Forecasting proportional representation elections from non-representative expectation surveys

CrossMark

Andreas Graefe

LMU Munich & Macromedia University, Germany

ARTICLE INFO

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.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Response rates in traditional phone surveys have decreased below 10% in recent years (Kohut et al., 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 et al., 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 and Stegmaier, 2011; Murr, 2011, 2015a, 2015b; Rothschild and 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 sub-samples that contained only Democratic or only Republican voters yielded more accurate forecasts than the complete sample of vote intentions (Rothschild and 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.

2. 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.

2.1. 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 (Labor 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 Labor 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 and 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 and 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 chose 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 et al., 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 et al., 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.

2.2. 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