Missing the Trees in the Forest: Polling Error and Partisanship

Merlin Heidemanns* Andrew Gelman†

March 2, 2021

Abstract

Looking at state polls from U.S. presidential elections since 2000, we find a consistent pattern of polling errors that favor Democratic candidates in Republican states, and vice versa, with the result that state elections typically look closer than they actually are. Similar patterns appear when looking at state polls for Senate and House races but not for elections for governor. Focusing on the presidential polls, we would like to frame the result in terms of differential nonresponse, but we do not yet have a good explanation.

Keywords: polling error, differential non-response, turnout, U.S. elections

1 A data pattern and a puzzle

This paper presents a data pattern and a puzzle relating to state polls for presidential elections. The data pattern is that polling errors are correlated with state partisanship. For example, the polls in Wyoming and West Virginia consistently show much more support for the Democratic candidate than actually happens in the election, while the polls in Hawaii and California consistently show pro-Republican errors. The puzzle is: why is this happening?

2 Background

Survey estimates suffer from two types of error, sampling and non-sampling error, which are usually discussed in a total survey error framework. The former refers to random variation inherent in computing an estimate from a sample rather than the population, and the latter addresses more pernicious problems such as a mismatch between the target and sampling frame of a survey.

For the purpose of forecasting elections, pollsters have to estimate both who will vote and how they will vote. In U.S. elections, public polling has experienced considerable

*Department of Political Science, Columbia University, New York. Email: mnh2123@columbia.edu.
†Department of Statistics and Department of Political Science, Columbia University, New York

misses in many high profile elections, most notably in 1948, 2016, and, to a lesser extent, 2020. In this paper we study a specific pattern of predictable partisan bias in polling errors in elections since 2000. Polling error is larger in more partisan states and favors the underdog, making candidates appear artificially more competitive; see 1.

This research addresses practical and theoretical issues. First, forecasting models that rely on public opinion polling should adjust for bias within the polls to make forecasts more accurate. For example, polling that is relatively favorable to Republicans in New Jersey should be taken with a similar level of skepticism as polling favorable to Democrats in Kentucky, when predicting the election outcome. Second, campaign spending is frequently dependent upon the perceived viability of a campaign. If spending on the aforementioned states appears artificially viable, this can influence allocation of campaign resources (Gelman et al., 1998). Third, while the prior cases suggest potential adjustments after the fact, ideally pollsters would change their procedures according to the observed problem to eliminate the pattern. This addresses the question of what specifically pollsters estimate incorrectly.

3 Polling Error

Polling error is the difference between the true population parameter θ and its survey estimate $\hat{\theta}$. Polling error consists of sampling error and non-sampling error. The former arises from estimating θ using a sample rather than the population. For example, given a yes or no question posed to 1000 respondents we expect the estimate of the mean $\hat{\theta}$ to fall within 1.5 percentage points of θ 68% of the time. We can reduce sampling error by increasing the sample size and the error decreases with its square root.

Non-sampling error on the other hand usually arises from substantial differences between the target and the sampling population. If individuals who are part of the target population are substantially different than the sampling population with respect to the attribute we are interested in, bias can arise in the absence of specific adjustments. Non-sampling error is difficult to detect and pollsters generally assume it to be absent when reporting uncertainty estimates.

3.1 Polling error in U.S. elections

Understanding the nature of non-sampling polling error in U.S. elections is the topic of ongoing research. Shirani-Mehr et al. (2018) show that in practice observed error is larger by a factor of two than would be expected under simple random sampling (SRS). This suggests considerable non-sampling error. Separately, Gelman et al. (2016) show that some part of observed variation in polling results over an election cycle can be attributed to differential non-response rates as partisans are more or less willing to answer surveys. In the context of election forecasting, scholars also consider observed polling error to be correlated across geographical units. The idea here is that underlying correlation in a cause of polling error will be reflected in the former's distribution.

Here, we investigate a separate pattern observed previously for the 2016 election in Gelman & Azari (2017). We start by estimating state-year polling errors using a Binomial model of the form,

$$y_i \sim \text{Binomial}(n_i, \text{logit}^{-1}(\alpha_{s[i],t[i]} + \xi_{s[i],t[i]})) \tag{1}$$

where i indexes polls three weeks before the presidential election; y is the number of respondents indicating their vote intention for the Democratic candidate and n the number of respondents who indicate a preference for either major party candidate; α is the election outcome as the two-party vote share of the Democratic candidate and ξ the polling error; and s and t index states and election years, respectively. The polling error is the amount by which the polling average overestimates the Democratic vote share. We just use polls in the before the election because vote intentions are close to fixed at that point (Erikson & Wlezien, 2012; Shirani-Mehr et al., 2018). The model is fit in Stan Stan Development Team (2020) through R (R Core Team, 2020) with polling data collected by FiveThirtyEight (2021). The pattern becomes slightly more pronounced when the polling error is differentiated into national, state, regional, division, and time components to allow for information to be shared across states.

We graph the posterior estimates of the polling error for each state against the Republican vote share in the previous election. We do not plot vs. the vote share in the current election out of concern for regression-to-the-mean effects (see for example section 11.3 of Gelman et al. (2020) to account for a potential regression to the mean effect. The graph in 1 shows the relationship (a version with state-labels is in the appendix). The state-level polling error is correlated with the state's partisanship. Specifically, states that voted Republican in previous election tend to show polling errors that favors the Democratic candidate, and vice versa. We observe this relationship to some extent in every Presidential election from 2000 to 2020 as can be seen in Figure 1. While the average error varies, the slope is positive in every year.

An immediate question is whether this pattern is a general feature of elections in the U.S. Following the same approach as for presidential elections, we investigate elections for the U.S. Senate and House of Representatives as well as for governors. While data is at points more sparse, there are some similarities for Senate and House elections even if they are less consistent. The pattern is inconsistent for elections for governor. These graphs can be found in the appendix.

Before discussing two potential explanations for this pattern, we start by considering whether survey respondents are lying. As this argument goes for example in Enns et al. (2017), respondents could be subject to social desirability bias when asked for their vote intention and misrepresent it. As outlined in section 5 of Gelman & Azari (2017), this theory could explain polling error favoring Democratic candidates in the more Democratic states (which does not occur) but is less plausible in strongly Republican states, where we would expect social desirability bias to be weaker or to go in the other direction. The observation in 1 provides further evidence against this argument, as these between-state patterns of polling bias also arose for less objectionable candidates prior to 2016. In essence, the social-desirability argument predicts a flat mean shift, but not a positive slope.

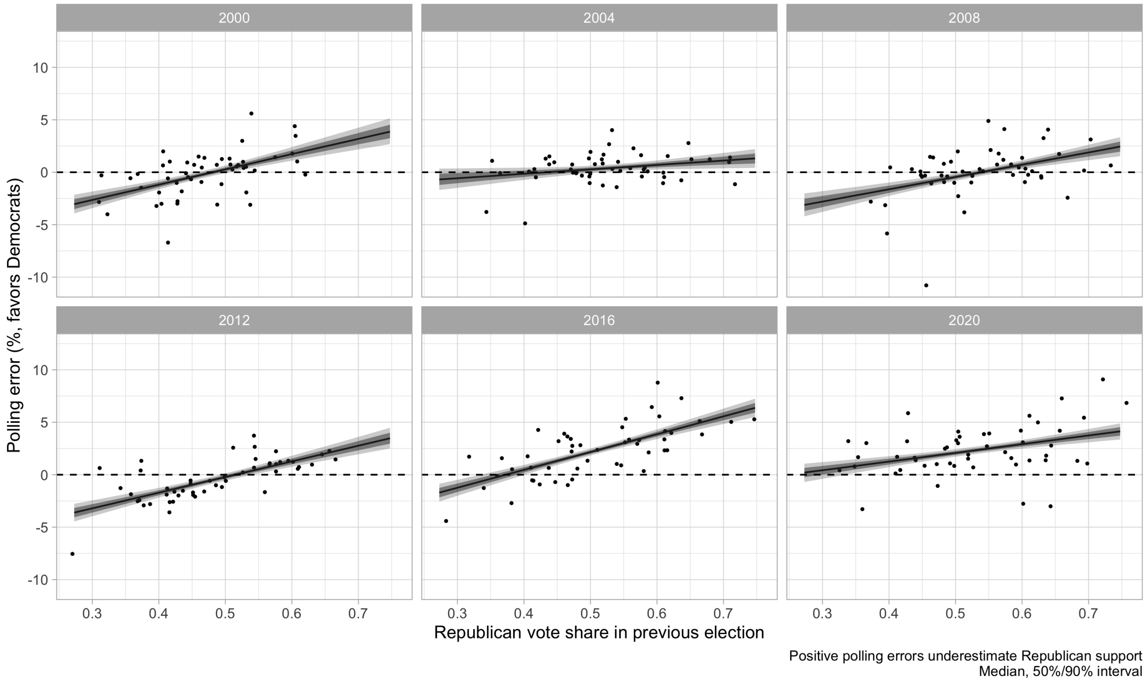


FIGURE 1: Polling error in U.S. presidential elections

3.2 Differential nonresponse

Sampling theory posits the simple random sample as its ideal whereby the probability of inclusion in the sample is identical for each individual in the target population. Nonresponse bias arises when the feature such as a respondent's vote intention is correlated with their propensity to respond to the survey. Survey sampling or post-survey adjustment measures are meant to account or adjust for known average differences between groups such as the differential propensity between genders to respond to surveys. Differential nonresponse bias then arises when these adjustment are insufficient as demonstrated by Gelman et al. (2016) who have shown that so-called convention bounces whereby a candidate gains in the polls after their respective convention are largely attributable to a greater share of supporters of that candidate in the sample.

In this case, the argument would start by assuming that Republican voters are less trustful towards pollsters. Therefore, in more Republican states Democratic voters are overrepresented leading pollsters to overestimate Democratic support. As the general Republican partisanship in the state increases, for example as we move from Ohio to Wyoming, these effect becomes more pronounced.

Yet, this argument would predict negligible polling error as the state becomes less Republican. Conversely one would have to assume that Democratic voters are equally distrustful of pollsters and that the two effects cancel each other out in swing states that are close to 50/50. In the absence of evidence showing similar patterns of distrust among Democratic voters this argument is unable to explain the observed relationship specifically in 2012 or 2000 for example.

3.3 Unexpected patterns in turnout

A turnout based explanation would consider that pollster underestimate turnout among Democratic voters in more Democratic states and among Republicans in more Republican states. We suspect that unexpectedly high Republican turnout explains some of the level shift in 2020, when nearly all the states showed polling errors favoring the Democrats), but this would not seem to be a good explanation of the slopes of the lines in Figure 1.

4 Alternative explanations

4.1 Events and poll timing

A potential alternative explanation for this pattern is that less competitive or more partisan states are polled less frequently.[1] Given that the state is not competitive, poll results become less news worthy. Phrased differently, the claim is that more partisan are only polled when there is a temporary shock to public opinion that makes the state more competitive. When the state is then polled, the polling error will be inflated and by more the more partisan the state is.

The empirical expectation is that the pattern disappears when poll timing is independent of local events. We probe this conjecture by considering only a single pollster, in this case YouGov for the Cooperative Election Study (Ansolabehere, 2010; Ansolabehere & Schaffner, 2013, 2017). Comparing polling error against the Republican vote share of the two party vote this pattern persists for 2008, 2012, and 2016. This suggests that poll timing is not responsible for the pattern.

5 Discussion

5.1 Why is this happening?

We think it's differential nonresponse but we don't have a good story yet, nor do we have any great ideas of how to test whatever theory we do come up with.

5.2 Other data

Geographical units within states.

State results from other national polls.

Data in other countries.

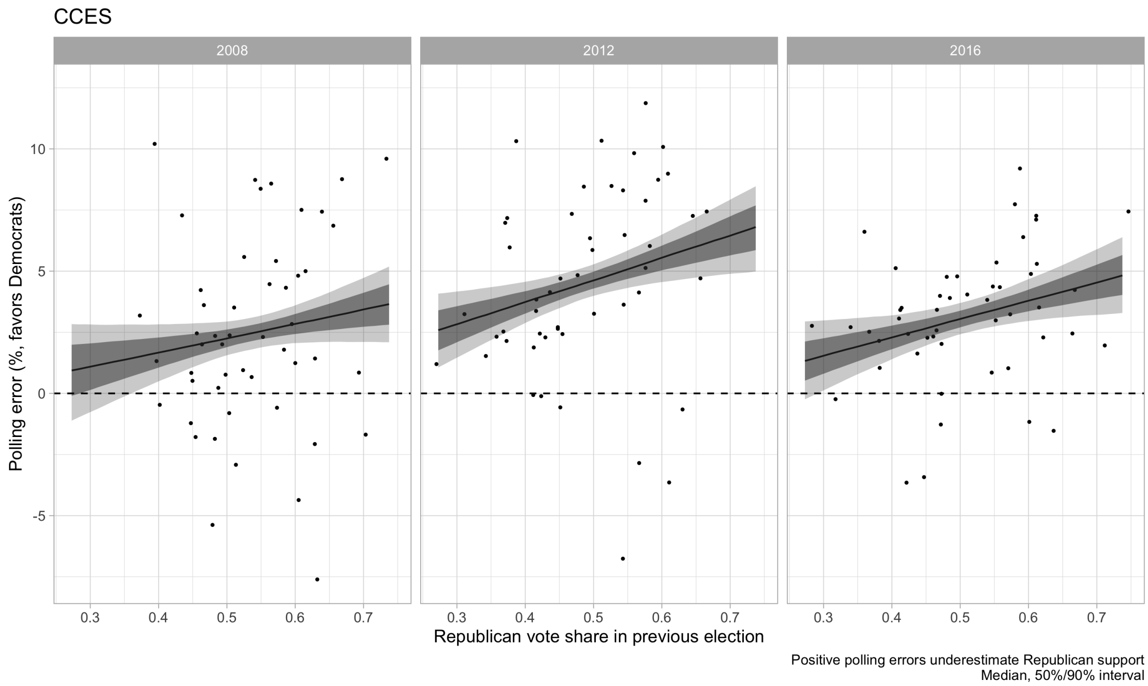
Years before 2000.

5.3 Political implications

Election forecasting

Campaigning

[1] We thank Yair Ghitza for this consideration.



References

Ansolabehere, S. (2010). CCES, Common Content, 2008 https://doi.org/10.7910/DVN/YUYIVB.

Ansolabehere, S. & Schaffner, B. (2013). CCES Common Content, 2012 https://doi.org/10.7910/DVN/HQEVPK.

Ansolabehere, S. & Schaffner, B. F. (2017). CCES Common Content, 2016 https://doi.org/10.7910/DVN/GDF6Z0.

Enns, P. K., Lagodny, J., & Schuldt, J. P. (2017). Understanding the 2016 US Presidential Polls: The Importance of Hidden Trump Supporters. *Statistics, Politics and Policy*, *8*(1).

Erikson, R. S. & Wlezien, C. (2012). *The Timeline of Presidential Elections: How Campaigns Do (and Do Not) Matter*. University of Chicago Press.

FiveThirtyEight (2021). Pollster ratings https://github.com/fivethirtyeight/data/tree/master/pollster-ratings.

Gelman, A. & Azari, J. (2017). 19 things we learned from the 2016 election. *Statistics and Public Policy*, *4*(1), 1–10.

Gelman, A., Goel, S., Rivers, D., & Rothschild, D. (2016). The mythical swing voter. *Quarterly Journal of Political Science*, *11*(1), 103–130.

Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories*. Cambridge University Press.

Gelman, A., King, G., & Boscardin, W. J. (1998). Estimating the probability of events that have never occurred: when is your vote decisive. *Journal of the American Statistical Association*, *93*(441), 1–9.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing https://www.R-project.org/.

Shirani-Mehr, H., Rothschild, D., Goel, S., & Gelman, A. (2018). Disentangling bias and variance in election polls. *Journal of the American Statistical Association*, *113*(522), 607–614.

Stan Development Team (2020). Stan Modeling Language http://mc-stan.org/.

6 Appendix

6.1 Other U.S. elections

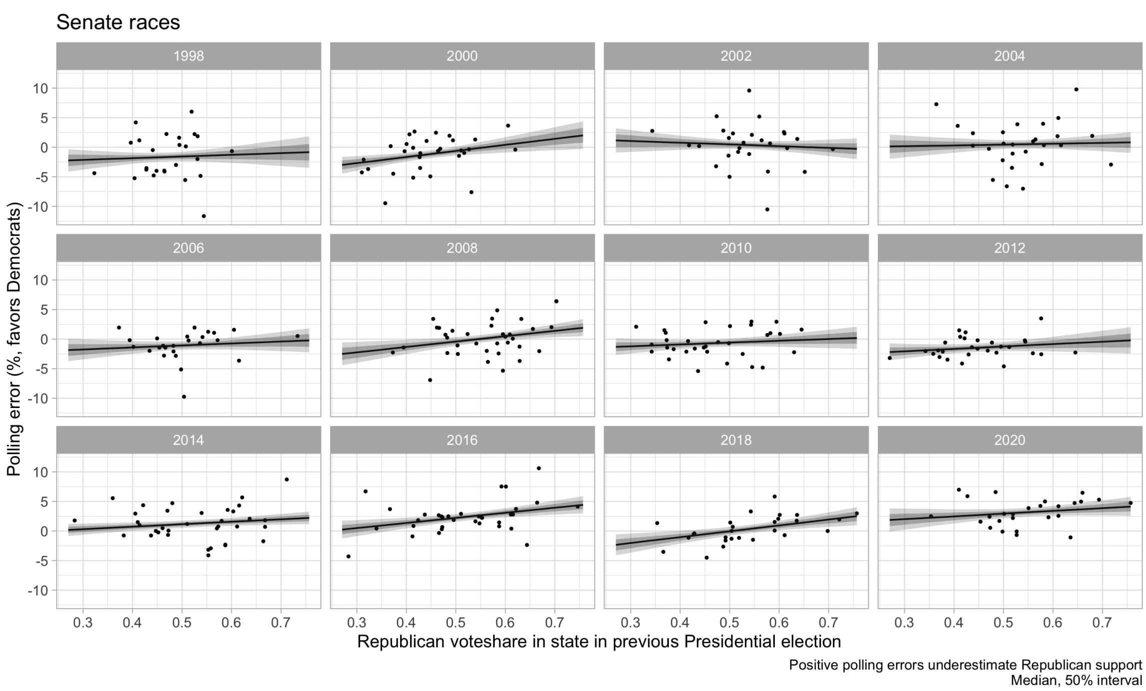


FIGURE 2: Polling error in U.S. Senate elections

6.2 Presidential elections with state labels

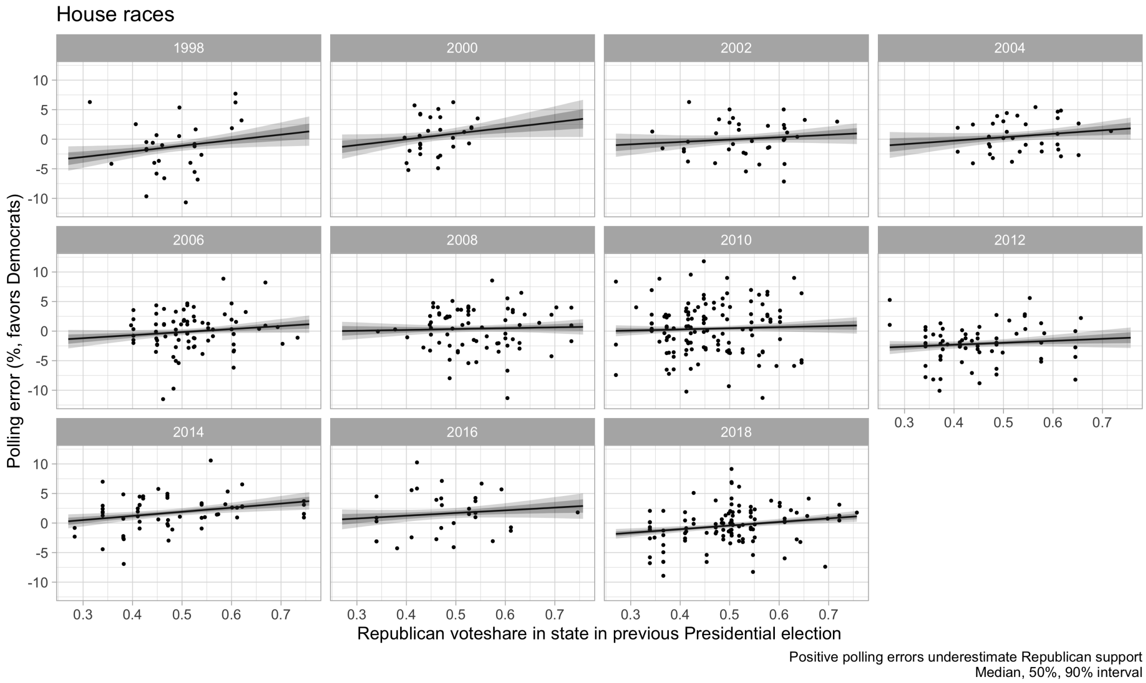


FIGURE 3: Polling error in U.S. House elections

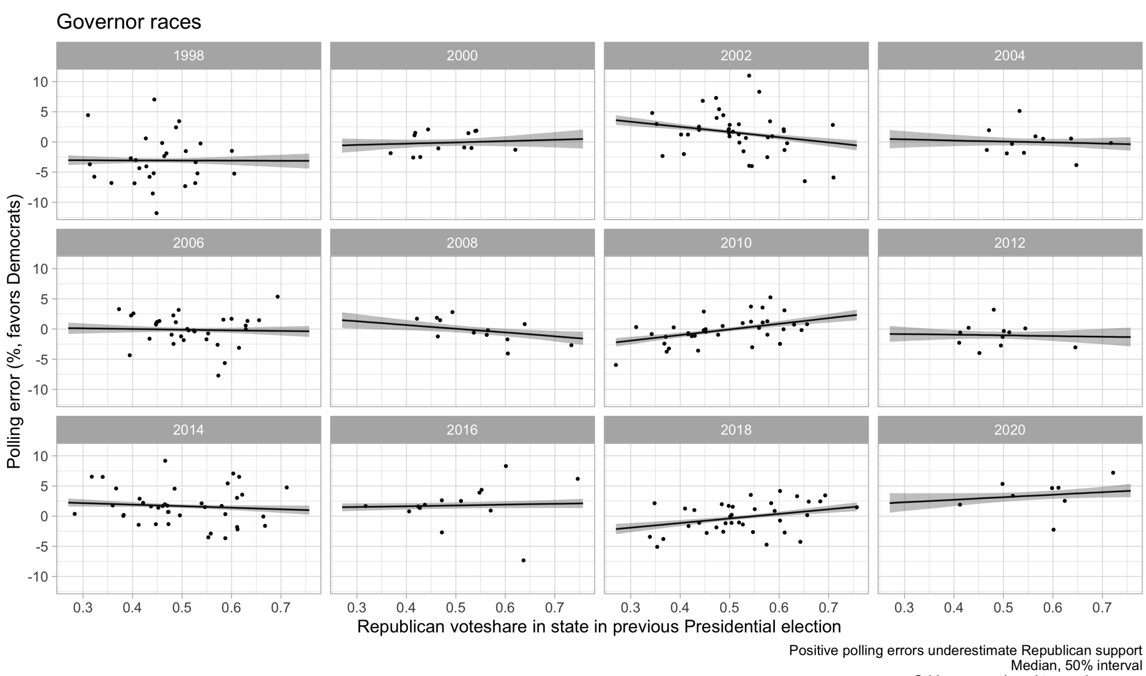


FIGURE 4: Polling error in state governor elections

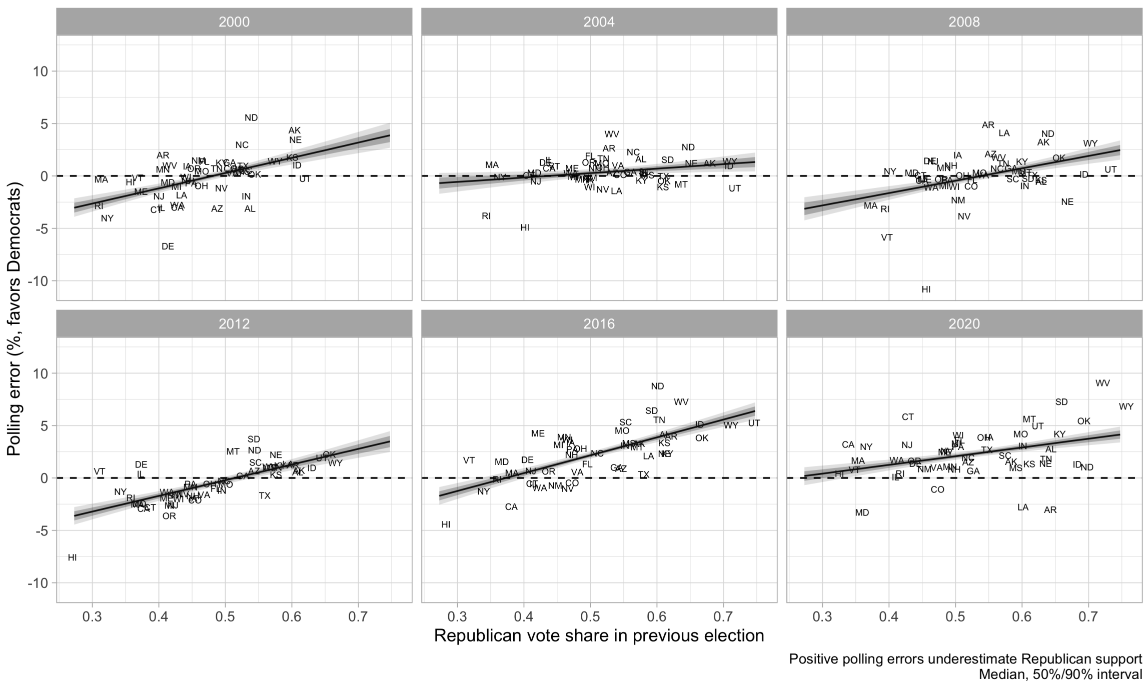


FIGURE 5: Polling error in U.S. Presidential elections with labels