Social Networks and Citizen Election Forecasting: The More Friends the Better

Debra Leiter, University of Missouri – Kansas City
Andreas Murr, University of Oxford
Ericka Rascon, Middlesex University
Mary Stegmaier, University of Missouri

ABSTRACT
Most citizens correctly forecast which party will win in most elections, usually with greater accuracy than voter intention polls. How do they do it? We argue that social networks are a big part of the answer: much of what we know as citizens comes from our communication with others. Previous research has considered only indirect characteristics of social networks to analyze why citizens are good forecasters. Using a unique German survey, we consider direct measures of social networks to explore their role in election forecasting. We find that three network characteristics – size, political composition, and frequency of political discussion – are among the most important variables when predicting the accuracy of citizens' election forecasts.

KEYWORDS Social networks, election forecasting, citizen forecasting, public opinion, political interest, expectations, Germany

Social Networks and Citizen Election Forecasting:
The More Friends the Better

In most elections, the majority of citizens are able to correctly predict the election winner, regardless of who they plan to vote for (Lewis-Beck and Skalaban 1989; Lewis-Beck and Tien 1999; Miller et al. 2012; Murr 2011, 2015, 2016). Most US citizens typically predict not only which presidential candidate will win their state, but also who will win the presidency (e.g., Graefe 2014); most British citizens are usually correct about which party will win their constituency and which will garner a parliamentary majority (e.g., Lewis-Beck and Stegmaier 2011; Murr 2016). How do they do it?

A small body of work suggests that social networks are a big part of the answer. Much of what we know as citizens comes from our social networks (e.g., Huckfeldt and Sprague 1995), and so we base our election prediction – like so many of our beliefs – on information from people in our network (Uhlaner and Grofman 1986; Lewis-Beck and Tien 1999; Meffert et al. 2011). However, previous studies on social networks and citizen forecasting accuracy have been hampered by the lack of direct measures of social network characteristics, and instead relied on indirect or proxy measures. For example, Lewis-Beck and Tien (1999) find that people with higher levels of education are better able to predict who will win. This is likely because people with higher education levels have developed skills to acquire and process information. They also intimate that level of education tells us something about the size of a person's network, with more educated individuals possessing a larger network. Uhlaner and Grofman (1986) and Meffert et al. (2011) use electoral differences between the citizen's electoral district and the national level to indirectly capture the network's partisan composition, because the surveys they use do not collect measures of social network party leanings. Yet these indirect measures may miss important aspects of the role of social networks on citizen forecasting.

In this study, we use direct measures of network size and composition, along with other network characteristics, in order to build a more complete model of citizen forecasting. Using a unique cross-sectional survey that collected both citizen election forecasts and direct measures of several social network characteristics in Germany in the autumn of 1990, we demonstrate that social networks are as predictive of citizen forecasting accuracy as the most important predictors identified by previous research: vote intention and political interest. In addition, we show which social network characteristics have predictive power – size, political composition, and frequency of discussion – and which ones do not – heterogeneity and level of expertise – in influencing election forecasts. We additionally provide guidance for future surveys on what network measure to include for improving the accuracy of citizen election forecasts. Using a cross-validation exercise we demonstrate that a single, abbreviated measure of network size improves out-of-sample predictions.

WHY CITIZEN FORECASTS?

As the field of election forecasting has grown, scholars have experimented with different measures and methods to find the most accurate predictors (for reviews, see Stegmaier and Norpoth 2017; Lewis-Beck and Stegmaier 2014). Often, such models include vote intention or government approval ratings a few months prior to the election as a gauge of the electorate's preferences.[1] Such variables are found in models of elections in the US (Campbell 2016; Erikson and Wlezien 2016), Britain (Ford et al. 2016; Stegmaier and Williams 2016) and Germany (Norpoth and Gschwend 2017; Jérôme et al. 2017) among others. Both the approval and vote intention items reflect the respondent's personal assessment of the incumbent government or the candidates. However, a developing branch of the election forecasting literature has begun to utilize electoral expectations measured by the question "who do you think will win the election?" This approach is referred to as "citizen forecasting" and has been used for election prediction in the US (Lewis-Beck and Skalaban 1989; Lewis-Beck and Tien 1999; Graefe 2014; Murr 2015) and Britain (Lewis-Beck and Stegmaier 2011, Murr 2011, 2016).

In the citizen forecasting models, survey responses are aggregated to the level of prediction – at the national or constituency level. And most often, citizens get it right. For instance, in their pioneering study, Lewis-Beck and

[1] In addition to voting intention polls or approval ratings, such models often include economic performance measures, number of terms the party has held office, and previous election results.

Skalaban (1989) looked at citizen forecasts of eight US presidential elections between 1956 and 1984. They found that on average 69 percent of citizens correctly forecast the election winner, and that the majority of citizens correctly forecasted 75 percent (6 of 8) of the elections. In other words, moving from individual to aggregate forecasts improved the accuracy from 69 to 75 percent – an increase of 6 percentage points. Their two main findings – that most citizens correctly forecast most of the time, and that groups forecast better than individuals – have subsequently been replicated at different levels (subnational and national) and in different countries (Britain and United States) (e.g., Graefe 2014, Lewis-Beck and Stegmaier 2011; Murr 2011, 2015, 2016).

In addition to demonstrating that citizen forecasts are accurate, several studies have shown that citizen forecasts are more accurate than any other forecasting approach, including voter intention polls. Using national-level data from the last 100 days before the seven US presidential elections between 1988 and 2012, Graefe (2014) compared the relative accuracy of citizen forecasts, voter intentions, prediction markets, expert surveys, and quantitative models. He found that citizen forecasts are better than any other approach at forecasting both election winners and vote shares. Similarly, using national-level data from the 48 months before 18 British general elections between 1950 and 2015, Murr et al. (2016) compared the relative accuracy of citizen forecasts and voter intentions. They found that citizen forecasts are better than voter intentions at forecasting both the winning party and its seat share.

As Murr (2015) has shown, the accuracy of citizen forecasts can even be increased by optimally weighting and delegating the individual forecasts based on the citizens' competence (e.g., Grofman 1975; Kazmann 1973; Shapley and Grofman 1984). The method proceeds in two steps: first, predict the probability that a citizen will correctly forecast; then, delegate the forecasting to the most competent citizen and weight their forecasts by their level of competence. Using data from eleven US presidential elections between 1952 and 2012, Murr (2015) showed that doing so increases the forecasting accuracy of both the candidates' vote shares in states and of which candidate will carry the state. Therefore, being able to predict the chance that a citizen will correctly forecast the election is crucial for improving forecasting accuracy.

WHY CAN CITIZENS FORECAST CORRECTLY?

The explanation of why citizen forecasts are accurate has two parts (Murr 2017). The first part explains why groups forecast better than individuals. This explanation rests on the assumption that individuals forecast better than chance on average, and the second part of the explanation rests on why individuals are able to do so.

Murr (2011) explains the fact that groups predict better than individuals with Condorcet's jury theorem and its generalizations (Condorcet 1785). Condorcet proves under which conditions group decisions reached by plurality rule are better, equal, or worse than individual decisions. His proof assumes that the group faces one correct and one incorrect alternative, that the *k* group members vote independently of one another, and that each member has one vote and the same probability *p* of choosing the correct alternative. Then the probability of a correct group decision by majority vote is

$$P = \sum_{m=\lfloor k/2 \rfloor+1}^{k} \binom{k}{m} p^m (1-p)^{k-m}.$$

He shows that if each member chooses the correct alternative with more than 50 percent probability, then as the group size increases to infinity, the probability of a correct group decision approaches unity ("wisdom of crowds"). He also shows that if each member chooses the correct alternative with less than 50 percent probability, then as the group size increases to infinity, the probability of a correct group decision approaches zero ("folly of crowds").

Although Condorcet's jury theorem refers to group sizes approaching infinity, even small groups show the effect of aggregating individual choices. Consider a group of three independent members each with a probability of choosing the correct alternative of 0.6. This group chooses the correct alternative using majority vote if at least two out of three members vote correctly. Using the above formula, the probability of a correct group decision is $P = 3 \times 0.6^2 \times 0.4 + 0.6^3 = 0.648$, an increase in accuracy of about 5 percentage points. This probability increases as the group size increases: with five independent members this probability is 0.6824, with seven it is 0.7102, with nine it is 0.7334, and so on. In other words, even though individually members may only be slightly better than chance in getting it right, collectively they may choose the correct alternative with almost certainty, if the group has enough members. Table 1 displays the probabilities of a correct group decision for different individual probabilities of getting it right (p = 0.6, 0.7, 0.8, and 0.9) as well as different group sizes (k = 3, 5, 7, and 9).

Table 1. The probability of a correct majority vote of *k* members with individual probability of getting it right *p*.

	$k = 3$	$k = 5$	$k = 7$	$k = 9$
$p = 0.6$	0.6480	0.6826	0.7102	0.7334
$p = 0.7$	0.7840	0.8369	0.8740	0.9012
$p = 0.8$	0.8960	0.9421	0.9667	0.9804
$p = 0.9$	0.9720	0.9914	0.9973	0.9991

In deriving the theorem, Condorcet made three assumptions: each member chooses between only two alternatives, votes independently of the others, and has the same probability of voting correctly. Since the publication of his theorem, several other authors have relaxed each of these assumptions and generalized the theorem accordingly. The theorem still holds even with more than two alternatives (List and Goodin 2001). This is important because in many elections voters choose between more than two parties. Further, Ladha (1992) generalizes the theorem to correlated votes. This is important because citizens might share the same information, talk to each other, or tend to "groupthink" (e.g., Janis 1982). Finally, Grofman et al. (1983) prove that the theorem still holds if members differ in their chance of getting it right as long as they are better than chance on average. This is important because Lewis-Beck and Skalaban (1989) show that citizens differ in their probability of making a correct forecast. In sum, these generalizations make the theorem useful for explaining why groups of citizens predict better than individuals.

Because the explanation of why groups predict better than individuals rests on the fact that individuals predict better than chance on average, the next step is to explain why they should do so. Murr (2017) explains the fact that individuals predict better than chance with Uhlaner and Grofman's Contact Model (Uhlaner and Grofman 1986). Echoing Condorcet's jury theorem, the Contact Model proves under which conditions a citizen's forecast, reached by choosing the party supported by the plurality of information available to the citizen, is better or equal or worse than chance. The proof assumes that the citizen forecasts a two-party election, that she receives and accepts information from the environment independently of one another, and that she counts each piece of information equally.

The Contact Model implies that if a citizen receives and accepts only information consistent with her vote intention ("selective sampling"), then citizen forecasts will always be better than chance on average, though always as informative as voter intentions. However, if a citizen receives and accepts information representative of the public's voter intentions ("random sampling"), then citizens will always be better than both chance and voter intentions on average. As the number of randomly sampled bits of information increases to infinity, the probability of a correct forecast approaches unity. In other words, as soon as citizens receive and accept at least some information that is representative of the public's vote intention, they will do better than chance and voter intentions, which indeed they do (e.g., Lewis-Beck and Skalaban 1989; Graefe 2014).

Because much of what we know as citizens comes from interpersonal communication, we argue that citizens' social networks predict their election forecast. And that the network offers the representative information necessary to forecast better than chance.

SOCIAL NETWORKS AND CITIZEN FORECASTS

The study of social networks—the social context through which individuals are tied to others—has shed light on how and to what extent friends, family, neighbors, and peers influence both electoral belief formation and voting behavior. Along with learning from previous cohorts and personal experience (Manski 2004, Blais and Bodet 2006) and the media (Entman 1989), networks provide contextual information to allow voters to form expectations about elections and influence their choices. Meffert et al. (2011), for example, analyze various factors that influence electoral expectations, such as political motivations (knowledge and interest), rational and strategic considerations (perceived distance between parties), and social context (regional differences as proxy of personal networks) and how these expectations influence voting behavior. The authors find that voters can form reasonable expectations about the winning party and that these beliefs are used to cast "fairly sophisticated votes", such as strategic coalition voting.

Complementarily, Pattie and Johnston (1999) have shown that conversations with partisan discussants influence vote decisions and these can even lead citizens to switch their vote to another party. Similarly, Huckfeldt and Sprague (1991) show that vote preferences are not only determined by the voter characteristics, but also by their discussant partners' characteristics and political preferences. And Nickerson (2008) provides evidence about the influence of couples on voting behavior. Other studies have shown that variations in the composition and size of an individual's network affect political attitudes and the amount of political information. These in turn affect their behavior and their beliefs (Huckfeldt 2007; Mutz 1998; Huckfeldt and Mendez 2008; Partheymüller and Schmitt-Beck 2012; Pietryka 2015).

But how do people form electoral expectations? Citizens may gather information and update their beliefs about electoral victories based on: 1) their network members' characteristics, through observing how their members behave and think about political, social and economic matters, 2) direct information from their network by discussing who they think will win the election and which party they support, 3) previous electoral experiences, and 4) the news and opinion polls.

The nature of social networks makes this source of information more likely to influence citizen electoral expectations and behavior than other sources such as the news media or polls. For instance, Schmitt-Beck and Mackenrodt (2010) show that personal communication appears more influential than mass communication on turnout in a German local election. Despite that the media and polls may provide more reliable and balanced information about the electoral environment than social networks, information from social networks may provide more personalized information by using language and terms that are more familiar and closer to the local context.

While obtaining information from the news media and polls is a passive source of information, social networks give citizens the chance to actively disagree with dissonant information and to learn from it by debating with network members. Hence, all sources of information might be complementary, but social networks provide the citizen the opportunity to engage in back-and-forth debate and to learn from disagreement. As suggested by McClurg (2006), social networks can encourage higher levels of political involvement, as well as increased openness toward differing viewpoints. In other words, people can learn from their networks.

The magnitude of the network's influence on citizens' beliefs about who will win the election may depend on the network's size, frequency of political discussion, political expertise and composition (heterogeneity), along with additional sources of political information.[2] Citizens embedded in larger social networks may have an advantage in forecasting elections, as they frequently have higher levels of political knowledge (Kwak et al. 2005). In addition, the larger the social network, the more likely it is that the network will reflect the vote intentions of the population, making the aforementioned indirect inference more accurate (Banerjee and Fudenberg 2004)[3].

Citizens without a network (*isolated* citizens) may form their beliefs about who will win the election based on media or poll information, as well as on their own electoral preferences. Yet when these citizens incorrectly perceive who will win the election, they lack the social contextual pressure or ability to update their expectations. In contrast, citizens embedded in networks with initially wrong or uncertain beliefs may retrieve information from their network and revise their expectations using information about their network's voting preferences (Chandra 2009).

[2] Similarly, Millner and Ollivier (2016) discuss three main factors that determine the public's beliefs in the context of environmental policies: individual inference (how updating of beliefs takes place), social learning and media.

[3] This is true only if the most important agent's influence diminishes as the number of network members increases (Golub and Jackson 2010).

Having large networks may influence beliefs and behavior, but the information citizens obtain from them should be frequently updated. The more political discussions citizens have with their network, the more information they collect from its members and the more they are able to remember it. Additional information may also make the network's information more salient than the citizen's own information when this brings new information to the citizen. Moreover, the increased frequency of discussion encourages citizens to become more informed and, therefore, it improves their ability to forecast (Eveland 2004; Eveland and Hively 2009).

Both informed and uninformed citizens use networks to gather information about the political system and elections (Pietryka 2015). They seek out political experts, even if they do not share the same partisan affiliation, to help evaluate an election. Citizens are more likely to be influenced by those they perceive as having expertise (Huckfeldt and Mendez 2008, Ryan 2011 and Ahn, Huckfeldt, and Ryan 2014; Huckfeldt, Pietryka, and Reilly 2014) than by non-experts. Thus, these experts within the network should help improve citizens' accuracy in forecasting by providing accurate, if still biased, information. Political expertise can also help to recognize dissonant information and reject it (McClurg 2006).

In general, social networks play a role in both disseminating information and acquiring information that reduces ambiguity (Manski 2004; Ahn, Huckfeldt, and Ryan 2014; Eveland and Hively 2009; Finkel and Smith 2011). However, in some cases, information acquired from social networks may decrease the likelihood of a correct election prediction. When the political network leans toward the losing parties, or when the citizen is unsure of how network members will vote, this will undermine the citizen's ability to offer an accurate election prediction. Those embedded in homogeneous networks may assume greater support for a political party than in fact exists. Homogeneous networks may also reinforce "wishful thinking" and therefore, citizens belonging to these networks may overestimate the chances of victory of a party that presents little chance of success.

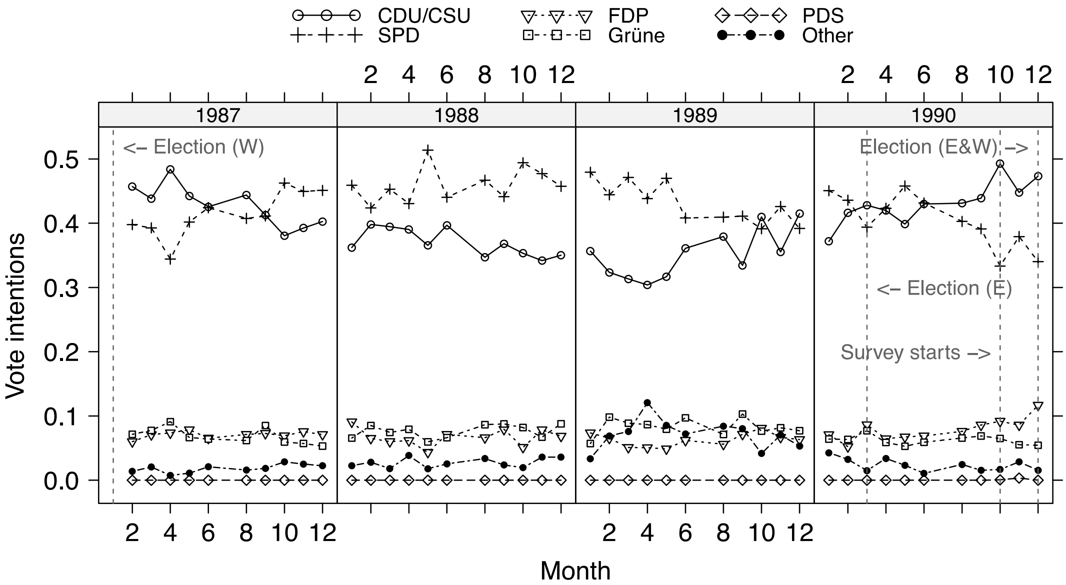
While political disagreement in networks persists even in multiparty electorates (Huckfeldt, Ikeda, and Pappi 2005; Huckfeldt and Johnson 2004), individuals frequently find themselves in social networks with like-minded others. Homogeneity (homophily) of the network may increase or decrease the likelihood of a successful forecast. Individuals in heterogeneous networks tend to show higher levels of political knowledge, as they frequently seek out additional information when they interact with those who do not share their views, which should improve an electoral forecast (Eveland and Hively 2009). However, to the extent that individuals rely on their networks to act as representative samples, more homogenous networks, particularly those allied with an unlikely winner, will decrease the likelihood of a correct forecast. Thus, in such case, the inclusion of more people into a person's network will not add new information. As such, social networks may improve the ability of citizens to make accurate electoral forecasts, but it is dependent on the size and composition of these networks.

DATA AND MEASURES

The 1990 German federal election offers a unique electoral context to examine how social networks predict citizens' ability to forecast the election, as it provides a direct comparison between citizens with long-term democratic experience (West Germans) and citizens new to democratic elections (East Germans), without varying the institutional or electoral context. West Germany held its first democratic election on 14 August 1949, whereas East Germany held its first democratic election on 18 March 1990. The December 2, 1990 Bundestag federal election was the first Federal Republic of Germany election for East Germans, who had voted only four months earlier to unify with West Germany.

After its electoral victory in January 1987, the governing Christian Democratic Union (CDU) had been losing support (Figure 1). This loss benefited the main opposition party, the Social Democratic Party (SPD), which then led the polls from October 1987 to September 1989. However, the CDU started to recover midway through the electoral cycle and led for the first time again in October 1989, starting a period of uncertainty about whether the CDU or the SPD would win in the subsequent election. From March 1990 onward, it looked increasingly likely that the CDU would be victorious in December. They won the East German general election in March, leading the SPD by 19 percentage points. In April Oskar Lafontaine, the candidate of the SPD, fell victim to an assassination attempt and was unable to campaign for three months. From August onwards, opinion polls showed the CDU in the lead, in large part due to the public perception that the CDU was the party best able to handle the economic consequences of unification (Pulzer 1991). However, even though the outcome was fairly certain, as we discuss in the next section, not everyone correctly forecasted a CDU win.

Figure 1: Voting intentions, 1987–1990.



Source: Forschungsgruppe Wahlen (2017).

To examine how social networks predict the ability of citizens to forecast, we use the 1990 German section of the Comparative National Elections Project, a cross-national survey that collects both traditional individual-level data, as well as information on the respondents' media, organizational, and, most importantly for this project, social network characteristics (Gunther et al. 2015; Gunther, Puhle, and Montero 2007). The German section of this survey relies on face-to-face interviews in the pre-election period (October and November 1990) and includes a network battery that asked respondents to name up to five people with whom they discuss important matters. Our sample includes a total of 1547 respondents, 487 of whom are from East Germany. This survey, to our knowledge, uniquely provides both information on the extensiveness and character of respondent's social networks and the respondent's electoral forecast.

To measure the ability of citizens to correctly forecast the winner of the election, we rely on a survey item that asks respondents whether they believe that a CDU-led government or an SPD-led government was likely to win the election, or if they did not know[4]. Based on previous literature, we code all respondents who predict a CDU victory as correct forecasters, and all other respondents as incorrect. The majority of respondents correctly forecasted the winner; however, approximately 25% of West Germans and 18% of East Germans had incorrect forecasts about the election. It is notable, here, that despite their limited experience with democratic elections, the East Germans were better forecasters than the West Germans.

To differentiate between uncertain and inaccurate answers, we create a categorical variable, where those who answer SPD are treated as inaccurate, those who respond with 'don't know' are uncertain, and correct CDU forecasts are treated as the reference category. While the proportion of inaccurate forecasts is similar between East and West Germans, 9.9% and 9.5% forecasted an SPD victory respectively, more than 15.7% of West Germans were uncertain about the election outcome compared to only 8.9% of East Germans.

To test how social networks predict the accuracy of election forecasts, we examine four network characteristics: network size, frequency of political discussion in the network, political expertise in the network, and network ideology (heterogeneity). Network size ranges from 1 to 5[5], and is based on how many discussants the respondent named in the network battery[6].

---

[4] Question wording can be found in the appendix.

[5] We exclude respondents without a discussant because for them the other network characteristics cannot be calculated.

[6] Subsequent to the creation of this survey in 1990, there has been growingly scholarly discussion about network size generators. Although Mardsen (2003) demonstrates that less than 10% of respondents generate more than 5 names, and Merluzzi and Bert (2013) provide evidence suggesting five is a cost effective number of network responses, Eveland, Hutchens, and Morey (2013) argue that the type of name generator used in this survey consistently underestimates network size. Given our theoretical expectation, however, we argue that this underestimation provides a conservative test for

Frequency of political discussion measures how often, on average, the respondent discusses political matters with members of the network, based on respondent evaluation ranging from always to never (Network Discussion). Network expertise is based on the average evaluation of each network member's level of political knowledge. Network ideology is measured as the proportion of the network that the respondent believes will vote for a left leaning party (Network Left), and the proportion of the network for whom the respondent does not know the political party preference (Network Unknown)[7]. Finally, network heterogeneity is operationalized as one minus the absolute difference between the proportions of left and right leaning members in the respondent's network.[8] While network size, frequency of discussion, network expertise, and network heterogeneity may be expected to improve the ability of the respondent to correctly forecast the outcome of the election, network ideology, particularly left-leaning networks, may decrease the likelihood of a correct election forecast – as suggested in the previous section. Table 2 displays summary statistics of the network variables.

Table 2: Summary Statistics of Network Variables.

	West Germans				East Germans			
	Average	SD	Min	Max	Average	SD	Min	Max
Network Size	2.46	1.21	1	5	2.67	1.24	1	5
Network Discussion	1.64	0.80	0	3	2.39	0.63	0	3
Network Expertise	1.08	0.54	0	2	1.27	0.50	0	2
Network Left (Proportion)	0.31	0.40	0	1	0.27	0.37	0	1
Network Unknown (Proportion)	0.29	0.41	0	1	0.28	0.40	0	1
Network Heterogeneity	0.42	0.43	0	1	0.43	0.43	0	1

our hypotheses. In addition, summary network measures cannot measure network characteristics besides size (Eveland, Hutchens, and Morey 2013).

[7] While there are some concerns of projection effects in using respondents' evaluation of their discussion partner's party preference, previous research has demonstrated that voters are surprisingly accurate in identifying their discussion partners' political preferences (Huckfeldt and Sprague 1995).

[8] The proportion of right leaning members is one minus the proportion of left leaning members and the proportion of members for whom the respondent does not know the political party preference. Respondents with equal proportions of left and right leaning members in the network reach the highest value of one on the measure, indicating complete heterogeneity, while respondents with network members of only one ideological direction reach the lowest value of zero on this measure, indicating complete homogeneity. Respondents with an ideologically mixed network reach a value between these two extremes.

In addition to these network effects, we consider other factors that previous studies suggest might predict the accuracy of the forecast (e.g., Lewis-Beck and Tien 1999; Meffert et al. 2011): political, media, and demographic factors, as well as how many days before the election the survey interview took place. To capture individual partisanship and the effects of 'wishful thinking', we create three dummy variables based on the respondent's reported vote intention on the second ballot, including SPD Voters, CDU Voters, and voters uncertain about how they will vote, with minor party supporters treated as the referent category.[9] We also control for self-reported levels of political interest, attention to television news, and attention to news in newspapers. The sociodemographic measures we include are gender, age (transformed into four quartiles), and education (transformed into three categories). Finally, since the survey was conducted over a number of weeks, we account for the number of days before the election that the respondent was surveyed.

Because we argue that social networks provide citizens with information to forecast correctly, it is instructive to examine how our network measures differ from other measures related to information such as formal education, political interest, and media attention (TV and print). To measure how network characteristics relate to these other informational measures, we calculated Pearson's correlation coefficient *r* (Table 3). While there is an association between network characteristics and political education, interest, and media attention, it is very weak ($|r| < .20$) or weak ($.20 \leq |r| < .40$) most of the time. This means that while many people have personal characteristics (e.g., low political interest) that might make an accurate forecast less likely, they nevertheless have social network characteristics (e.g., many discussants) that might make an accurate forecast more likely. In other words, for many citizens their social network can potentially compensate for the lack of information from the media, while for others it can also correct or complement the media information they receive. The weak correlation between network characteristics

9 Germany uses a mixed member proportional electoral system, which provides voters with the opportunity to cast a candidate vote (first ballot) and party vote (second ballot) for the Bundestag, with the party vote determining the overall share of seats in the legislature. This measure of vote intention creates the most comparable measure between East and West Germany, as partisanship was not asked of East German respondents.

and political interest, education, as well as media attention, together with our theoretical arguments, justifies considering network characteristics as additional predictors of forecasting accuracy.

Table 3: Correlation between network characteristics and education, political interest, and media attention.

	Education	Political Interest	TV News Attention	Print News Attention
Network Size	0.20	0.18	0.12	0.12
Network Frequency	0.33	0.46	0.43	0.31
Network Expertise	0.22	0.34	0.28	0.24
Network Left	0.01	0.09	0.03	0.04
Network Unknown	-0.05	-0.13	-0.11	-0.10
Network Heterogeneity	-0.01	-0.06	-0.06	-0.06

The regression analyses reported below weight the respondents by inverse sampling probability in East and West, because East Germans where oversampled relative to their population proportion, and cluster the standard errors by sampling point.

RESULTS

Correct and incorrect forecasts

First, we examine the variables that predict the accuracy of Germans' election forecasts. The outcome in the logit model shown in Table 4 is whether the respondent correctly forecasted the CDU victory or not. In this analysis, the incorrect forecasts include responses that the SPD would win as well as "don't knows". While we are most interested in the difference in forecasting accuracy between respondents with different social network characteristics, looking at other variables that could predict forecast accuracy enables us to compare these results to the handful of other studies that have looked at the characteristics of accurate forecasters.

The results of the binary logit model in Table 4 indicate that social networks predict forecast accuracy in ways consistent with our expectations, even when controlling for a host of other political, media, and demographic characteristics.[10] We observe that both the number of people in the respondent's network and the frequency of political discussion have positive and statistically significant coefficients. This means that both more people in the network and more frequent discussions in the network correspond to a positive difference in probability of a correct forecast. Conversely, we observe that the share of the network with left or unknown political leanings have negative and statistical significant coefficients. This means that the larger the share of the network with left or unknown political leanings, the less likely the respondent's forecast will be correct. The coefficients of both network expertise and network heterogeneity are in the expected positive direction, but miss conventional levels of statistical significance.

Table 4: Correct Forecast of CDU victory
Pooled Binary Logit Estimates
Log-Odds

	Estimate	(Std. Error)
Constant	-0.89	(0.65)
East	-0.09	(0.19)
Age	0.09	(0.06)
Female	-0.01	(0.13)
Education	0.06	(0.13)
Political Interest	0.29**	(0.09)
TV News Attention	-0.03	(0.09)
Print News Attention	0.03	(0.07)
SPD Voter	0.03	(0.18)
CDU Voter	2.10**	(0.27)
Undecided Voter	0.53**	(0.25)
Days Until Election	-0.01	(0.01)
Network Size	0.22**	(0.07)
Network Discussion	0.19*	(0.11)
Network Expertise	0.24	(0.16)

[10] These and the following computations were performed on a Mac OS X 10.11.6 with Stata/SE 12 using the logit, mlogit, margins, and lincom commands.

Network Left	-0.77**	(0.24)
Network Unknown	-0.59*	(0.35)
Network Heterogeneity	0.15	(0.33)
N	1547	

Note: * p<0.10, ** p<0.05. Standard errors clustered by sampling points. Data weighted by inverse sampling probabilities in East and West.

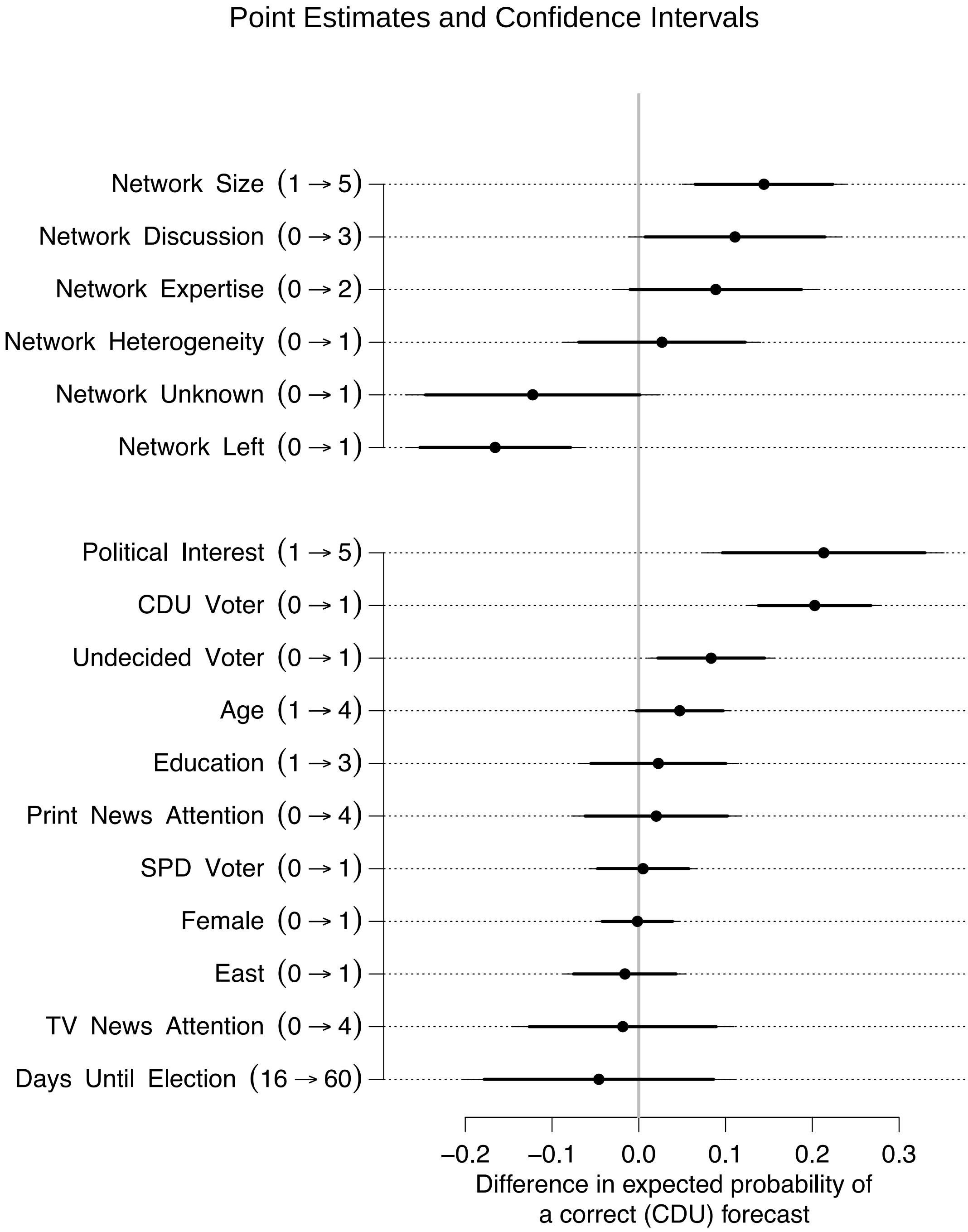
Of the other variables, only a few coefficients attain statistical significance. We corroborate findings from earlier studies that respondents with higher levels of political interest are more likely to make an accurate forecast. And, we find evidence that CDU voters are more likely to correctly forecast a CDU victory relative to the excluded "minor party vote" category. We also observe that respondents who say they don't know whom they will vote for (undecided voters) are also more likely to correctly forecast compared to minor party voters, though the coefficient is smaller than for CDU voters. By contrast, SPD voters are just as likely to get it right or wrong as minor party voters.

Notably, the coefficient of the "East" variable, designed to capture systemic differences between East and West Germans in this pooled analysis, is not statistically significant. Furthermore, the demographic variables, media exposure, and days before the election are not predictive of forecasting accuracy.

To understand the results of the full binary model and the subsequent multinomial logit models we compute first differences (King 1989: 107f). First differences estimate how much the fitted values would differ on average when comparing two respondents that have different levels of a given predictor while being identical in all the other variables. We compute first differences by subtracting the expected probability of an outcome given the maximum value of a predictor from the expected probability given its minimum value, holding all other variables at their median.

Figure 2 provides a visual assessment of the difference in the expected probabilities of a CDU forecast when comparing two respondents who have the minimum and maximum level of a predictor, while holding all the other variables at their medians. The bold line depicts the 90% confidence interval around the point estimate of the difference in expected probability, while the thinner and slightly longer line shows the 95% confidence range. Here, the predictive power of the social network variables is apparent, reinforcing the importance of the network characteristics. Network size and the ideological leanings of the network show large differences in the expected probability of forecast accuracy, differences that are rivaled only by political interest and respondent vote intention for the CDU or not known. For instance, if we compare a respondent who has five network members with someone who has one network member (the maximum and minimum values for network size), we expect to see that the respondent with the larger network has a 15 percentage point higher chance of making a correct forecast on average. As another example, if we compare a respondent whose network consists of only left leaning members with someone whose network consists of no left leaning members, we expect to see that the one with the more left leaning network has a 16 percentage point smaller chance of making a correct forecast on average. (Online Appendix Table A1 provides the difference in expected probabilities and their confidence intervals that correspond to this figure.)

Figure 2: Difference in Expected Probabilities for Pooled Binary Logit Model



Note: Difference in expected probabilities between two respondents with maximum and minimum values of the predictor while holding the other predictors constant at their median value. Predictors are sorted by increasing effect, separately for network characteristics and controls. Bold segments indicate 90% confidence intervals and thin segments indicate 95% confidence intervals.

Correct and incorrect forecasts and the "don't knows"

Next we recognize that "wrong" forecasts are not all the same. A respondent could provide an incorrect forecast of an SPD victory, or the respondent could report not knowing who will win, and the covariates that predict these results are likely to be different. To assess this, we estimate multinomial logit models where those who offer an incorrect (SPD) or uncertain (don't know) response are assessed relative to those who forecasted correctly. We estimate this for the pooled survey, and also in the form of an interactive model where we assess whether differences between East and West Germans exist when it comes to the coefficients of the various predictors.

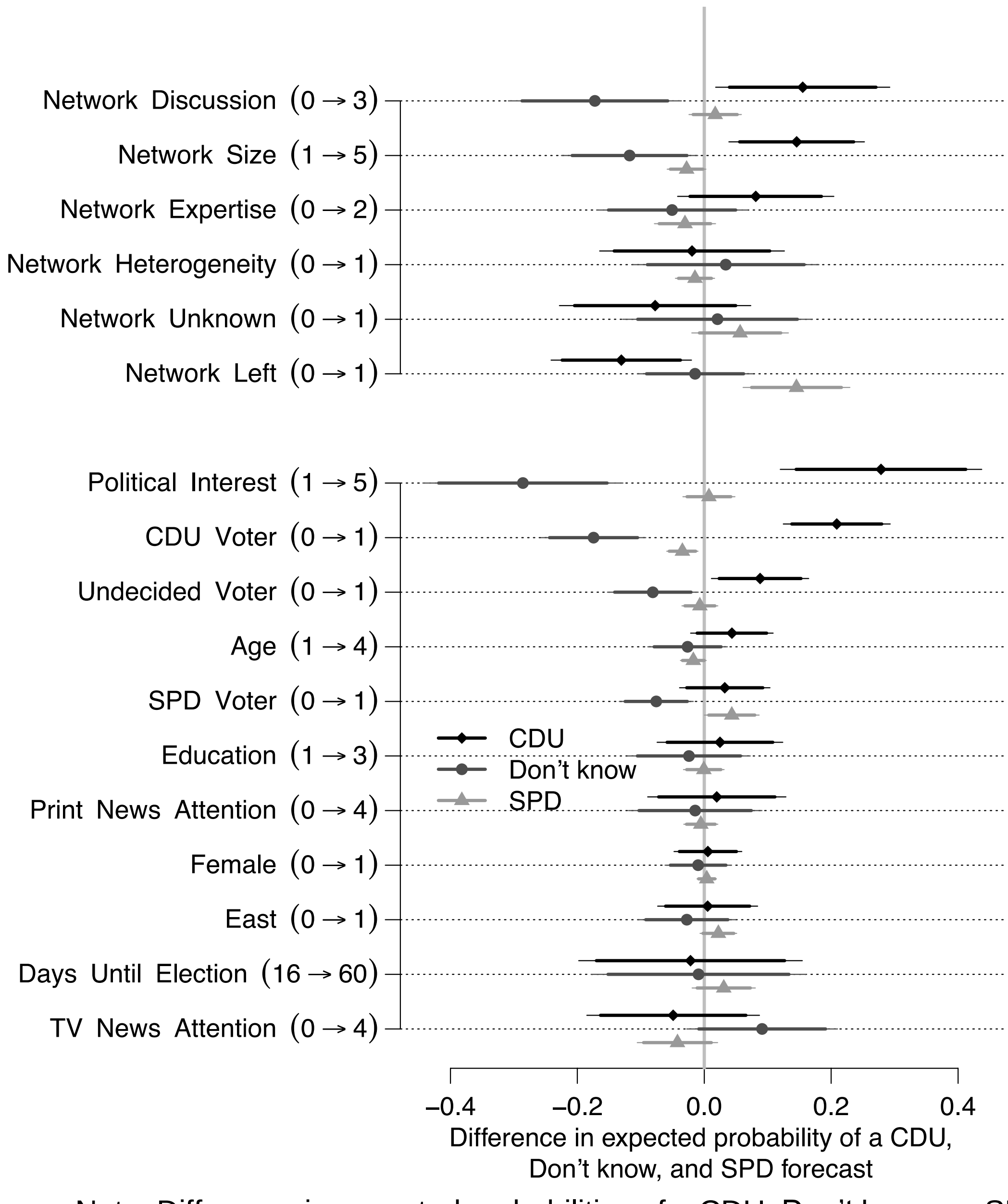
To understand the results of the multinomial logit model, we again compute first differences[11]. Figure 3 presents the difference in expected probabilities and the corresponding confidence intervals for each predictor and for each forecast (CDU, SPD, don't know) based on the estimates of the pooled multinomial logit model (Full results reported Table A2 in the online appendix). Here we observe that the social network variables differ in their predictive power across the three distinct forecasts. In general, we observe that respondents who had a higher share of their network with left or unknown leanings and who have a network with lower levels of expertise were more likely to provide an incorrect SPD forecast. By contrast, respondents who had less frequent discussions with those in their network were more likely to give a "don't know" response. Specifically, if we compare a respondent whose network has five members to someone whose network has one member, we expect that the respondent with the larger network is 14 percentage points more likely to make a correct CDU forecast, 11 percent less likely to give a "don't know" response, and 2 percentage points less likely to make an incorrect SPD forecast. In other words, we expect that respondents with varying network sizes differ in their chances of giving a CDU forecast or "don't know" response, but that they are similar in their chances of giving a SPD forecast on average. In sum, the larger the network, the more accurate and certain citizen forecasts are. (Online

[11] Online Appendix Table A2 reports the full results of our pooled and interactive multinomial logit models.

Appendix Table A3 reports the differences in probabilities and the values for the 95% confidence intervals).

Figure 3: Difference in Expected Probabilities for the Pooled Multinomial Logit Model

Point Estimates and Confidence Intervals



Note: Difference in expected probabilities of a CDU, Don't know, or SPD forecast between two respondents with maximum and minimum values of the predictor while holding the other variables constant at their median value. Predictors are sorted by increasing effect on giving a CDU response, separately for network characteristics and controls. Bold segments indicate 90% confidence intervals and thin segments indicate 95% confidence intervals.

Figure 3 also shows large differences in expected probabilities for respondents who differ in their vote intention and political interest. Comparing two respondents with high and low political interest, we expect that the one who is more interested in politics has a 27 percentage points higher chance of correctly forecasting the CDU to win, a 27 percentage points lower chance of a “don’t know” response, but does not differ in the probability of an incorrect SPD forecast on average.[12]

So far, in the binary and multinomial logit models and in the difference in expected probability figures, we have demonstrated that social network characteristics are highly predictive of the accuracy of an election forecast, and they help us distinguish between incorrect forecasts and respondent uncertainty. These network measures, in addition to political interest and vote intentions, by far outperform demographics and media variables. The number of days before the election that the interview took place is not predictive of the type of prediction given by the respondent.

Allowing the coefficients to vary between East and West Germans

German reunification ended 40 years of political division between East and West Germany. It has been of general interest to describe the similarities and differences in public opinion and behavior between East and West Germans in order to understand the extent to which the country has developed a unified political culture (e.g., Gabriel 1997; van Deth, Rattinger, and Roller 2000; Fuchs, Roller, and Wessels 2002; Gabriel, Falter, and Rattinger 2005; Falter et al. 2006). In our context, we expect East Germans to rely more on social network information than West Germans given the challenges that new democracies are likely to be subjected to, such as weak partisan cues, low levels of partisan identification, and volatile voters (Baker, Ames and Renho

[12] While we cannot reject the null hypothesis that the first difference for political interest is the same for network size related to a CDU response (b=0.13; Std. Err.=0.09; z=1.54) and a SPD response (b=0.04, Std. Err.=0.03; z=1.34), we can reject the null hypothesis for a Don't Know response (b=-0.17; Std. Err.=0.08; z=-1.98).

2006). Hence, now we examine whether the coefficients of our predictors differ in the East and West.

To examine possible heterogeneous coefficients between East and West, we follow the recommendations of Tsai and Gill (2013) on interactions in generalized linear models. We first add to the pooled multinomial logit regression equation product terms between each of the predictors and the East dummy variable. (The last two columns of Table A2 display the estimates of this interacted multinomial logit model.) We then calculate first differences of the predictors, separately for East and West. And, finally, we compare the first differences of a predictor across East and West to assess the statistical significance and magnitude of the interaction. (Online Appendix Figure A1 and Table A4 show all of these first differences.)

By following this procedure, we found statistically significant interactions for only two network variables (the size of the network and the share of the network with left political leanings) on just one outcome ("don't know"). In other words, of the 18 possible interactions – six network variables multiplied by three outcomes – 16 are statistically insignificant. Because with 20 such comparisons we would expect 1 of them to be statistically significant by chance, we do not want to emphasize the differences that we found. The results of the interacted model suggest, therefore, that there are no major differences in how network characteristics predict forecast accuracy between East and West Germans. For both groups social networks predict forecast accuracy in the same way.[13]

A Simple Network Measure for Improving Accuracy of Out-of-sample Predictions

The above analysis described which citizens were more likely than others to correctly forecast the election. Next, we would like to provide

[13] We also considered possible interactions between the most important predictors (Gelman and Hill 2007: 69): network size, network discussion, and network left as well as political interest and vote intention. We tested whether the network variables interact with each other, and whether they interact with the other predictors following again the procedure recommended by Tsai and Gill (2013). (In the online appendix, Tables A5, A6 and A8 show the estimated regression models while the Tables A7 and A9 as well as Figures A2 to A6 show the first differences.) We found one statistically significant interaction: the importance of the frequency of discussion decreases with higher levels of political interest for the outcomes CDU and Don't Know.

guidance for people who want to use citizen forecasts to forecast future election outcomes. As mentioned before, aggregated citizen forecasts are most accurate when weighting individual forecasters by their forecasting competence. The above analysis improves the researcher's ability to identify which individuals to weight more heavily: because social network characteristics predict forecasting competence, future aggregated citizen forecasts will be more accurate when they use these network characteristics to calculate the individual weights.

However, network batteries take a great deal of space on a questionnaire. The survey we used in our analysis included five questions identifying network members plus follow-up items for each identified member measuring their political preference, expertise, frequency of discussion, etc. Is including network batteries in new surveys worth it in terms of improving election forecasting accuracy? Below, we show that even a single, abbreviated measure of network size – asking citizens with how many people they discussed an important personal matter – improves out-of-sample predictions.

We compared the out-of-sample predictive accuracy of all possible subsets of the predictors considered above, with three modifications. First, as the response variable, we chose whether the citizen correctly forecasted the winner (0 = "no"; 1 = "yes"). We excluded the response "don't know" because only actual forecasts can be weighted. Second, as the only network characteristic, we considered network size (0 = "no discussants" to 5 = "five discussants"). We do so because the above descriptive analysis found that it strongly correlated with forecasting accuracy, and because this predictor also applies to citizens without a discussant, while the other network characteristics apply only to citizens with at least one discussant. (Excluding "don't knows" and including citizens without networks changes the number of observations to 1,592.) Finally, we replaced the three vote intention predictors with a single dummy variable indicating whether citizens forecasted the same party to win as the one they intend to vote for (0 = "no"; 1 = "yes"). We do so because this predictor can be used without the researcher knowing in advance which party will win (Murr 2015). This leaves us with ten predictors: East, Age, Female, Education, Political Interest, TV News Attention, Print News Attention, Forecast Intention, Days Until Election, and Network Size.

We used k-fold cross-validation (e.g., Ward 2010, Murr 2015) to compare the out-of-sample predictive accuracy of all 2^10=1,024 possible subsets of predictors. Cross-validation randomly splits the data into k folds. It first fits the models to the k-1 folds and then tests them on the k-th one, iterating these two steps from 1 to k to get a distribution of the predictive accuracy. We set k = 10, which is the typical value in the literature, and repeated k-fold cross-validation with ten different splits. We measured the predictive accuracy with the area under the receiver operator characteristic curve (AUROC), which is a common measure of accuracy in the forecasting literature for binary classification tasks (e.g., Ward 2010, Murr 2015). An AUROC value of 50 per cent indicates a random classifier and a value of 100 per cent indicates an optimal classifier. The AUROC can be interpreted as the probability that a randomly chosen correct citizen forecaster is ranked as more likely to be correct than a randomly chosen incorrect citizen forecaster (Fawcett 2007).

Including network size as a predictor improved the predictive accuracy (Table 5). Overall, the model with the largest AUROC of 62.57 per cent included only five of the nine predictors: Age, TV News Attention, Forecast Intention, Days Until Election, and Network Size. By contrast, the best model excluding network size achieved an AUROC of 61.40 per cent – 1.17 percentage points lower compared to the best model including network size. Averaging across all 1,024 models, the AUROC of models including network size was 1.4 percentage points larger than the AUROC of models excluding network size. By comparison, only Forecast Intention and Age had a larger increase of 3.98 and 3.22 percentage points, respectively. Including some predictors even decreased predictive accuracy on average. For instance, the AUROC of models including Print News Attention was on average 0.17 percentage points lower than the AUROC of models excluding Print News Attention. This all demonstrates that it is worth including network size as a measure on a new survey because it does a better job predicting forecasting competence than many commonly available measures (e.g., print news attention). And as elections grow increasingly competitive and election results grow tighter, even minor improvements on forecasting measurements may play a critical role in increasing forecast accuracy.

Table 5: Out-of-sample accuracy of all possible 1,024 subsets of variables in predicting correct forecasts (0 = "no"; 1 = "yes") of 1,592 citizens before the German Bundestag election in 1990 using binary logistic regression.

	Predictors (0 = "excluded"; 1 = "included")										
Rank	East	Age	Fem.	Educ.	Pol. Int.	TV	Print	Forec. Intent.	Days	Net. Size	AUCOC
1	0	1	0	0	0	1	0	1	1	1	62.57
2	0	1	0	0	1	0	0	1	1	1	62.48
3	0	1	0	0	0	0	0	1	1	1	62.36
4	0	1	0	0	1	0	0	1	0	1	62.32
5	0	1	0	0	0	1	0	1	0	1	62.28
6	0	1	0	0	1	1	0	1	1	1	62.27
7	0	1	0	0	0	1	1	1	1	1	62.21
8	0	1	0	0	0	0	1	1	1	1	62.19
9	0	1	0	1	0	1	0	1	1	1	62.18
10	0	1	0	1	0	0	0	1	1	1	62.18
…											
91	0	1	0	0	0	1	0	1	1	0	61.40
…											

Note: Entries are sorted by decreasing average area under the receiver operator characteristic curve (AUROC) in 10-fold cross-validation across ten repetitions and by decreasing number of predictors. Due to space constraints, only the ten best models are presented, as well as the best model without network size as a predictor.

CONCLUSION

In this study, we have examined how social networks predict the ability of citizens to correctly forecast the election winner when controlling for other variables such as political interest, gender, education, media attention, and vote intention. Specifically, we have found that citizens with larger social networks and who engage in more frequent political discussion are better at forecasting the winner than people who do not share these network characteristics. Our analysis also shows that the political leanings of the network matter. Those whose network is composed of a higher proportion of left-wing party supporters were less likely to correctly forecast that the right-wing CDU would win. Furthermore, respondents who were unsure of their friends' party preferences were less likely to provide a correct forecast. Essentially, those voters with an extensive, communicative, and varied group of friends – and, of course, neighbors, colleagues, family members, and peers – are best able to accurately forecast the election winner.

Finding such robust results for social network characteristics might be a surprise in this particular election when, at the time of the survey in autumn 1990, public opinion polls pointed to a decisive CDU victory. We view this particular election as a conservative test of our social networks theory. With the election all but a forgone conclusion, one might expect the social network's predictive power of the respondents' forecasts to be limited. Yet even in this context, networks demonstrably predicted citizens' forecasts. In more competitive elections, where there is greater uncertainty about the likely winner, social networks and their characteristics would likely play an even more important role in predicting voters' election forecasts.

In addition to examining the predictive power of social network characteristics on election forecasts, we have also considered how experience with democratic elections might predict citizens' ability to give an accurate forecast, based on whether respondents resided in East or West Germany. Perhaps surprisingly, East Germans were more likely to correctly forecast the victor than West Germans. And, while we might have expected that less democratic experience would mean that networks were more important for East than for West Germans in predicting their expectations – given the challenges faced by new democracies (e.g. weak partisan cues, low levels of partisan identification and volatile voters as discussed by Baker, Ames and Renho 2006), our analysis indicates that such differences do not exist.

The robustness of our findings in both East and West Germany suggests that the predictive power of social networks should appear in both new and established democracies. However, since the institutional and political context of the 1990 German election is the same for both regions, future research should examine whether social networks predict citizen forecasts similarly in countries with different party systems and electoral rules.

Future research could study how the internet and the emergence of online social networks have influenced citizens' ability to forecast. Some studies have shown that the internet has neither increased nor decreased social capital, but supplemented it (e.g., Wellmann et al. 2001). Hence, citizens seem to bond (form closer connections with others) and bridge (form ties across social groups) to the same extent as before. Other studies have shown a high overlap of offline and online social networks (e.g., Subrahmanyam et al. 2008), hence, elicited networks using electoral surveys are likely to be a subset of the networks captured in online social networks. Online platforms are likely to increase citizens' ability to forecast because they provide wider access to information without additional cost. They enable citizens to be updated about their networks' electoral preferences without face-to-face discussions, and they allow citizens to be informed about all their network members, even those who are distant from the most influential people in their network.

A final lesson of our analysis is that social network characteristics, and questions on citizen forecasting are important elements in electoral surveys, and that their exclusion may inhibit our understanding of political learning and decision making. Social network size and composition are associated with the ability of citizens to correctly forecast elections, and as the demand for political forecasting continues, understanding how and why citizens correctly estimate the winners of elections is critical. Absent measures of social network characteristics, we cannot fully predict and utilize these forecasts. Additionally, understanding citizen forecasting reveals something important about how social networks predict political learning. Size and ideological make-up of networks compete with other factors to predict whether citizens can make correct inferences about not just local, but also national, political trends. In sum, just as social networks help us understand citizen forecasting, citizen forecasting informs us about how social networks predict contextual learning and political knowledge.

WORKS CITED

Ahn, T. K., Robert Huckfeldt, and John Barry Ryan. 2014. *Experts, Activists, and Democratic Politics: Are Electorates Self-Educating?* Cambridge University Press.

Baker, A., Ames, B. and Renno, L.R., 2006. Social context and campaign volatility in new democracies: networks and neighborhoods in Brazil's 2002 elections. *American Journal of Political Science*, 50(2), pp.382-399.

Banerjee, Abhijit, and Drew Fudenberg. 2004. "Word-of-mouth Learning." *Games and Economic Behavior* 46(1): 1-22.

Blais, A. and Bodet, M.A., 2006. How do voters form expectations about the parties' chances of winning the election?. *Social Science Quarterly*, 87(3), pp.477-493.

Campbell, James E. 2016. "The Trial-heat and the Seats-in-trouble Forecasts of the 2016 Presidential and congressional Elections." *PS: Political Science & Politics* 49: 664-68.

Chandra, Kanchan. 2009. "Why Voters in Patronage Democracies Split Their Tickets: Strategic Voting for Ethnic Parties." *Electoral Studies* 28(1): 21–32.

Condorcet, Marie-Jean-Antoine-Nicolas de Caritat, Marquis de. 1785. *Essai Sur L'application de L'analyse a La Probabilité Des Descisions Redues a La Pluralité de Voix*. Paris: De l'Imprimerie Royale.

van Deth, Jan, Hans Rattinger, and Edeltraud Roller (eds.). 2000. *Die Republik auf dem Weg zur Normalität? Wahlverhalten und politische Einstellungen nach acht Jahren Einheit*. Opladen: Leske + Budrich .

Entman, Robert M. 1989. "How the Media Affect What People Think: An information processing approach." *The Journal of Politics* 51: 347-370.

Erikson, Robert S., and Christopher Wlezien. 2016. "Forecasting the Presidential Vote with Leading Economic Indicators and the Polls." *PS: Political Science & Politics* 46: 669-72.

Eveland, William P. 2004. “The Effect of Political Discussion in Producing Informed Citizens: The Roles of Information, Motivation, and Elaboration.” *Political Communication* 21(2): 177–93.

Eveland, William P, and Myiah Hutchens Hively. 2009. “Political Discussion Frequency, Network Size, and ‘Heterogeneity’ of Discussion as Predictors of Political Knowledge and Participation.” *Journal of Communication* 59(2): 205–24.

Eveland, Jr, William P., Myiah J. Hutchens, and Alyssa C. Morey. 2013. "Political Network Size and Its Antecedents and Consequences." Political Communication 30(3): 371-94.

Falter, Jürgen W., Oscar Gabriel, Hans Rattinger, and Harald Schoen (eds.). 2006. *Sind wir ein Volk? Ost- und Westdeutschland im Vergleich*. Munich: Beck.

Fawcett, Tom. 2006. “An Introduction to ROC Analysis.” Pattern Recognition Letters 27(8): 861–874.

Finkel, Steven E., and Amy Erica Smith. 2011. “Civic Education, Political Discussion, and the Social Transmission of Democratic Knowledge and Values in a New Democracy: Kenya 2002.” *American Journal of Political Science* 55(2): 417–35.

Ford, Robert, Will Jennings, Mark Pickup, and Christopher Wlezien. 2016. “From Polls to Votes to Seats: Forecasting the 2015 British General Election” *Electoral Studies* 41: 244-49.

Forschungsgruppe Wahlen, Mannheim. 2017. *Partial Cumulation of Politbarometers 1977-2015*. GESIS Data Archive, Cologne. ZA2391 Data file Version 7.0.0, doi:10.4232/1.12733

Fuchs, Dieter, Edeltraud Roller, and Bernhard Wessels (eds.). 2002. *Bürger und Demokratie in Ost und West: Studien zur politischen Kultur und zum politischen Prozess*. Wiesbaden: Westdeutscher Verlag.

Gabriel, Oscar W. (ed.). 1997. *Politische Orientierungen und Verhaltensweisen im vereinigten Deutschland.* Opladen: Leske + Budrich.

Gabriel, Oscar W., Jürgen W. Falter, and Hans Rattinger (eds.). 2005. *Wächst zusammen, was zusammengehört? Stabilität und Wandel politischer*

Einstellungen im wiedervereinigten Deutschland. Baden-Baden: Nomos.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Golub, Benjamin, and Matthew O. Jackson. 2010. “Naive Learning in Social Networks and the Wisdom of Crowds.” *American Economic Journal: Microeconomics* 2(1): 112-49.

Graefe, Andreas. 2014. “Accuracy of Vote Expectation Surveys in Forecasting Elections.” *Public Opinion Quarterly* 78 (S1): 204-32.

Grofman, Bernard. 1975. “A Comment on 'Democratic Theory: A Preliminary Mathematical Model.” *Public Choice* 21: 99-103.

Grofman, Bernard, Guillermo Owen, and Scott L. Feld. 1983. “Thirteen Theorems in Search of the Truth, Theory and Decision.” 15 (3): 261–78.

Gunther, Richard, Paul A. Beck, Pedro C. Magalhães, and Alejandro Moreno, eds. 2015. *Voting in Old and New Democracies*. Routledge.

Gunther, Richard, Hans-Jürgen Puhle, and José R. Montero. 2007. *Democracy, Intermediation, and Voting on Four Continents*. Oxford University Press.

Huckfeldt, Robert. 2007. “Unanimity, Discord, and the Communication of Public Opinion.” *American Journal of Political Science* 51(4): 978–95.

Huckfeldt, Robert, and John Sprague. 1991. “Discussant Effects on Vote Choice: Intimacy, Structure, and Interdependence.” *The Journal of Politics* 53(01): 122–58.

Huckfeldt, Robert, Ken ’ichi Ikeda, and Franz Urban Pappi. 2005. “Patterns of Disagreement in Democratic Politics: Comparing Germany, Japan, and the United States.” *American Journal of Political Science* 49(3): 497–514.

Huckfeldt, Robert, and P. E. Johnson. 2004. *Political Disagreement: The Survival of Diverse Opinions within Communication Networks*. Cambridge Univ Pr.

Huckfeldt, Robert, and Jeanette Morehouse Mendez. 2008. “Moths, Flames, and Political Engagement: Managing Disagreement within Communication Networks.” *The Journal of Politics* 70(1): 83–96.

Huckfeldt, Robert, Matthew T. Pietryka, and Jack Reilly. 2014. “Noise, Bias, and Expertise in Political Communication Networks.” *Social Networks* 36: 110–21.

Huckfeldt, Robert, and John Sprague. 1995. *Citizens, Politics and Social Communication: Information and Influence in an Election Campaign*. Cambridge University Press.

Janis, Irving L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Boston: Houghton Mifflin.

Jérôme, Bruno, Véronique Jérôme-Speziari, and Michael S. Lewis-Beck. 2017. “The Grand Coalition Reappointed by Angela Merkel on Borrowed Time.” *PS: Political Science and Politics*. 50: 683-385.

Kazmann, Raphael G. 1973. “Democratic Organization: A Preliminary Mathematical Model.” Public Choice 16(1):17-26.

King, Gary. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Cambridge: Cambridge University Press.

Kwak, Nojin, Ann E. Williams, Xiaoru Wang, and Hoon Lee. 2005. “Talking Politics and Engaging Politics: An Examination of the Interactive Relationships between Structural Features of Political Talk and Discussion Engagement.” *Communication Research* 32(1): 87–111.

Ladha, Krishna K. 1992. “The Condorcet Jury Theorem, Free Speech, and Correlated Votes.” *American Journal of Political Science.* 36(3): 617-34.

Lewis-Beck, Michael S., and Andrew Skalaban. 1989. “Citizen Forecasting: Can Voters See into the Future?” *British Journal of Political Science* 19 (1): 419-27.

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 Mary Stegmaier. 2014. “US Presidential Election Forecasting.” *PS: Political Science and Politics* 47: 284-88.

Lewis-Beck, Michael S., and Charles Tien. 1999. "Voters as Forecasters: A Micromodel of Election Prediction." *International Journal of Forecasting* 15 (2): 175-84.

List, Christian and Robert E. Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9(3): 277–306.

Manski, C.F., 2004. "Social Learning from Private Experiences: The Dynamics of the Selection Problem." *The Review of Economic Studies 71*(2): 443-58.

McClurg, S.D., 2006. The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation. *American Journal of Political Science*, 50(3), pp.737-754.

Meffert, Michael F., Sascha Huber, Thomas Gschwend, and Franz Urban Pappi. 2011. "More than Wishful Thinking: Causes and Consequences of Voters' Electoral Expectations about Parties and Coalitions." *Electoral Studies* 30: 804-15.

Miller, Michael K., Guanchun Wang, Sanjeev R. Kulkarni, H. Vincent Poor, and Daniel N. Osherson. 2012. "Citizen Forecasts of the 2008 U.S. Presidential Election." *Politics & Policy* 40: 1019-52.

Millner, A. and Ollivier, H., 2016. Beliefs, politics, and environmental policy. *Review of Environmental Economics and Policy*, 10(2), pp.226-244.

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: 283-88.

Murr, Andreas E. 2017. '"Wisdom of Crowds"', in Arzheimer, Kai, Evans, Jocelyn, and Lewis-Beck, Michael S. (eds) Handbook of Political Behavior. Sage, 835-860.

Murr, Andreas E., Mary Stegmaier and Michael S. Lewis-Beck. 2016. "Citizen Forecasting v. Pocketbook Forecasting: Are Two Heads Better than One?" Paper presented at the Midwest Political Science Association Conference, April 2016.

Mutz, Diana C. 1998. *Impersonal Influence: How Perceptions of Mass Collectives Affect Political Attitudes*. Cambridge University Press.

Nickerson, David W. 2008. "Is Voting Contagious? Evidence from Two Field Experiments." *American Political Science Review* 102(01): 49-57.

Norpoth, Helmut, and Thomas Gschwend. 2017. "Chancellor Model Predicts a Change of the Guards." *PS: Political Science and Politics* 50: 686-688.

Partheymüller, Julia, and Rüdiger Schmitt-Beck. 2012. "A 'Social Logic' of Demobilization: The Influence of Political Discussants on Electoral Participation at the 2009 German Federal Election." *Journal of Elections, Public Opinion & Parties* 22(4): 457–78.

Pattie, Charles, and Ron Johnston. 1999. "Context, Conversation and Conviction: Social Networks and Voting at the 1992 British General Election." *Political Studies* 47(5): 877–89.

Pietryka, Matthew T. 2015. "Accuracy Motivations, Predispositions, and Social Information in Political Discussion Networks: Accuracy Motivations in Political Discussion." *Political Psychology* 37(3): 367-86.

Pulzer, Peter. 1991. "The German Federal Election of 1990." *Electoral Studies* 10(2): 145–54.

Ryan, John Barry. 2011. "Social Networks as a Shortcut to Correct Voting." *American Journal of Political Science* 55(4): 753–66.

Schmitt-Beck, R. and Mackenrodt, C., 2010. Social networks and mass media as mobilizers and demobilizers: A study of turnout at a German local election. *Electoral studies*, 29(3), pp.392-404.

Shapley, Lloyd, and Bernard Grofman. 1984. "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies." *Public Choice* 43(3): 329–43.

Stegmaier, Mary, and Helmut Norpoth. 2017. "Election Forecasting." In: Valelly, R. (Ed.), Oxford Bibliographies in Political Science. Oxford University Press. Available at http://www.oxfordbibliographies.com/view/document/obo-9780199756223/obo-9780199756223-0023.xml.

Stegmaier, Mary and Laron Williams. 2016. "Forecasting the 2015 British Election Through Party Popularity Functions." *Electoral Studies* 41: 260-63.

Subrahmanyam, K., Reich, S.M., Waechter, N. and Espinoza, G., 2008. Online and offline social networks: Use of social networking sites by emerging adults. *Journal of Applied Developmental Psychology*, 29(6), pp.420-433.

Tsai, Tsung-han and Jeff Gill. 2013. "Interactions in Generalized Linear Models: Theoretical Issues and an Application to Personal Vote-Earning Attributes." *Social Sciences* 2: 91-113.

Uhlaner, Carole J. and Bernard Grofman. 1986. "The Race May Be Close but My Horse is Going to Win: Wish Fulfillment in the 1980 Presidential Election." *Political Behavior* 8(2): 101-29.

Wellman, B., Haase, A.Q., Witte, J. and Hampton, K., 2001. Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist*, 45(3), pp.436-455.

QUESTION WORDING APPENDIX

DEPENDENT VARIABLE

FORECASTING:

From the present point of view: who would you say will win the next general election: The CDU/CSU or a coalition government led by CDU/CSU, or the SDP or a coalition government led by the SDP?

NETWORK VARIABLES

NETWORK SIZE

From time to time, most people discuss important personal matters with other people. Looking back over the last six months, who are the people with whom you discussed an important personal matter?

NETWORK FREQUENCY

When you talk with these persons, how often do you discuss political questions? Would you say almost always, sometimes, seldom, or never?

NETWORK EXPERTISE

How much do these persons, in your opinion, know about politics: much or very much, average, less much? Would you say: much/very much, average, or less much?

NETWORK IDEOLOGY

Which party do you think would these persons vote for in the general election of 2 December this year?

INDIVIDUAL LEVEL

VOTE CHOICE:

Second Vote: Which party will you vote for with your second vote?

POLITICAL INTEREST

Generally Speaking: How much are you interested in politics? Would you say:

very much, much, so-so, somewhat, or not at all?

TV NEWS ATTENTION

How attentively do you follow [television] news reports on political events in Germany and other countries? Would you say: very attentively, attentively, less attentively, or not attentively at all?

PRINT NEWS ATTENTION

Regardless of how often you read your daily newspaper: How attentively do you read the reports on the political events in Germany and other countries? Would you say: very attentively, attentively, less attentively, or not attentively at all?

EDUCATION

What education level do you have?

AGE

Please tell me what month and year you were born

GENDER

Sex of Respondent: Man or Woman.