Sources of Polling Accuracy: How Election Context Affects the Polls *

Jacob Sohlberg † J. Alexander Branham ‡

Abstract

Although polling accuracy increases throughout the election, polls are always at least a little wrong on election day. In this article, we attempt to understand how characteristics of particular elections may make them harder (or easier) to predict. In particular, we focus on estimating the impact of voter turnout, social trust, electoral change, and vote buying on polling error. We find support for three of the four hypotheses. There is little evidence that voter turnout affects polling error. However, polling errors tend to be higher in less trusting societies and where there have been large changes in parties' voteshare from the previous election. Moreover, we also find that higher prevalence of vote buying may be associated with larger polling errors.

Word Count: 4,872

*The authors would like to thank Mikael Gilljam and Christopher Wlezien for helpful comments and suggestions. A previous version of this paper was presented at the 2016 Annual Meeting of the American Association for Public Opinion Research.

†Department of Political Science, University of Gothenburg

‡Department of Government, University of Texas at Austin

1 Introduction

Pre-election polling exists in almost all democratic countries. The accuracy of the polls varies, however, with pollsters sometimes being close to the actual election outcome and sometimes missing by wide margins. In the latter category, there are notorious cases, as is illustrated by the infamous "Dewey Defeats Truman" headline that the Chicago Daily Tribune (incorrectly) ran after the 1948 U.S. presidential election. More recently, pollsters underestimated the vote share for the Conservatives in the 2015 U.K. election and the support for Donald Trump in several U.S. Midwestern states in the 2016 U.S. presidential election.

Polling failures often lead to soul-searching among survey practitioners, as they try to find out what failed. The commission who assessed where the 1948 polls went wrong pointed to flaws in sampling methods (Mosteller and Doob 1949), as did a similar commission in the U.K. (Sturgis et al. 2016) and the U.S. (Kennedy et al. 2017). These studies focus on single countries, which is also common in academic research on the causes of polling accuracy (for examples, Jowell et al. 1993; Durand, Blais, and Vachon 2001; Traugott 2001). Somewhat less frequently, scholars look at sources of polling inaccuracy within one country over multiple years (Shamir 1986; Mitofsky 1998; Desart and Holbrook 2003; Magalhaes 2005).

In contrast, we know much less about cross-national determinants of polling accuracy. Only recently have scholars gather enough data to be able to systematically analyze polling accuracy cross-nationally (Jennings and Wlezien 2016; Wlezien, Jennings, and Erikson 2017). And while those advances are interesting and important, the field is still relatively new. It is central to understand determinants of polling accuracy from a comparative perspective. Without this, survey practitioners may fret over the error of their survey instruments without realizing that the instruments are no worse than previous years; the election was simply harder to predict due to a more challenging polling environment. Similarly, an increase in accuracy compared to previous years may not signify that pollsters improved, but that the environment of that particular elcetion was more forgiving.

In this paper, we address this gap in the research by assessing what circumstances affect

accuracy of polls across countries and elections. Drawing on research from multiple fields, we identify four possible determinants. Our hypotheses are that polling bias is lower when (1) voter turnout is higher, (2) social trust is higher, (3) electoral change from one election to the next is lower, and (4) vote buying is less prevalent. To examine these propositions, we rely on a dataset (gathered by Jennings and Wlezien 2016) from 44 countries over 71 years.

2 Sources of variability in polling accuracy

Prior research has identified several factors that affect how close polls are to the election outcome. As expected based on probability theory, polls with larger sample sizes tend to be more accurate (Lau 1994; Desart and Holbrook 2003; Magalhaes 2005). We also know that polls are more accurate the closer they are to the election date (Jennings and Wlezien 2016).

The sampling method used can also affect polling accuracy. While opt-in internet surveys can be as accurate as probability-based polls if advanced techniques are used to adjust for the inherent biases of self-selected samples (Ansolabehere and Schaffner 2014; Wang et al. 2015), these samples are often associated with more bias in practice (Yeager et al. 2011; Sohlberg, Gilljam, and Martinsson 2017). In the U.K. 2015 election although both types of polls were biased, those with the most sophisticated probability-based sampling methods (such as the British Election Study) tended to outperform polls that did not rely on full probability sampling (Sturgis et al. 2016, for further analysis of this difference, see; Vavreck and Rivers 2008; Pickup et al. 2011).

With response rates falling worldwide (Curtin, Presser, and Singer 2005; Kohut et al. 2012), there has been much interest in determining whether lower response rates are associated with lower polling accuracy. Generally speaking, this does not seem to be the case. Even though non-response bias is a plausible concern as response rates drop, polls with low response rates are not more biased overall (Keeter et al. 2006; Groves and Peytcheva 2008; Kohut et al. 2012).

As for other factors, more accurate polls tend to screen for likely voters (Desart and Holbrook 2003), are fielded over a longer periods (Lau 1994), and are not sponsored by the parties or candi-

dates themselves (B. Hennessy and E. Hennessy 1961; Shamir 1986; Martin and Traugott 2005).

Institutional features also affect polling accuracy. Jennings and Wlezien (2016) find, for example, that there is a difference in polling accuracy between presidential and parliamentary systems (though that difference disappears by election day). Our study differs from theirs in several important ways. Whereas Jennings and Wlezien (2016) look at variation across electoral systems, we examine variation within an electoral system. They show that systems can explain quite a bit of polling accuracy, but there still is variation to explain, and that is what we examine here.

However, almost all studies on polling accuracy focus on specific countries. A large portion of these examine accuracy in the United States (for example, Traugott 2005; Panagopoulos 2009) or the United Kingdom (for example, Pickup et al. 2011; Sturgis et al. 2016), yet some use data from countries such as France (Arzheimer and Evans 2014), Italy (Callegaro and Gasperoni 2008), Portugal (Magalhaes 2005), Mexico (Cantú, Hoyo, and Morales 2015), and Sweden (Sohlberg, Gilljam, and Martinsson 2017). While these studies are helpful in improving our understanding of polling accuracy, we still know little of the specific electoral circumstances that make polling harder or easier. Drawing on research from multiple areas, we have identified four factors that are likely to influence cross-national polling accuracy over time.

2.1 Voter turnout

When most citizens cast a vote, pollsters potentially have an easier job because they become less reliant on likely voter models. Likely voter models attempt to account for the fact that there may be (and usually are) systematic differences between individuals who vote and those who do not. If we naively include everyone in the sample, then the sample may be biased, especially if nonvoters have different preferences than voters. Unfortunately, these likely voter models are difficult to calibrate, as there are major differences between elections in what kinds of people turn out to vote. For example, in the U.S. 2008 presidential election, African Americans voted at unprecedented rates (Philpot, Shaw, and McGowen 2009), which may have made likely voter models

harder to get right given the departure from historical norm. Another problem with, likely voter models is that they can produce artificial swings in estimates of the electorate because changing levels of enthusiasm affects the classification of voter and nonvoters (Erikson, Panagopoulos, and Wlezien 2004).

Given the problems associated with likely voter models, it might be tempting to disregard them entirely, but that tends to further exacerbate polling mistakes in low-turnout elections. In the U.S., for example, without a likely voter model support for Democratic candidates would be overestimated because nonvoters tend to prefer the Democratic Party to the Republican Party. There is also empirical support for relying on likely voter models, as polls that use them tend to be more accurate (Desart and Holbrook 2003).

In elections with high turnout, there is less need for pollsters to rely on likely voter models because most people who are polled also vote. Of course, elections that have, say, an 85 percent turnout may also be highly biased if the nonvoters strongly and systematically favor a specific party or candidate. In practice, however, we expect that such biases are less common.

A related proposition is the *law of dispersion*, which states that the higher the turnout, the more dispersed electoral participation is for different groups (Tingsten 1937; Persson, Solevid, and Öhrvall 2013). If electoral participation is more equal across societal groups, it follow that likely voter models are less needed to compensate for such imbalances.

Another reason that polling in low turnout elections may be harder is that it could take longer to reach someone who is likely to vote in the election with traditional sampling methods like random digit dialing. This increases the monetary cost to pollsters, who, in turn, may not be able to sample as many voters and conduct callbacks, which are oftentimes necessary. Conversely, in high turnout elections, it is easier to get in touch with voters. This allows pollsters to quickly recruit larger samples for less money.

To the extent that the effect of turnout on polling accuracy has been studied before, prior research has relied on data from single countries. Using Portuguese data, Magalhaes (2005) finds that the low turnout European Parliamentary elections are more difficult to predict than the high

turnout national election. While this finding is consistent with our prediction, it is difficult to rule out the possibility that European Parliamentary elections are harder to poll for other reasons. For example, these are generally considered second order, low salience elections (Vreese et al. 2006; Hix and Marsh 2011; Hobolt and Wittrock 2011).

- *Hypothesis 1: Higher levels of voter turnout reduce polling error on average.*

2.2 Social Trust

When individuals trust each other, a range of beneficial outcomes generally follow. For example, social trust is associated with charitable donations (Uslaner 2002), compliance with rules (Scholz and Lubell 1998), and economic prosperity (Bjørnskov 2012). While definitions of social trust vary somewhat, it is commonly defined in terms of "another." For example, one definition says that social trust is a "psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (Rousseau et al. 1998, p. 395). Expectations about a generalized "other" are also key in the definition put forward by Delhey and Newton (2003, p. 105), as they state that social trust is, "the belief that others will not, at worst, knowingly or willingly do you harm, and will, at best, act in your interest."

Trust influences cooperation (Deutsch 1958). In fact, as decades of research show, one of the most important effects of trust is that it enhances cooperation and alleviates collective action problems (for example Sønderskov 2008; Balliet and Lange 2013). That is, individuals who trust tend to work better with other people and put the greater good before their immediate self interest. Since survey researchers depend on the cooperation of willing participants, we believe that it is easier to conduct accurate polling with a population that is more trusting.

There are two reasons behind this expectation. First, when survey firms attempt to sample a population, respondents should be more likely to answer contact attempts because they do not expect that pollsters have harmful intentions. With positive expectations from the population, response bias should be less problematic. After all, it takes more effort from pollsters to complete surveys with individuals who are less trusting of other people (Keeter et al. 2006). Second, if

respondents trust polling firms, there is a reduced incentive to lie or refuse to answer sensitive questions, which questions about party preferences can be. Consequently, in a more trusting environment, systematic underestimates of support for controversial candidates or parties should be less of a concern.

- *Hypothesis 2: Higher levels of social trust reduce polling error on average.*

2.3 Electoral change

Electoral volatility — change in party support from election to election — is a central concept in research on party systems and may have an effect on polling accuracy. In the 1960s, Lipset and Rokkan (1967) published their influential thesis that the party systems of the time reflect cleavage structures that froze in place in the 1920s. However, there has been an intense discussion among scholars on the empirical support for the frozen party system thesis. Siding with Lipset and Rokkan, one perspective maintains that party systems are largely stable. While new parties arise and support for traditional parties fluctuates, there is remarkable stability in party support over time, according to this view (Mair 1997). In contrast, other researchers argue that traditional forces, such as class, party identification, or ideology, which are associated with stability, have weakened over time (Franklin, Mackie, and Valen 1992; Brug 2010). Overall, the perspective that argues for increasing electoral volatility seems to have more supporters (for example Pedersen 1979; Dalton, McAllister, and Wattenberg 2002; Drummond 2006), and even proponents of the opposing perspective acknowledge that volatility often describes current party systems (Mair 2008).

Electoral volatility is presumably also relevant in polling accuracy. When there is an unusual level of change in party support, it suggests that something important has happened in a society. It may be economic problems, scandals, or other events that shift the status quo. Which such changes, polling also becomes challenging because it is uncertain whether the normal polling procedures are still valid. The standard advice in all forecasting, sometimes described as the *Golden Rule* is to be conservative (Armstrong, Green, and Graefe 2015). However, when some-

thing suggests that the polling environment really has changed, as a comparison between the results from the previous election and polls before the new election indicate, it is presumably tempting to change polling practices. After all, a large polling error could mean a damaged reputation for individual pollsters and consequently less future revenue. Pollsters face a difficult choice; either they rely on their usual polling practices even though societal changes suggest that they will produce erroneous estimates, or they can alter their practices.[1] The latter option is fraught with problems. Since pollsters have little experience with this new situation, it is difficult to know which polling techniques that need to be adjusted (for example, sampling, weighting, or mode) and how much. In the face of such difficulties, it may be tempting to follow the lead of other polls, a phenomenon known as herding, which generally leads to inferior polling overall (Silver 2014; Enten 2014). Moreover, there is a risk that a large degree of electoral volatility triggers an overreaction. For example, if Party A won 60% of the vote last election, but pollsters see that they are polling around 30% of the vote in the polls, pollsters may worry that their sampling methods are off and adjust their sampling technique rather than accept that Party A's support is really around 30%.

- *Hypothesis 3: Larger changes in electoral results increase polling error on average.*

2.4 Vote buying

Vote buying is extensive in many areas of the world, yet its prevalence within countries varies across different social groups (for example Brusco, Nazareno, and Stokes 2004; Bratton 2008).[2] It is usually associated with negative consequences. For example, it makes it harder to hold politicians accountable and reduces the likelihood of good policymaking. Consequently, when citizens receive favors, gifts or money in compensation for their vote, it hinders economic development

[1] There is a lot at stake for pollsters around elections because they know that their estimates will be compared with election results. After all, a large polling error could mean a damaged reputation for individual pollsters and consequently less future revenue.

[2] Of course, we are actually interested in the concept of vote *selling* rather than vote buying, as it is individuals selling their votes that creates a discrepancy between polling and vote results. However, the phrase "vote buying" is much more prevalent in the literature, so we use that here.

and reduces the quality of government (Schaffer 2006). Many hoped that the secret ballot would end (or at least lessen) vote buying, as it is impossible to check whether the seller actually voted for the buyer; however, vote buying continues (Nichter 2008; Lawson and Greene 2014).

Vote buying may also affect average polling error. The intuition is straightforward. If vote buying is low or nonexistent, then people's responses to pollsters questions about their favored candidate should accurately reflect their intention to vote. On the other hand, if there is rampant vote buying, then people may respond that their favored candidate is candidate X but then associates of candidate Y buy their vote. This creates discrepancy between who people vote for and their actual political preferences.

Since few people openly want to admit that someone bought their vote, it can be challenging to assess the prevalence of vote buying. Fortunately, experimental methods such as list experiments can be used to address possible social desirability bias on this issue. They show that vote buying and attempts at buying votes is extensive in Nicaragua (Gonzalez-Ocantos et al. 2011), Mexico (Mizuno 2012) and Lebanon (Corstange 2012). For instance, 24 percent of respondents in Nicaragua report that they have been offered compensation for their vote, though only two percent admit this openly in direct questioning. Another method to gauge the commonness of vote buying is to rely on expert judgments. After an extensive data collection effort on this issue, the Varieties of Democracy project shows that there is marked variation between elections and countries in the frequency of vote buying (Coppedge et al. 2017).

- *Hypothesis 4: Increased levels of vote buying increase polling error on average.*

3 Data & Measurement

Our measures of polls, turnout, and the vote come from Jennings and Wlezien (2016). Whereas they are interested in analyzing the predictive power of the polls over the course of the election (see also Wlezien and Erikson 2002; Jennings and Wlezien 2016; Wlezien, Jennings, and Erikson 2017), we focus on the performance of the polls at the end of the campaign — election day.

Therefore, we exclude observations coming before election day. To be clear — we are not analyzing the predictive power of the polls over the course of the election; we are simply interested in the difference between what the polls indicated would happen on election day versus the actual election outcome.

To calculate the average polling error in an election, we took the mean of the absolute value of the difference between the prediction of the polls on election day and the actual vote outcome for each party. Thus, we have one observation for each country-election. Descriptive statistics of this measure and the main variables we examine are shown in Table 1. The mean is 2.62 and the median is 1.84. This means that on average, the polls get a party's vote share incorrect by about two percentage points the day of the election.

Since we know that polls generally predict the outcome fairly well on election day, one worry is that there is relatively little left to explain. It could be the case, for example, that nearly all of the variation in polling error is due to time (as polling may have improved over the years, for example) or place (if polling in one country is simply better on average than others) or both. If that is the case, this leaves us with relatively little to explain. This, however, does not seem to be the case. An analysis of variance due to years and countries suggests that about 26.2 percent of the variance is attributable to variation across years, 36.3 percent to country-level variation, and 58.1 percent of the variance is explained by country and year together. Therefore, while time and place account for quite a bit of variation in polling accuracy, there is still substantial variation even after that. It is this variation we seek to understand.

[Table 1 about here.]

To measure social trust, we would ideally rely on aggregate country scores based on data from the same polling question for all of the elections included in our data. Since such a measure is unavailable — after all, our election accuracy measures cover the period from the end of the Second World War for some countries — we have chosen to rely on a proxy measure instead. Rather than social trust, we use an index that captures variation in the strength of civil society.

Civil society consists of the space between public society and private life and is made up of organizations that are routinely involved in politics. We also know that places with robust civil society also tend to have high levels of political and social trust (Rothstein and Stolle 2008). In these places, people should be more likely to answer pollsters questions, and perhaps more likely to answer them honestly.

We merged the Jennings and Wlezien (2016) data with information from the *Varieties of Democracy* project (Coppedge et al. 2017) for information on the strength of civil society (our proxy measure of social trust) and the amount of vote buying. The strength of civil society (`v2x_cspart` in the data) are point estimates from a Bayes factor analysis model incorporating information about how centralized the candidate selection process is within party (on a national versus local scale), how often (or whether) major civil society organizations are consulted by policymakers, the extent of involvement in civil society organizations, and the extent to which women participate in civil society organizations.[3] We multiplied this variable by ten to put it on a similar scale with the other variables in our analysis.

Electoral change is calculated as the mean of the absolute value of the difference between a party's share of the vote in the most recent election minus its share in the previous election. We take the mean of that value for all parties in an election[4] Thus, a value of one means that the average party gained or lost one percentage point support from the previous election. Higher values indicate that parties gained or lost quite a bit of support. The lower bound (zero) would mean that parties' vote share did not change at all from the previous election (a value that we do not actually observe in the data).

Finally, vote buying is a variable meant to capture the amount of vote or turnout buying according to a survey of experts. The variable has been recoded from the original data so that

[3] See the Varieties of Democracy codebook for additional information. We also used `v2xcs_ccsi` instead of `v2x_cspart`, which is an alternative measure of the robustness of civil society. Using this variable instead of the other does not affect the results meaningfully.

[4] In calculating this average, we drop parties that competed in the previous election but not in the current or did not compete in the previous election but are in the current. Results are similar if we treat those values as a zero instead of missing data.

higher numbers indicate more vote buying and lower numbers indicate less.[5]

[Figure 1 about here.]

Figure 1 plots the five variables over time. Each observation is represented by a dot with lines connecting observations from the same country. Due to the number of observations, particularly in more recent years, there is considerable overplotting which makes it difficult to see all the observations. To rectify this somewhat, we plot a smoothed average of the variables over time with the solid black line. We also include a rugplot on the left side of each pane that shows the distribution of each variable unconditional on time.

We can see that average polling error has remained relatively stable over time. Pessimistically, this could mean that pollsters are not improving over time. Alternatively, pollsters could be improving over time but adding more countries in more recent years increases the mean error. That is ultimately a testable proposition, but one that is probably unfruitful with the available data given the relatively short time series for most countries.

Figure 1 shows that two variables, social trust and vote buying, do not vary dramatically within country over time (note how the colored lines — one representing each country — tend to be relatively flat). The remaining variables, however, do vary substantially both within country over time and within time across countries. Which of these variables is most predictive of polling errors, though?

4 Analysis

Table 2 shows a summary of the regression models. We present results from four bivariate regressions — including each variable separately — and one regression when we include all four variables at once.[6] Voter turnout seems to be unrelated to polling error. Without controlling for

[5] The scale in the expert surveys is a categorical zero to four scale which is converted to a continuous measure by V-Dem's measurement model.

[6] We also ran models including country and year fixed effects. These results are reported in Appendix A, Table 3 and Table 4. In short, including year fixed effects does not change the results much (it affects the magnitude of the

other variables, the average effect of a one percentage point increase in turnout is to change average error by just -0.01 points, and this is not statistically significant. This accounts for a relatively low amount of the total variance in polling error — just 0.6 percent. Also, once we control for other variables, the coefficient is estimated even closer to zero. So it seems likely that the true effect of turnout on polling error is likely either very small or zero.

[Table 2 about here.]

Social trust (operationalized with our proxy variable) is predictive of polling error. The effect indicates that a more trusting society is associated with decreased polling error, on average. Accounting for the level of social trust explains about 6.7 percent of the variance of polling errors.

Electoral volatility seems to make quite a difference to average polling error and is statistically reliable across various modeling strategies. Larger swings from the previous election are associated with larger errors in the polls on election day. This accounts for about 5 percent of the variance by itself. An average swing of one percentage point results in a change of 0.1 increase in polling error, on average.

We find mixed evidence with regard to vote buying. As vote buying increases polling error increases, on average. This effect is statistically significant at the usual confidence levels and remains so when we add fixed effects for years. However, we must caution against overinterpretation of this result. Adding fixed effects for country (see Table 4) produces an estimate very closer to zero and not statistically significant. This is likely because vote buying is relatively constant within countries across years.

Of course, while it is interesting to note whether one variable affects polling error, we also care about the size of the effect. Although Table 2 can give us some indication, since the scale of the variables differ it is difficult to see which variable (or variables) have a substantially large

coefficients, but not much and does not change the significance levels reported). Including country-level fixed effects decreases our confidence in the effect of vote buying, most likely because this variable does not change much over time within country, as mentioned above. We also ran models controlling for whether the election was a presidential or legislative election, as Jennings and Wlezien (2016) find that this is associated with differences in polling error over the course of the election (though not at election day). Results are reported in Appendix B, Table 4. Controlling for election type makes very little difference to the results.

effect on polling error. For this reason, we present the predicted change in polling error from a one standard deviation increase in each of the independent variables in Figure 2.

[Figure 2 about here.]

From Figure 2, we can see that all variables have similar effect magnitudes, with the exception of turnout, which does not predict polling error. A one standard deviation increase in any of the variables changes the predicted polling error by about half a percentage point. A change in polling error of 0.4 is a substantially meaningful change — many elections have been decided by a tighter margin of victory than that. The point estimate for the effect of a one standard deviation shift in electoral change is substantially higher than the others at around 0.5 if we control for the other variables and 0.7 if we do not.

5 Conclusion and Discussion

The polling industry is rapidly changing in response to various challenges. For example, declining survey participation rates and rising costs have forced organizations to try new methods of contacting people and new survey modes. However, even with these innovations, polls are always at least a little wrong on election day, and sometimes they are off by an embarrassing amount. With so much at stake for both pollsters and academics, it is no wonder that a large number of studies have examined variation in polling accuracy. These studies have greatly improved our understanding of factors that increase or decrease accuracy. However, we still know relatively little about the effects of election-specific variables.

In this paper, we began to address the possibility that not all elections are equally easy to predict by looking at election-specific variables. Our findings show that elections in countries with low social trust are difficult to poll. Prior research demonstrates that a higher level of social trust is associated with a number of sought-after characteristics (for example, economic prosperity). We can now add polling accuracy to that list, though with the caveat that we were forced to rely on a proxy measure of civil society strength rather than a direct operationalization of social trust.

We also found support for the notion that elections conducted around large changes in party support are particularly difficult to poll. Party scholars have documented an increase in electoral volatility, in particular after the turn of the new millennium, which is due in part to the replacement of more long-term and stable forces like party identification with short-term economic evaluations (Kayser and Wlezien 2011; Dassonneville and Hooghe 2015). The increase in electoral change is obviously not behind all polling mistakes, yet it is possible that the polling industry is getting its reputation hurt by a factor beyond its control.

Normative arguments against vote buying are quite strong; many believe that it corrupts the democratic process. From a more empirical perspective, it also appears to increase polling error. This implies that pollsters should not assume that knowledge gained from decades of polling in places that have little vote buying applies equally well in contexts where vote buying is more prevalent.

In contrast to a popular narrative, different levels of voter turnout do not seem to be systematically associated with increased polling error. Of course, absence of evidence of an effect is not the same as evidence of absence of an effect. Frequentist statistics lacks a straightforward way to test whether the null hypothesis is true. However, given the relatively large number of observations in our sample, the absence of serious measurement problems for turnout and polling error, and the fact that the (statistically insignificant) coefficients are quite close to zero, we conclude that turnout is unrelated to polling bias. Perhaps pollsters in low-turnout elections are better at compensating for the lower levels by, for example, relying on likely voter models. Alternatively, our expectation was wrong, which indicates that voters and non-voters are more similar than we thought.

Effect sizes for social trust, electoral change, and vote buying are substantially meaningful. For example, a typical swing in electoral change increases polling error by about 0.4 percentage points on average (see Figure 2 and accompanying discussion), and the effect size for social trust and vote buying is estimated to be similar, perhaps slightly smaller. Half a percentage point (or more if the effects are combined) can mean the difference between confidently stating which

parties that will be the eventual winner and losers, as we have seen in several recent elections.

We have already hinted that several actors can benefit from our results. For instance, when pre-election polls suggest that party support has changed substantially compared to the last election, pollsters would do well to inform its audience that predictions are particularly uncertain. Moreover, election post-mortems should vary in harshness depending on electoral change. Lastly, individuals who construct forecasting models would be advised to change their levels of uncertainty depending on fluctuation in party support between elections. That is, the point estimates of support based on poll aggregates need to be accompanied by larger uncertainty if there has been high electoral volatility. Likewise, pollsters, commentators, and forecasters should take into account levels of social trust and the prevalence of vote buying. Knowing these levels determines whether the polling environment is easy or hard.

6 References

Ansolabehere, Stephen and Brian Schaffner (2014). "Does Survey Mode Still Matter? Findings From a 2010 Multi-Mode Comparison." *Political Analysis* 22.03, pp. 285–303.

Armstrong, J. Scott, Kesten Green, and Andreas Graefe (2015). "Golden Rule of Forecasting: Be Conservative." *Journal of Business Research* 68.8, pp. 1717–1731.

Arzheimer, Kai and Jocelyn Evans (2014). "A New Multinomial Accuracy Measure for Polling Bias." *Political Analysis* 22.01, pp. 31–44.

Balliet, Daniel and Paul Van Lange (2013). "Trust, Conflict, and Cooperation: A Meta-Analysis." *Psychological Bulletin* 139.5, pp. 1090–1112.

Bjørnskov, Christian (2012). "How Does Social Trust Affect Economic Growth?" *Southern Economic Journal* 78.4, pp. 1346–1368.

Bratton, Michael (2008). "Vote Buying and Violence in Nigerian Election Campaigns." *Electoral Studies* 27.4, pp. 621–632.

Brug, Wouter van der (2010). "Structural and Ideological Voting in Age Cohorts." *West European Politics* 33.3, pp. 586–607.

Brusco, Valeria, Marcelo Nazareno, and Susan Carol Stokes (2004). "Vote Buying in Argentina." *Latin American Research Review* 39.2, pp. 66–88.

Callegaro, Mario and Giancarlo Gasperoni (2008). "Accuracy of Pre-Election Polls for the 2006 Italian Parliamentary Election: Too Close To Call." *International Journal of Public Opinion Research* 20.2, pp. 148–170.

Cantú, Francisco, Verónica Hoyo, and Marco A. Morales (2015). "The Utility of Unpacking Survey Bias in Multiparty Elections: Mexican Polling Firms in the 2006 and 2012 Presidential Elections." *International Journal of Public Opinion Research* 28.1, pp. 96–116.

Coppedge, Michael, John Gerring, Staffan Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Joshua Krusell, Anna Luhrmann, Kyle Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Constanza Sanhueza Petrarca, Johannes von Romer, Laura Saxer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova, and Steven Wilson (2017). *V-Dem Dataset v7.0*.

Corstange, Daniel (2012). "Vote Trafficking in Lebanon." *International Journal of Middle East Studies* 44.3, pp. 483–505.

Curtin, Richard, Stanley Presser, and Elanor Singer (2005). "Changes in Telephone Survey Nonresponse Over the Past Quarter Century." *Public Opinion Quarterly* 69.1, pp. 87–98.

Dalton, Russell, Ian McAllister, and Martin Wattenberg (2002). "The Consequences of Partisan Dealignment." *Parties Without Partisans*. Parties Without Partisans. Oxford University Press, pp. 37–62.

Dassonneville, Ruth and Marc Hooghe (2015). "Economic Indicators and Electoral Volatility: Economic Effects on Electoral Volatility in Western Europe, 1950-2013." *Comparative European Politics*.

Delhey, Jan and Kenneth Newton (2003). "Who Trusts?: The Origins of Social Trust in Seven Societies." *European Societies* 5.2, pp. 93–137.

Desart, Jay and Thomas Holbrook (2003). "Campaigns, Polls, and the States: Assessing the Accuracy of Statewide Presidential Trial-Heat Polls." *Political Research Quarterly* 56.4, pp. 431–439.

Deutsch, Morton (1958). "Trust and Suspicion." *Journal of Conflict Resolution* 2.4, pp. 265–279.

Drummond, Andrew (2006). "Electoral Volatility and Party Decline in Western Democracies: 1970-1995." *Political Studies* 54.3, pp. 628–647.

Durand, Claire, André Blais, and Sébastien Vachon (2001). "Review: A Late Campaign Swing Or a Failure of the Polls? The Case of the 1998 Quebec Election." *Public Opinion Quarterly* 65.1, pp. 108–123.

Enten, Harry (2014). *Are Bad Pollsters Copying Good Pollsters*. URL: `https://fivethirtyeight.com/features/are-bad-pollsters-copying-good-pollsters/`.

Erikson, Robert, Costas Panagopoulos, and Christopher Wlezien (2004). "Likely (and Unlikely) Voters and the Assessment of Campaign Dynamics." *Public Opinion Quarterly* 68.4, pp. 588–601.

Franklin, Mark, Thomas Mackie, and Henry Valen (1992). *Electoral change: Responses to Evolving Social and Attitudinal Structures in Western Countries*. ECPR Press.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet de Jonge, Carlos Meléndez, Javier Osorio, and David Nickerson (2011). "Vote Buying and Social Desirability Bias: Experimental Evidence From Nicaragua." *American Journal of Political Science* 56.1, pp. 202–217.

Groves, Robert and Emilia Peytcheva (2008). "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* 72.2, pp. 167–189.

Hennessy, Bernard and Erna Hennessy (1961). "The Prediction of Close Elections: Comments on Some 1960 Polls." *Public Opinion Quarterly* 25.3, pp. 405–411.

Hix, Simon and Michael Marsh (2011). "Second-Order Effects Plus Pan-European Political Swings: An Analysis of European Parliament Elections Across Time." *Electoral Studies* 30.1, pp. 4–15.

Hobolt, Sara Binzer and Jill Wittrock (2011). "The Second-Order Election Model Revisited: An Experimental Test of Vote Choices in European Parliament Elections." *Electoral Studies* 30.1, pp. 29–40.

Jennings, Will and Christopher Wlezien (2016). "The Timeline of Elections: A Comparative Perspective." *American Journal of Political Science* 60.1, pp. 219–233.

Jowell, Roger, Barry Hedges, Peter Lynn, Graham Farrant, and Anthony Heath (1993). "Review: The 1992 British Election: The Failure of the Polls." *Public Opinion Quarterly* 57.2, p. 238.

Kayser, Mark Andreas and Christopher Wlezien (2011). "Performance Pressure: Patterns of Partisanship and the Economic Vote." *European Journal of Political Research* 50.3, pp. 365–394.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill (2006). "Gauging the Impact of Growing Nonresponse on Estimates From a National Rdd Telephone Survey." *Public Opinion Quarterly* 70.5, pp. 759–779.

Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua Clinton, Claire Durand, Charles Franklin, Kyley McGeeney, Lee Miringoff, Kristen Olson, Doug Rivers, Lydia Saad, Evans Witt, and Christopher Wlezien (2017). *An Evaluation of 2016 Election Polls in the United States*. Tech. rep. American Association of Public Opinion Research.

Kohut, Andrew, Scott Keeter, Carroll Doherty, Michael Dimock, and Leah Christian (2012). *Assessing the Representativeness of Public Opinion Surveys*. Tech. rep. Washington DC: Pew Research Center.

Lau, Richard (1994). "An Analysis of the Accuracy of "Trial Heat" Polls During the 1992 Presidential Election." *Public Opinion Quarterly* 58.1, pp. 2–20.

Lawson, Chappell and Kenneth Greene (2014). "Making Clientelism Work: How Norms of Reciprocity Increase Voter Compliance." *Comparative Politics* 47.1, pp. 61–85.

Lipset, Seymour Martin and Stein Rokkan (1967). "Cleavage Structures, Party Systems And Voter Alignments: An Introduction." *Party Systems and Voter Alignments: Cross-National Perspectives*. Ed. by Seymour Martin Lipset and Stein Rokkan. Free Press.

Magalhaes, Pedro (2005). "Pre-Election Polls in Portugal: Accuracy, Bias, and Sources of Error, 1991-2004." *International Journal of Public Opinion Research* 17.4, pp. 399–421.

Mair, Peter (1997). *Party System Change: Approaches and Interpretations*. Oxford University Press.

— (2008). "The Challenge To Party Government." *West European Politics* 31.1-2, pp. 211–234.

Martin, Elizabeth and Michael Traugott (2005). "A Review and Proposal for a New Measure of Poll Accuracy." *Public Opinion Quarterly* 69.3, pp. 342–369.

Mitofsky, Warren (1998). "Review: Was 1996 a Worse Year for Polls Than 1948?" *Public Opinion Quarterly* 62.2, p. 230.

Mizuno, Elda Lorena Becerra (2012). "Does Everyone Have a Price? The Demand Side of Clientelism and Vote-Buying in an Emerging Democracy." PhD thesis. Duke University.

Mosteller, Frederick and Leonard William Doob (1949). *The Pre-Election Polls of 1948*. Social Science Research Council.

Nichter, Simeon (2008). "Vote Buying Or Turnout Buying? Machine Politics and the Secret Ballot." *American Political Science Review* 102.01, pp. 19–31.

Panagopoulos, Costas (2009). "Polls and Elections: Preelection Poll Accuracy in the 2008 General Elections." *Presidential Studies Quarterly* 39.4, pp. 896–907.

Pedersen, Mogens (1979). "The Dynamics of European Party Systems: Changing Patterns of Electoral Volatility." *European Journal of Political Research* 7.1, pp. 1–26.

Persson, Mikael, Maria Solevid, and Richard Öhrvall (2013). "Voter Turnout and Political Equality: Testing the 'law of Dispersion' in a Swedish Natural Experiment." *Politics* 33.3, pp. 172–184.

Philpot, Tasha, Daron Shaw, and Ernest McGowen (2009). "Winning the Race: Black Voter Turnout in the 2008 Presidential Election." *Public Opinion Quarterly* 73.5, pp. 995–1022.

Pickup, Mark, J. Scott Matthews, Will Jennings, Robert Ford, and Stephen D. Fisher (2011). "Why Did the Polls Overestimate Liberal Democrat Support? Sources of Polling Error in the 2010 British General Election." *Journal of Elections, Public Opinion & Parties* 21.2, pp. 179–209.

Rothstein, Bo and Dietlind Stolle (2008). "The State and Social Capital: An Institutional Theory of Generalized Trust." *Comparative Politics* 40.4, pp. 441–459.

Rousseau, Denise, Sim Sitkin, Ronald Burt, and Colin Camerer (1998). "Not So Different After All: A Cross-Discipline View OF Trust." *Academy of Management Review* 23.3, pp. 393–404.

Schaffer, Frederic Charles (2006). *Elections For Sale : The Causes and Consequences of Vote Buying*. Rienner.

Scholz, John T. and Mark Lubell (1998). "Adaptive Political Attitudes: Duty, Trust, and Fear As Monitors of Tax Policy." *American Journal of Political Science* 42.3, p. 903.

Shamir, Jacob (1986). "Preelection Polls in Israel: Structural Constraints on Accuracy." *Public Opinion Quarterly* 50.1, p. 62.

Silver, Nate (2014). *Here's Proof Some Pollsters Are Putting a Thumb on the Scale*. URL: `https://fivethirtyeight.com/features/heres-proof-some-pollsters-are-putting-a-thumb-on-the-scale/`.

Sohlberg, Jacob, Mikael Gilljam, and Johan Martinsson (2017). "Determinants of Polling Accuracy: the Effect of Opt-In Internet Surveys." *Journal of Elections, Public Opinion and Parties*.

Sønderskov, Kim Mannemar (2008). "Environmental Group Membership, Collective Action and Generalised Trust." *Environmental Politics* 17.1, pp. 78–94.

Sturgis, Patrick, Nick Baker, Mario Callegaro, Stephen Fisher, Jane Green, Will Jennings, Jouni Kuha, Ben Lauderdale, and Patten Smith (2016). *Report of Inquiry into the 2015 British General Election Opinion Polls*. Tech. rep. Market Research Society and British Polling Council.

Tingsten, Herbert (1937). *Political behavior: Studies in election statistics*. 7. PS King.

Traugott, Michael (2001). "Assessing Poll Performance in the 2000 Campaign." *Public Opinion Quarterly* 65.3, pp. 389–419.

— (2005). "The Accuracy of the National Preelection Polls in the 2004 Presidential Election." *Public Opinion Quarterly* 69.5, pp. 642–654.

Uslaner, Eric (2002). *The Moral Foundations of Trust*. Cambridge University Press.

Vavreck, Lynn and Douglas Rivers (2008). "The 2006 Cooperative Congressional Election Study." *Journal of Elections, Public Opinion and Parties* 18.4, pp. 355–366.

Vreese, Claes de, Susan Banducci, Holli Semetko, and Hajo Boomgaarden (2006). "The News Coverage of the 2004 European Parliamentary Election Campaign in 25 Countries." *European Union Politics* 7.4, pp. 477–504.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman (2015). "Forecasting Elections With Non-Representative Polls." *International Journal of Forecasting* 31.3, pp. 980–991.

Wlezien, Christopher and Robert Erikson (2002). "The Timeline of Presidential Election Campaigns." *Journal of Politics* 64.4, pp. 969–993.

Wlezien, Christopher, Will Jennings, and Robert Erikson (2017). "The "Timeline" Method of Studying Electoral Dynamics." *Electoral Studies*.

Yeager, David, Jon Krosnick, LinChiat Chang, Harold Javitz, Matthew Levendusky, Alberto Simpser, and Rui Wang (2011). "Comparing the Accuracy of Rdd Telephone Surveys and Internet Surveys Conducted With Probability and Non-Probability Samples." *Public Opinion Quarterly* 75.4, pp. 709–747.

A Fixed effects

Here we report results when we include year fixed effects (Table 3) and country fixed effects (Table 4). Because several observations only had one year or one country in the data, we are forced to drop these, so these estimates rely on a slightly reduced dataset.

[Table 3 about here.]

[Table 4 about here.]

B Type of election

Here we control for whether the election is a legislative or presidential election, as Jennings and Wlezien (2016) find that this matters for poll accuracy. Note that there are only a small number of presidential elections in the dataset; just 39 total.

[Table 5 about here.]

List of Tables

Table 1: Descriptive statistics of the dependent variable and four independent variables.

Statistic	N	Mean	St. Dev.	Min	Max
Polling error	295	2.62	2.24	0.02	13.47
Voter turnout	308	73.55	15.81	32.50	96.32
Social trust	308	9.07	0.75	5.72	9.88
Electoral change	295	5.15	4.63	0.08	35.40
Vote buying	308	−1.70	0.92	−3.64	1.43

Table 2: Predicting polling error. *Models 1–4 are results from bivariate regression; Model 5 includes all four independent variables in a single regression.*

	Average Polling Error				
	Model 1	Model 2	Model 3	Model 4	Model 5
Voter turnout	−0.011				−0.001
	(0.008)				(0.008)
Social trust		−0.774***			−0.459**
		(0.169)			(0.185)
Electoral change			0.101***		0.084***
			(0.026)		(0.026)
Vote buying				0.515***	0.350**
				(0.137)	(0.158)
Constant	3.379***	9.637***	2.018***	3.483***	6.914***
	(0.599)	(1.533)	(0.184)	(0.263)	(1.751)
Observations	294	294	286	294	285
R^2	0.006	0.067	0.050	0.046	0.122
Adjusted R^2	0.003	0.064	0.047	0.043	0.109
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3: Predicting polling error including year fixed effects

	Average Polling Error				
	(1)	(2)	(3)	(4)	(5)
Voter turnout	-0.016^{*}				-0.006
	(0.009)				(0.009)
Social trust		-0.683^{***}			-0.422^{**}
		(0.191)			(0.208)
Electoral change			0.093^{***}		0.083^{***}
			(0.029)		(0.028)
Vote buying				0.563^{***}	0.450^{**}
				(0.156)	(0.178)
Observations	283	283	275	283	274
R^2	0.246	0.272	0.265	0.272	0.339
Adjusted R^2	0.042	0.075	0.059	0.076	0.141

Note: *p<0.1; **p<0.05; ***p<0.01
Includes year fixed effects

Table 4: Predicting polling error including country fixed effects

	Average Polling Error				
	(1)	(2)	(3)	(4)	(5)
Voter turnout	0.032				0.009
	(0.021)				(0.021)
Social trust		-1.269^{***}			-1.115^{**}
		(0.439)			(0.489)
Electoral change			0.080^{***}		0.081^{***}
			(0.029)		(0.029)
Vote buying				0.614^{**}	0.001
				(0.307)	(0.333)
Observations	286	287	278	287	278
R^2	0.328	0.370	0.360	0.360	0.381
Adjusted R^2	0.231	0.280	0.267	0.267	0.282
Note:					*p<0.1; **p<0.05; ***p<0.01
					Includes country fixed effects

Table 5: Predicting polling error controlling for election type

	Average Polling Error				
	(1)	(2)	(3)	(4)	(5)
Voter turnout	−0.011				−0.0003
	(0.008)				(0.008)
Social trust		−0.685***			−0.468**
		(0.176)			(0.187)
Electoral change			0.089***		0.088***
			(0.029)		(0.029)
Vote buying				0.439***	0.360**
				(0.141)	(0.161)
presidential	0.967***	0.557*	0.324	0.665**	−0.129
	(0.310)	(0.326)	(0.349)	(0.325)	(0.354)
Constant	3.181***	8.716***	2.019***	3.217***	6.988***
	(0.594)	(1.620)	(0.184)	(0.293)	(1.765)
Observations	294	294	286	294	285
R^2	0.038	0.077	0.053	0.060	0.122
Adjusted R^2	0.032	0.070	0.046	0.053	0.107

Note: *p<0.1; **p<0.05; ***p<0.01

Rerun of main models, controlling for legislative or presidential elections

List of Figures

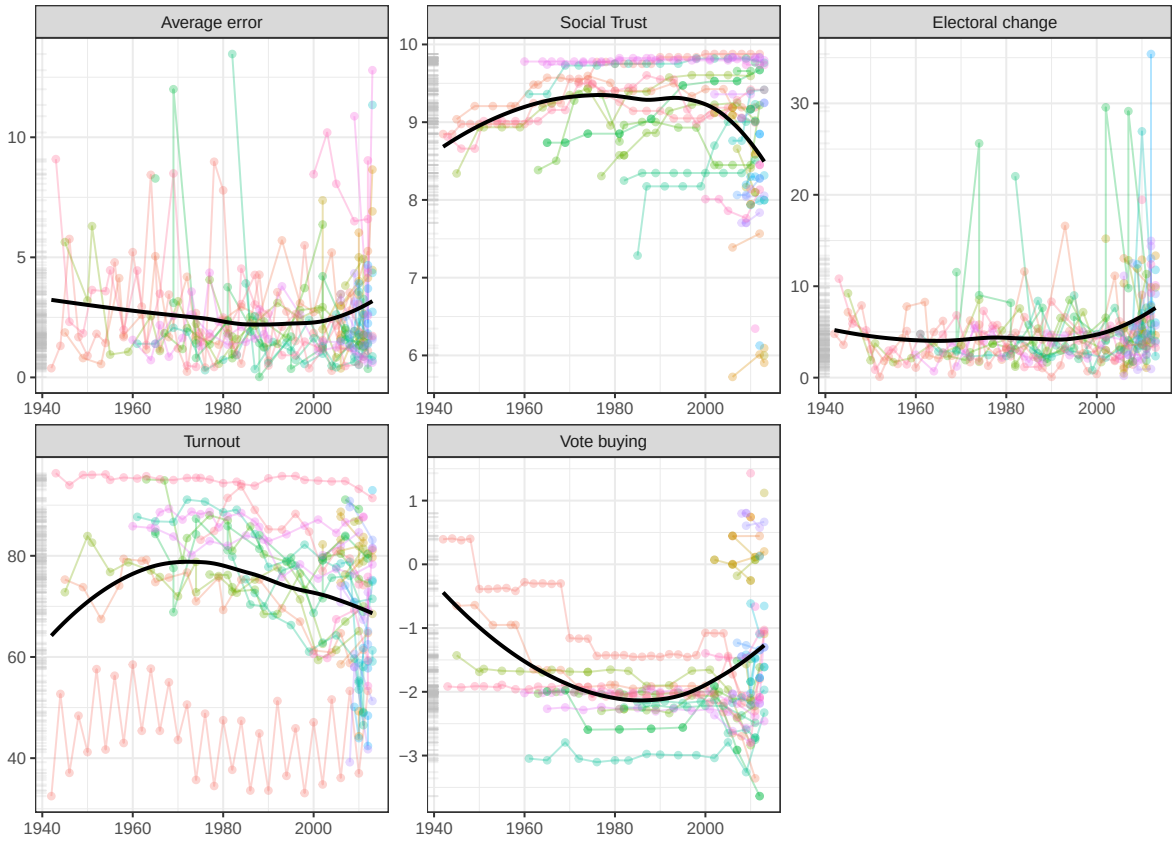

Figure 1: Black lines plot a smoothed average of the variables by year over time. Faded lines plot all the observations with one line per country. Grey lines on the side are a rugplot of the variables unconditional on year.

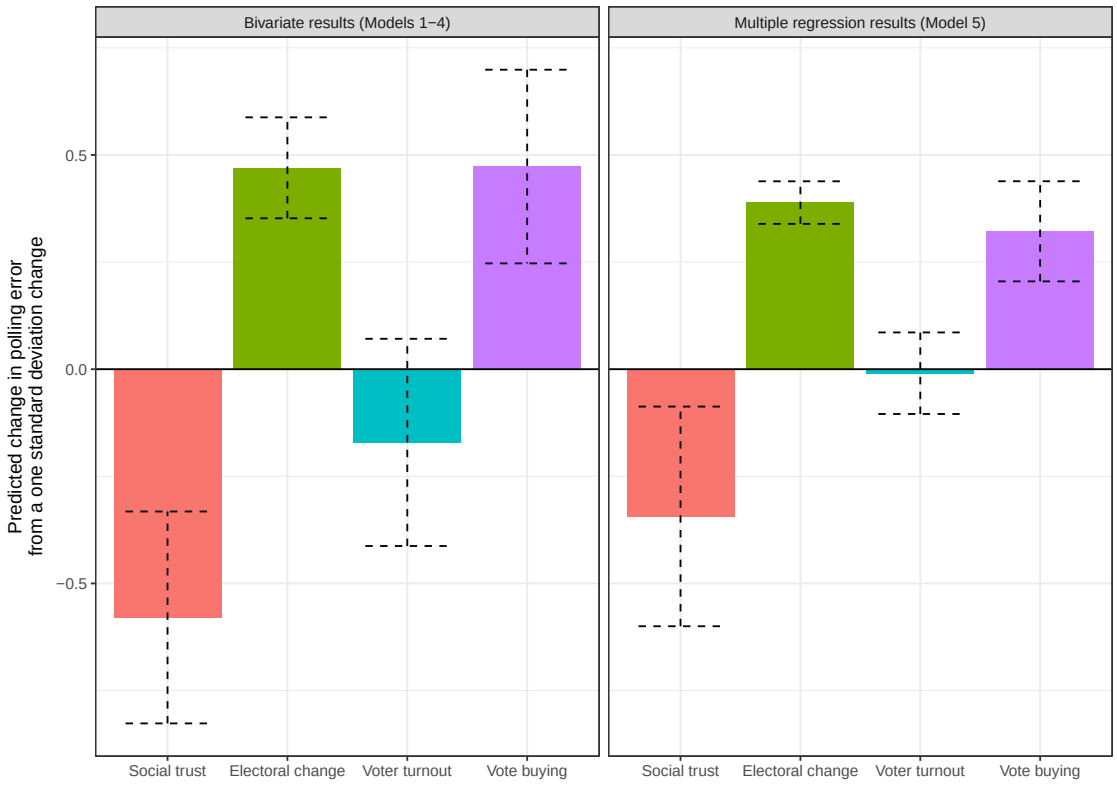

Figure 2: Effect of a one standard deviation change in each variable. *The left panel reports estimates from four bivariate models. The right panel reports estimates from a single regression model. Dashed lines represent* 95% *confidence intervals*