on the election outcome is a difficult task.

Furthermore, decreasing turnout has made German election outcomes increasingly volatile and thus unpredictable. The results of the past two elections illustrate this (Rossteutscher, Faas, & Arzheimer, 2015). In the 2009 election, the vote shares of both CDU/CSU (33.8%) and SPD (23.0%) fell to historical lows, while the Free Democrats (FDP) achieved the best result in their party's history (14.6%). Only four years later, however, the 2013 election produced a very different outcome. For the first time in history, the FDP failed to pass the 5% threshold and thus did not enter parliament. With a vote share of 41.5%, which was the party's best result since the 1990 election, the CDU/CSU almost achieved an absolute majority of votes. The SPD scored 25.7% of the vote and became the new junior partner in a grand coalition of CDU/CSU/SPD. Finally, the Alternative for Germany (AfD), a new right-wing euroskeptic party founded little more than half a year before the election, almost gained 5% of the vote and thus only narrowly missed representation in parliament.

Compared to plurality elections such as in the US and the UK, multi-party elections with proportional representation, such as in Germany, are more uncertain and should thus be more difficult to predict. Therefore, one would expect expectation surveys to be less useful for forecasting elections in such a system.

Methodological barriers

Responses to the traditional expectation question provide direct forecasts of who will win but cannot be interpreted as vote share predictions. For example, a survey that shows that 55% of respondents expect the Democrats to win the 2016 US presidential election does not mean that the Democrats will gain 55% of the vote. However, one can use simple linear regression analysis to translate expectation survey results into numerical predictions of vote-shares or seats achieved by a certain candidate or party. For example, to predict the incumbent's vote-share in US presidential elections, Lewis-Beck and Tien (1999) regressed the incumbent's actual vote shares on his support in the expectation surveys using data from historical elections. Similar approaches have been used to calculate seat forecasts in UK general elections by using the winning party's actual seat share as the dependent and the winning party's support in the expectation surveys as the independent variable (Lewis-Beck & Stegmaier, 2011).

This procedure of using regression analysis for generating numerical forecasts from categorical responses to expectation surveys can only be applied under certain conditions, however. First, sufficient data need to be available to run the analysis. That is, there need to be enough observations with expectation surveys from historical elections. Second, the data need to be valid and measure the same information. That is, the questions used in historical surveys should be phrased similarly. Third, it has to be clear what the dependent and the independent variables are (Lewis-Beck & Stegmaier, 2011). The selection of the dependent variable is obvious in plurality elections, where the incumbent party's vote is commonly used. However, this question is tricky in multi-party systems with proportional representation, as shown in the case of the 2013 German election. First, established parties may fail to exceed the electoral threshold, such as the FDP in 2013. Second, new parties may enter parliament, which almost happened with the AfD. Third, coalition signals may change over time. For example, a new coalition

between CDU/CSU and the Green party seemed possible prior to the 2013 election. In such a situation, it becomes difficult to determine the dependent variable for calculating numerical forecasts. In addition, and perhaps more importantly, while a forecast that provides the vote share of the strongest party is sufficient in plurality systems as it reveals the outcome of the election, such a forecast is usually of little value in multi-party systems with proportional representation. For example, what if the strongest party is unlikely to find potential coalition partners? Quantitative election forecasting models try to account for this problem by using the vote share of the outgoing coalition as the dependent variable (e.g., Kayser & Leininger, 2016; Norpoth & Gschwend, 2013). But again, such a forecast is of little practical value if the coalition may not persist, such as in 2013, when the FDP failed to enter parliament. In any case, much information would be lost. The reason is that such forecasts do not reveal whether or not a particular party is likely to enter parliament or which alternative coalitions are possible.

Ideally, one would like to have vote share predictions for each party (as provided by traditional polls), as such figures would allow for calculating different scenarios of what is likely to happen. Unfortunately, traditional regression analysis does not allow for calculating vote-share forecasts for multiple parties and the limited data available precludes using more sophisticated statistical methods.

A straightforward alternative is to directly ask respondents about what percentage of the vote they expect each party to receive. Two previous studies have used such an approach. Sjöberg (2009) asked citizens and three groups of experts (political scientists, journalists, and editors) to predict the vote-shares of seven parties in the 2006 Swedish Parliamentary election. The combined citizen forecast performed well and outperformed the combined expert forecast. Ganser and Riordan (2015) analyzed data from a representative telephone survey, in which 1,000 respondents were asked to predict the vote-shares for seven parties in the 2013 German election. They found that the aggregated expectations performed similar to the raw intention data but provided less accurate forecasts than the finally published (and therefore adjusted) poll results. The authors then concluded that voters are unable to provide vote-share forecasts for the difficult task of predicting German elections. There are two caveats, however. First, providing numerical vote share predictions during a phone interview may indeed be too difficult for participants.

Second, the authors compared the accuracy of people's expectations only to intention polls, which performed particularly well in 2013 compared to historical elections. In general, more evidence is necessary to draw definite conclusions about a method's adequacy under certain conditions.

The present study builds on this work and reports results from an online panel survey that was conducted prior to the 2013 German election and which is described in the following section.

Survey Design

The online survey was conducted in five waves prior to the 2013 German election, which was held on September 22nd. The five waves started on July 17th, August 13th, September 3rd, September 10th, and September 17th. Participants could enter the survey at any wave. At the beginning of the questionnaire, respondents were asked to state their vote intention and to rate their knowledge about (and interest in) politics as well as their self-interest in the election outcome on a five-point scale from 1 (very little) to 5 (very high). Then, respondents were asked about their expectations of several election outcomes. First, respondents were asked (i) which of two candidates (Angela Merkel, CDU/CSU, and Peer Steinbrück, SPD) will be the next Chancellor, (ii) which of five small parties (Green Party, Left Party, FDP, AfD, and Pirate Party) will exceed the 5% electoral threshold and thus enter parliament, and (iii) which of seven possible coalitions will form the next government. For each of the 14 possible outcomes, respondents expressed their expected likelihood of occurrence on a four-point scale: ("definitely not", "quite uncertain", "quite certain", "completely certain"). Second, respondents were asked to predict the vote-shares received by the seven largest parties (as indicated by the national polls) as well as the remaining vote share that would be received by all other parties. Respondents were told that the vote shares should sum up to 100 but the survey software did not enforce this. Finally, respondents were asked to provide demographic information such as gender, age, and level of education.

Respondents were divided into two groups (panelists and webusers), depending on how they were recruited. *Panelists* were recruited by sending 4,000 email invitations to the noncommercial online access convenience SoSci Panel, whose members agreed to participate in scientific surveys. The panel has two major advantages compared to traditional student samples. First, it allows for recruiting a large number of

participants. Second, the resulting samples are more heterogeneous than student samples regarding age, geography, professional background, and personal interest. The panelists are, however, not representative of the German population. In particular, panelists are younger and better educated than the general public, since a large share (about 40%) of the about 60.000 active panel members are students. For more information on the composition and limitations of the panel see Leiner (2014). The second group, *webusers*, received no personal invitation but found a public link to the online survey, which was posted on a German election forecasting website (pollyvote.de). That is, webusers were self-selected participants that can be expected to have a strong interest in the election.

In order to obtain benchmark estimates, an expert survey was conducted in parallel. A research assistant collected names and corresponding email addresses of 332 political journalists from a public relations and media database (i.e, *Zimpel Online*, www.zimpel.de) as well as 180 scholars of German elections with at least a doctoral degree who worked at universities or think tanks.

Survey Data

The data presented in this section are based on completed questionnaires; incomplete questionnaires were removed from the dataset. The complete data and calculations are publicly available (Graefe, 2016).

Survey response

Part A of Table 1 shows the survey response per group. A total of 1,276 panelists participated in at least one of the five survey waves, which corresponds to a response rate of 32%. The total number of panelist responses was 3,399; that is, a typical respondent participated in 2.7 survey waves. The average number of responses per wave was 680. In addition, 437 webusers completed the survey. Note, that the webusers' average response per wave (87 responses) is skewed due to the large response in the last wave. This is the result of increased traffic to the website that posted the survey link as the election came closer.

In the expert survey, a total of 210 experts (106 scholars and 104 journalists) participated, which corresponds to a response rate of 41%. On average across the five waves, a total of 122 expert responses were collected.

Table 1: Response, political knowledge, demographics, and vote intention of survey participants

A. Survey response	Citizens		Experts
	Panelists	Webusers	
Invited	4,000	NA	512
Unique respondents	1,276	437	210
Response rate	32%	NA	41%
Total number of responses	3,399	437	612
Average number of responses per wave	680	87	122

B. Political knowledge and interest (column: mean)	Citizens		Experts
	Panelists	Webusers	
Political knowledge	3.4	4.1	4.5
Interest in politics	3.6	4.5	4.8
Self-interest in election outcome	4.0	4.5	4.3

C. Demographics (column: percent, relative to election turnout)	Citizens		Experts	Election turnout
	Panelists	Webusers		
<i>Gender</i>				
Men	0	32	28	48
Women	0	-32	-28	52
<i>Age</i>				
20-	-1	-1	-3	3
21-30	27	-5	-10	11
31-40	3	-2	11	12
41-50	-9	-2	6	18
51-60	-12	-2	4	19
61-70	-9	15	-7	15
71+	0	-3	-1	21

D. Vote intention (column: percent, relative to election result)	Citizens		Experts	Election result
	Panelists	Webusers		
Christian Democratic Union (CDU/CSU)	-19	-20	-18	41.5
Social Democrats (SPD)	-7	9	-4	25.7
Green Party (Grüne)	20	9	26	8.4
Left Party (Linke)	2	-2	-2	8.6
Free Democratic Party (FDP)	2	2	7	4.8
Alternative for Germany (AfD)	-1	5	-3	4.7
Pirates (Piraten)	4	0	-2	2.2
Other Parties (Sonstige)	-1	-3	-4	4.1
<i>MAE</i>	7	6	8	

Political knowledge and interest

Part B of Table 1 shows participants' ratings on their knowledge about—and interest in—politics as well as their self-interest in the election outcome. Experts were most interested and rated themselves as highly knowledgeable, although webusers also showed high ratings. This was expected as webusers were self-selected respondents who are only likely to visit the website and participate in the survey if they are interested. In comparison, panelists showed much lower levels of interest and knowledge in politics. Regarding self-interest in the election outcome, the results are similar in that experts and webusers

assigned higher ratings than panelists. However, it is interesting to note that webusers were most concerned about the election result. Again, this supports the assumption that webusers were highly interested citizens.

Demographics

Part C of Table 1 provides demographic information about the respondents, depicted as percentage point deviations from the demographics of the actual electorate (listed in the last column). As expected, neither group is by any means close to a representative sample.

For example, according to official election statistics from Bundeswahlleiter (2014), there were 52% female and 48% male voters in the 2013 election. This gender distribution is similar to that of the panelist group. However, male respondents clearly dominated groups of webusers and experts. For example, in the webusers group, the share of male respondents was 32 percentage points higher (i.e., 80% of webusers were male) than in the electorate. In other words, there was only about one woman out of four respondents in both the webusers and the expert group. There were also substantial differences in terms of the respondents' age across the three groups. The panelists formed the youngest group, with a mean age of 37 years, followed by experts (49 years), and webusers (56 years). In addition, the age distribution in all three groups differed substantially from that of the actual electorate. For example, the panelists' share of the 21 to 30-year-olds was 27 percentage points higher than the corresponding share in the actual electorate.

Vote intention

Part D of Table 1 shows the aggregated results of the vote intention question per group, relative to the final election result. A negative value means that the share of respondents who intended to vote for a party was smaller than that party's actual vote share, and vice versa.

As one would expect, given the highly non-representative samples, the results differ widely from the official election results. The differences are most striking for the Conservatives and the Green Party. In all three groups, the share of respondents who intended to vote for the Conservatives was about 20

percentage points lower than in the general public. In comparison, support for the Green Party was much higher among all groups, in particular for panelists and experts.

Needless to say, one would obtain large forecast errors if the responses to the vote intention question of such highly non-representative samples would be directly interpreted as forecasts of what will happen on Election Day. For example, the mean absolute error (MAE) of the panelists vote intention was 7 percentage points. The MAE is the unweighted average of the errors for all parties, which are calculated as the absolute percentage point deviation between a party's forecast and actual vote (Mitofsky, 1998). The lower the MAE, the more accurate the forecast.

Accuracy of expectations

While forecasts generally became more accurate closer to Election Day, the results across waves did not differ between groups. Therefore, results across waves were averaged in order to simplify presentation.

Relative accuracy of citizen and expert surveys

Predictions for each of the 14 dichotomous outcomes were counted as correct if (a) outcomes that did occur (e.g., Angela Merkel will be elected Chancellor) were judged as “quite certain” or “completely certain” or (b) outcomes that did not occur (e.g., the Liberal Party will enter the parliament) were judged as “quite uncertain” or were expected to “definitely not” occur. Otherwise, judgments were counted as wrong. There were no differences in accuracy across groups. On average across the five waves, panelists and webusers correctly predicted 12 of the 14 outcomes and thus performed similar to the expert group. The two outcomes that were commonly missed were FDP's failure to enter parliament and the resulting formation of a grand coalition of CDU/CSU/SPD.

However, forecast accuracy across groups differed for vote share forecasts. The webuser survey yielded a MAE of 1.49 percentage points and was thus slightly more accurate than the expert survey (1.52 percentage points). In contrast, the panelist survey yielded a MAE of 1.87 percentage points, which is 24% higher than the error of the expert survey.

Accuracy of citizens' vote share forecasts compared to benchmark methods

We also compared the accuracy of citizens' vote share predictions to forecasts from three benchmark methods that are commonly used to forecast elections, namely polls, prediction markets, and quantitative models. Polls ask people for whom they intend to vote if the election was held today. Although, strictly speaking, polls do not provide predictions of what will happen on Election Day but measure public opinion at a certain point in time, polling results are commonly projected to election day and interpreted as forecasts (Hillygus, 2011). Prediction markets allow people to bet on the election outcome. Participants win (or lose) money depending on the accuracy of their predictions, and thus have an incentive to make accurate predictions. The resulting betting odds can then be interpreted as forecasts of the election results (Erikson & Wlezien, 2012). In eliciting expectations, prediction markets are thus closely related to expectation surveys. Quantitative models predict the election outcome from information about a country's economic (e.g., growth, unemployment, inflation) and/or political (incumbency, job approval) condition several months before the election. Most of these so-called political economy models rely on the idea of retrospective voting. That is, voters are assumed to reward (or punish) the incumbent government based on past performance. Two models that have long been used to forecast German elections are those by Jérôme, Jérôme-Speziari, and Lewis-Beck (2013) and Norpoth and Gschwend (2013).

Table 2 shows the absolute error of the citizens' expectations, relative to the most recent forecast of each benchmark method's typical component at the last day of each survey wave. Data for prediction markets, models, and polls were derived from a publicly available dataset that contains 2013 German election forecasts from six pollsters, six prediction markets, and five quantitative models. For a description of each individual method included in the dataset see Graefe (2015b). A relative absolute error above (below) 1 means that the citizen forecast was less (more) accurate than the benchmark. For example, compared to the typical prediction market, the panelists' relative absolute error was 0.89, which means that the error of the panelists' expectations was 11% ($= 1 - 0.89$) lower than the corresponding error of the typical prediction market. In general, both citizen groups did well compared to prediction markets, which

also rely on expectations of self-selected participants, but were outperformed by forecasts from traditional intention polls. Compared to quantitative models, performance of panelists and webusers was mixed.

Table 2: Absolute error of citizens' vote share forecasts relative to the typical benchmark forecast

Typical benchmark	Panelists	Webusers
Prediction market	0.89	0.71
Quantitative Model	1.27	1.01
Poll	1.47	1.17

Discussion

Prior research found that representative surveys of citizens' expectations of outcomes of plurality elections provide forecasts that are often at least as accurate as forecasts derived from other methods. The present study builds on this work in two ways by, first, testing the performance of expectation surveys of highly non-representative samples and, second, applying expectation surveys to the more complex problem of predicting an election in a multi-party system with proportional representation, namely the 2013 German federal election.

The results conform to prior evidence from plurality elections in that the majority of citizens were able to provide correct answers to questions that are probably most important to the general election observer, such as who will lead the government or which parties will be represented in the parliament. Both non-representative citizen samples correctly predicted 12 of 14 dichotomous election outcomes and thus performed similar to the expert survey.

However, in contrast to plurality elections, methodological barriers make it difficult to translate citizens' aggregated responses to categorical questions into numerical multi-party vote share forecasts. In such a situation, it is necessary to directly ask respondents to provide forecasts of the parties' expected vote shares. For this task, accuracy differed between both citizen samples. While the group of self-selected webusers performed equally well than the experts, the panelists provided less accurate forecasts. I also compared the accuracy of both citizen groups to three established benchmark methods. Interestingly, both citizen samples provided more accurate forecasts than prediction markets. However, similar to findings by

Ganser and Riordan (2015), they were less accurate than regular polls. Compared to quantitative models, performance was mixed, with webusers performing similar to the typical model.

The different performance of webusers and panelists suggests that the procedure used for sampling the citizen groups affected the results. Webusers, who rated themselves high in interest in politics and the election outcome, accessed the survey by following a public link that was posted on the German PollyVote.de forecasting website. This website published and combined forecasts from four different methods (i.e., polls, prediction markets, econometric models, and experts' judgment), following an approach that has proved successful in forecasting election in the US (Graefe, Armstrong, Jones, & Cuzán, 2014b) and in Germany (Graefe, 2015b). That is, people who visited the website had access to a large number of forecasts as well as to a combined forecast that can be expected to provide highly accurate forecasts of the election outcome. The findings from the present study suggest that webusers were able to harness this information to form accurate expectations of what will happen on Election Day. Future research should focus on developing methods to identify the most accurate forecasters in a sample; see Murr (2015a) for an important first step in this direction.

Needless to say, the present analysis, although based on five survey waves, uses data from only one election and, therefore, any conclusions about the methods' relative performance should be tentative. In particular, since we know from prior research that the accuracy of different forecasting methods often varies widely across elections. The reason is that every election is held in a different context (e.g., economic conditions, government popularity, foreign affairs) and has its idiosyncrasies and, therefore methods that worked well in the past might not work well in the future. For example, across the six US presidential elections from 1992 to 2012, methods that provided the most accurate forecasts in one election were often among the least accurate in another election (Graefe et al., 2014b). Similarly, the polls' good performance in the 2013 German election might just be an outlier given that polls provided less accurate forecasts than prediction markets in both the 2005 and the 2009 election (Groß, 2012; Schaffer & Schneider, 2005). Therefore, a general rule is to not rely on a single method. Instead, the best

way to generate accurate forecast and to avoid large errors is to combine forecasts from different methods that use different data (Graefe, 2015b; Graefe, Armstrong, Jones, & Cuzán, 2014a).

The present paper is only a first step to assess the value of non-representative expectation surveys for forecasting proportional representation elections. However, initial results are promising and call for further research on elections in different countries and electoral systems. Thereby, a particular focus should be on small-scale or local elections for which it may not be feasible or cost-effective to conduct a representative survey, run a prediction market, or develop a quantitative model. In such situations, expectations from non-representative samples could provide a useful low-cost alternative.

References

- Bundeswahlleiter. (2014). *Heft 4: Wahlbeteiligung und Stimmabgabe der Männer und Frauen nach Altersgruppen*. Retrieved from https://www.bundeswahlleiter.de/de/bundestagswahlen/BTW_BUND_13/veroeffentlichungen/BTW2013_Heft4.pdf.
- Erikson, R. S., & Wlezien, C. (2008). Are political markets really superior to polls as election predictors? *Public Opinion Quarterly*, 72(2), 190-215.
- Erikson, R. S., & Wlezien, C. (2012). Markets vs. polls as election predictors: An historical assessment. *Electoral Studies*, 31(3), 532-539.
- Ganser, C., & Riordan, P. (2015). Vote expectations at the next level. Trying to predict vote shares in the 2013 German federal election by polling expectations. *Electoral Studies*, 40, 115-126.
- Graefe, A. (2014). Accuracy of vote expectation surveys in forecasting elections. *Public Opinion Quarterly*, 78(S1), 204-232.
- Graefe, A. (2015a). Accuracy gains of adding vote expectation surveys to a combined forecast of US presidential election outcomes. *Research & Politics*, 2(1), 1-5.
- Graefe, A. (2015b). German election forecasting: Comparing and combining methods for 2013. *German Politics*, 24(2), 195-204.
- Graefe, A. (2016). Replication Data for: Forecasting proportional representation elections from non-representative expectation surveys (Publication no. doi/10.7910/DVN/25XOBR). from Harvard Dataverse.
- Graefe, A., Armstrong, J. S., Jones, R. J. J., & Cuzán, A. G. (2014a). Accuracy of Combined Forecasts for the 2012 Presidential Election: The PollyVote. *PS: Political Science & Politics*, 47(2), 427-431.
- Graefe, A., Armstrong, J. S., Jones, R. J. J., & Cuzán, A. G. (2014b). Combining forecasts: An application to elections. *International Journal of Forecasting*, 30(1), 43-54.
- Groß, J. (2012). Märkte und Prognosen. In N. Braun, M. Keuschnigg, & T. Wolbring (Eds.), *Wirtschaftssoziologie II: Anwendungen* (pp. 111-126). München: Oldenbourg.
- Hayes, S. P. J. (1936). The predictive ability of voters. *Journal of Social Psychology*, 7(2), 183-191.
- Hillygus, D. S. (2011). The evolution of election polling in the United States. *Public Opinion Quarterly*, 75(5), 962-981.
- Jérôme, B., Jérôme-Speziari, V., & Lewis-Beck, M. S. (2013). A political-economy forecast for the 2013 German elections: Who to rule with Angela Merkel? *PS: Political Science & Politics*, 46(3), 479-480.
- Kayser, M., & Leininger, A. (2016). A Predictive Test of Voters' Economic Benchmarking: The 2013 German Bundestag Election. *German Politics*, 25(1), 106-130.

- Kohut, A., Keeter, S., Doherty, C., Dimock, M., & Christian, L. (2012). *Assessing the representativeness of public opinion surveys*. Washington, D.C.: Pew Research Center .
- Leiner, D. (2014). Convenience samples from online respondent pools: A case study of the SoSci Panel. *Working paper*, Available at <https://www.researchgate.net/publication/259669050>.
- Lewis-Beck, M. S., & Stegmaier, M. (2011). Citizen forecasting: Can UK voters see the future? *Electoral Studies*, *30*(2), 264-268.
- Lewis-Beck, M. S., & Tien, C. (1999). Voters as forecasters: a micromodel of election prediction. *International Journal of Forecasting*, *15*(2), 175-184.
- Meffert, M. F., Huber, S., Gschwend, T., & Pappi, F. U. (2011). More than Wishful Thinking: Causes and Consequences of Voters' Electoral Expectations about Parties and Coalitions. *Electoral Studies*, *30*(4), 804-815.
- Mitofsky, W. J. (1998). Review: Was 1996 a Worse Year for Polls Than 1948? *Public opinion quarterly*, *62*(2), 230-249.
- Murr, A. E. (2011). "Wisdom of crowds"? A decentralised election forecasting model that uses citizens' local expectations. *Electoral Studies*, *30*(4), 771-783.
- Murr, A. E. (2015a). The wisdom of crowds: Applying Condorcet's jury theorem to forecasting US presidential elections. *International Journal of Forecasting*, *31*(3), 916-929.
- Murr, A. E. (2015b). The wisdom of crowds: What do citizens forecast for the 2015 British General Election? *Electoral Studies*, doi:10.1016/j.electstud.2015.1011.1018.
- Norpoth, H., & Gschwend, T. (2013). Chancellor model picks Merkel in 2013 German election. *PS: Political Science & Politics*, *46*(3), 481-482.
- Rossteutscher, S., Faas, T., & Arzheimer, K. (2015). Voters and Voting in Multilevel Systems – An Introduction. *German Politics*, *24*(1), 1-7.
- Rothschild, D., & Wolfers, J. (2012). Forecasting elections: voter intentions versus expectations. *Working paper*, Available at: ssrn.com/abstract=1884644.
- Schaffer, L.-M., & Schneider, G. (2005). Die Prognosegüte von Wahlbörsen und Meinungsumfragen zur Bundestagswahl 2005. *Politische Vierteljahresschrift*, *46*(4), 674-681.
- Sjöberg, L. (2009). Are all crowds equally wise? A comparison of political election forecasts by experts and the public. *Journal of Forecasting*, *28*(1), 1-18.
- Wang, W., Rothschild, D., Goel, S., & Gelman, A. (2015). Forecasting elections with non-representative polls. *International Journal of Forecasting*, *31*(3), 980-991.