Spring 4-12-2021

Aiming to Misfire: Intentional Bias and its Relationship with Polling Error

Tyler Redmon
tyler.redmon@pop.belmont.edu

Aiming to Misfire:
Intentional Bias and its Relationship with Polling Error

Tyler C. Redmon
Belmont University
Nashville, Tennessee
USA

tyler.redmon@pop.belmont.edu

Abstract: With growing public dissatisfaction with election polling, finding the source of polling error is key to resolving common frustrations. Many things have been blamed for recent polling misses but the method by which polls are conducted deserved a closer look. This paper seeks to find that source of error through polling conducted in the 2020 Presidential election. This paper shows that polling error at the national level can be attributed to intentional bias but there remain questions when it comes to state polling.

Keywords: election polling, bias, polling error, 2020 election, push polling

Introduction:

To know exactly how an election will conclude is an impossible task. Without omniscience, we settle for the next best thing: polling. Polling has been the friend and foe of prognosticators as they have attempted to accurately predict elections for the last few decades. As recently as the 2008 and 2012 Presidential elections, we saw polling being praised for its extreme accuracy at both the state and national levels. However, the recent 2016 and 2020 Presidential elections have shown the world that polling is not as infallible as it seemed. With so much perceived error in one of the few instruments we have to predict elections, we naturally reach an important question: *What is the largest contributor to polling error in Presidential elections?*

To answer this question we must first understand the factors that play a part in creating polling error. Nate Cohn, a writer for the New York Times with an emphasis on elections, wrote a post-2020 election article that breaks down a few theories as to why polling did not work perfectly in the most recent election. His analysis puts the failures of polling on three major factors: bias, methodology, and differences in state versus national polling. His findings were quite preliminary given their proximity to the election itself, but his work puts on display just how many issues there are with inconsistencies between polling and actual results. While each explanation has merit, this analysis will be looking at the differences between intentional and unintentional error in polling. State and national polling will be separated since there are such large disparities in state and national error depending on which state is considered.

The common missteps of polling are well documented over the last few decades, but I believe they are culminating in a single direction. As Edwards-Levy et al. (2017) wrote, polling is increasingly moving online. This is true nationally but especially at the state level where samples are becoming increasingly difficult to obtain. Online polls have, on balance, worse methodology and generate more error than live-interview polls. It stands to reason that the polling averages would have increased error over the last few cycles coming largely from the shift online.

The explosion of online polling has also lowered the barrier-to-entry for cheap, partisan polls. The result of this has been a flood of polls funded by campaigns and partisan organizations with worse methodology than would have existed in the past. This culmination of online polling and intentionally biased pollsters has created a dangerous concoction. Not only will these firms produce error in the direction of their favored candidate or position, but they will also have a greater range of random error caused by the online method. This would serve to explain why error has seemingly increased in recent elections, especially in some states.

Literature Review:

BIAS

Bias can take many forms in polling. One common example of pollsters putting weight on the scales is known as “herding.” This practice was observed by Walsh et al. (2009) in their evaluation of the 2000 Presidential election. They analyzed polling released just prior to the election and found that the results provided by pollsters were far too close to one another to account for the ever-existent statistical randomness. In other words, this often caused the final polling to be unrepresentative of the actual environment leading up to the election. Pollsters will

do this as a way to hedge their bets before an election. No pollster wants to badly miss the mark in their final poll, but the margin of error within polling means that at least a few of them should be off by a significant amount. As Walsh et al. explain, "our chief argument is that pre-election presidential polling is an activity more akin to forecasting next year's GDP or the winner of a sporting match than to scientific probability sampling" (316). This effect of herding can make a somewhat predictive science seem like it provides certainty, something which it cannot do. These intentional acts of data manipulation remove the ability for the proper range of outcomes to be known which complicates any ability to draw conclusions from them.

If herding is contributing to error in polling aggregates, a different way to model the polls could be a solution. In attempting to correct bias with a new model for polling aggregation, Martin et al. (2005) evaluated the polling in the 1948, 1996, and 2000 elections to find a method which produced less error than typical aggregation would. Martin et al. recognize the difficulty in separating random error and actual bias in a single poll, but find that when a wide array of polls over time are analyzed, "...twice as many polls overstated the Democratic share as understated it" (349). The researchers also cite other examples of articles which evaluated error but concluded that absolute error, instead of bias, was the typical method by which researchers analyzed polling error. It is clear from the work by Martin et al. that commonly used methods for aggregating polling fail to account for this possibility of bias largely because it is difficult to distinguish from simple error. They also concluded that, unsurprisingly, polls conducted by party-affiliated firms tend to be more biased towards the firm's contracted party which can add another layer of bias should aggregators include those firms' polls.

An older but relevant argument comes from Michalos' (1991) article on the ethics of the polling industry. Michalos feared that as the polling industry grows a larger following, their ability to influence public opinion on their own will become dangerous. One of his concerns is with polling firms which release their results without explaining any methodology including basics such as sample size, party composition, and margin of error. Michalos proposed following in the footsteps of proposed Canadian legislation requiring all firms which publish polling data publicly to also include those basic methodological facts. Given that our current system does not force firms to release this information, Michalos would argue that a large amount of bias could be easily introduced and used to change public opinion in a desired direction. As he stated, "Communities have as much right to protect themselves against intellectual pollution as they do against environmental pollution" (420).

POLLING METHODOLOGY

Beyond bias, there also exists another means by which pollsters can introduce large errors into their work: poor methodology. In his book evaluating polling from the 1980s, Crespi (1988) argues that it is, above all else, methodology which causes polls to deviate from reality. These errors come when pollsters ask leading questions, have sampling errors, or fail to accurately model expected turnout for the election. Each of these factors alone is enough to cause a poll to miss the mark. Crespi argues that while these are well known sources of error, especially after examination of early polling misses in the 1940s-50s, these issues still exist. Crespi also discusses interesting turnout models, but such conclusions lose some value when the increase in non-response to pollsters, since his book was published, has continued.

With the understanding that controlling methodology is critical for accurate polling, it makes sense that Northcott (2015) argued in his research that methodology is the key to good

polling and not the more abstract metaphysical aspects of political science. Northcott does not argue against the idea that an election is an open system with variables which cannot be controlled. However, he also does not think that this makes good polling impossible. His work found that polling is more accurate when a strong methodology is used– from phrasing of questions asked to the sample itself. In looking at the 2012 election, Northcott complimented the polling methodology that led to many prognosticators being able to accurately predict every or most-every state correctly in the days leading up to the election. The demographic weight in most polls that cycle matched well with the actual electorate which made them more accurate to reality. This work reinforces the idea that a better methodology increases the poll's accuracy which also helps explain the inverse idea (bad methodology makes for bad polling) that is explored in other works.

In a modern look at methodological issues, Campbell (2020) uses hindsight to describe the failures of election polling through a partially methodological lens. In the past the biggest issue with many polls was the sample size which was either too small or too exclusive to be representative. Campbell shows that issues with turnout modeling are a more widespread issue. After the introduction of caller-ID and the rapid increase of spam calls, people become far less likely to respond to pollsters over the phone. Even as it gets harder to get responses, the sample for a properly representative sample exists if pollsters are willing to find it. One place where it is easy to find a sample of convenience is when people are currently voting. Surveying them through exit polls is often a way for pollsters and the media to interpret the possible outcomes of the election in the days or hours before polls close. However, recent elections have proven that exit polls are often inaccurate because they tend to paint incorrect pictures of the actual electorate. On top of this, more recent issues with certain groups refusing to respond or lying to pollsters has made their job of modeling the electorate much more difficult, creating error. These incorrect responses cause pollsters to create poor representations of reality which means polling numbers differ from the population.

As technology changes, so too does polling methodology. Many pollsters have shifted from telephone interviews to online polling as gathering a sizable sample over the phone has become more difficult. Edwards-Levy et al. (2017) write that this has been a largely controversial move within the polling industry because of how samples are gathered. Normally, a sample would be conducted through probability sampling. Using probability sampling, "all persons in the target population have a known chance of being interviewed and, ideally, no one is left out" (154). Online polling cannot be conducted in this manner. Instead, online pollsters get a much larger sample size from internet responses and then weigh those responses to fit the target population. The novelty of this method likely caused some issues in the 2016 election. Because of growing costs, 2016 was a massive year for internet polling. Edwards-Levy et al. specifically point to state level polling has increased over 408% from 2012 to 2016 (155). This influx of internet polls likely accounted for the larger degree of error in state averages compared to the national average before the 2016 election. While they do not cover it because it occurred after publishing, the same effect was seen in the 2020 election as well, although even national polls were not perfect in that election.

STATE VS NATIONAL

In thinking about state level versus national polling, some researchers believe that state level indicators may actually be the key to predicting elections accurately. Berry et al. (2012)

argue that too many election predictions are made about the national popular vote when the real way to win an election is through the Electoral College. Their evaluation of the 2008 election showed that most predictions simply provided an Electoral College map estimate based on the popular vote instead of a more sophisticated state level model. In landslide elections like 2008 this is not a major problem, however, as Berry et al. argue, "It becomes problematic, however, at precisely the point that forecasts are most interesting: when elections are close" (669). If someone told the researchers of the events of the 2016 and 2020 elections, they would likely have agreed that those elections reinforce this idea. Berry et al. believe that the future of election modeling relies on accurate and in-depth predictions of outcomes in individual states instead of the simple national view. This is something they attempt by modeling the outcomes of past elections using a novel system. In short, their model of state outcomes correctly predicted the winner of every election tested even as national polls occasionally missed the mark.

Supporting the idea that state indicators are more important than national ones is Rothschild et al. (2012) in their work studying different indicators than just head-to-head polling. Their research focused on polling asking people who they expected to win the election instead of who they intended to vote for. Not only did they find this method to be more accurate than head-to-head polling for past Presidential elections, but they also discovered that this method, when applied to state level polling, was even more accurate than just using national polling. The expectation question performed 11.6% better than the typical polling in states, showing its power as a predictor for elections (8). While this work is somewhat tangential to the intended research, it still reinforces the idea that we may have the right indicators around us already to predict elections accurately and that we are just using them wrong or not using them at all.

This research generally shows that polling can fall into a few pitfalls which harm its ability to accurately predict Presidential elections. If a poll has too much bias, it will provide incorrect information. If the methodology used fails at any point, predicting based on faulty assumptions would lead to incorrect predictions that could not be falsified until the election occurs. Even the general use of national polling instead of state polling could explain the general hesitation to rely on polling these days, but there are also concerns with online polling which is widely practiced in state level polling. These sources provide a guide that will help create measures which can show where and with what magnitude error is introduced into polling.

Model:

Even with all of these theories, we still do not have a clear answer as to what factor contributes most to polling error. However, I do think that this research, especially the more contemporary works, point in the direction of bias playing the largest part. While it is easy to simply blame bias for the introduction of error, that explanation alone does not tell the full story. More interestingly, if bias is the largest contributor to error, is this bias intentional or unintentional? If both exist, which is the larger factor?

This model theorizes that bias, systematic error which researchers are themselves introducing, exists in polling and that bias can be differentiated by intentionality. There is a clear difference between a university conducting a poll with poor methodology because it is their first time and a partisan firm using faulty methods with an intended result. The reason either of these types of polls even exist is largely due to a lowering of the barriers-to-entry for pollsters. Much like how our media landscape used to be dominated by three networks but is now littered with hundreds of options of varying accuracy, the polling industry too has seen a rapid growth in

unorthodox and untraditional firms conducting polls. Online polling, for instance, saw its share of total polls conducted jump from 18% to 41% nationally and from 10% to 44% at the state level between 2012 and 2016 (Edwards-Levy et al. 2017). The increase in online polling and its availability has made it possible for any firm to conduct polling beyond those polls sanctioned by major polling institutions. Due to this, the general reliability of the body of polling has likely gone the same way the body of media work has since the growth in social media and fringe news sites.

This information is what leads me to postulate that intentional bias will play a larger part in error than unintentional bias. Using the previous example, if a university conducting its first poll finds that its results are far from reality, they have the opportunity and inclination to adjust and produce better polling in the future. However, if a firm has malice in the creation of poor methodology, they have no incentive to represent reality accurately and instead would see a partisan result as successful even if it is error prone. If this is the case, the results would show a larger difference in the means of error between intentionally biased polling and unbiased polling than unintentionally biased polling would show. However, if there is no significant distinction between these means then it would force us to reevaluate where error comes from and look for alternative solutions.

Method:

The cases for this analysis will be polls conducted in the closing weeks of the 2020 Presidential election. These polls will consist of the final polling conducted by a diverse set of firms which exist within the categories of intentional, unintentional, or no bias as well as differentiating between state and national polling to control for any error caused by the many factors there. This data will be gathered from polling aggregate sites such as *RealClearPolitics* and *FiveThirtyEight* with others taken directly from the websites of the polling firm. The aim is to gather as many polls as possible conducted in late October and early November as those polls would, in theory, be the closest approximations to compare with the actual electorate. This should provide an adequate number of cases in each category so that the data collected can be significant. By selecting data in this way, enough information is available to back up the hypothesis that intentional bias creates the largest amount of error as opposed to unintentional bias or simply no bias whatsoever. This data will also be discussed further for demonstrative purposes to show that many of the firms exhibiting intentional bias have the capability to produce unbiased results and thus are making the conscious decision to introduce bias themselves.

Each poll considered will be evaluated on the basis of the source of bias. If a poll is conducted by a partisan firm, on behalf of a political candidate, or on the behalf of a partisan group then any methodological issues will be considered intentional. This classification of intentional bias has to be created in this way because pollsters do not reveal when they have introduced bias on purpose and this intent has to be assumed from outside factors as opposed to outright confession. As for unintentional bias, a poll is considered methodologically poor if it receives a grade of lower than a C+ from the FiveThirtyEight Pollster Ratings. These grades are based on associations with polling agencies and the method by which a sample is collected. A rating below C+ shows enough evidence that a poll suffers from large methodological issues and thus bias is introduced.

If no rating exists for a pollster, they will be catalogued as methodologically poor if they are not a member of National Council of Public Polls, the American Association for Public Opinion Research's transparency initiative, or the Roper Center for Public Opinion Research's data archive as well as not conducting the poll by live interview (landline or cellphone). If a firm receives a score of C+ or greater from FiveThirtyEight or meets the previously listed criteria, they are considered methodologically sound and thus not biased. Polls are also split based on whether they are state or national polls as a way to control for greater amounts of error that exist in state polling as opposed to national polling. This should allow for consistent comparisons between like polls. Finally, error should be considered the absolute value of the difference between the predicted margin of victory for Joe Biden in any given poll compared to his actual margin of victory, national or state, in the election. Therefore, if a poll reported the final national head-to-head matchup as being 53% Biden and 45% Trump, the error would equal that margin of 8% minus the actual margin of 4.5%, meaning the error, or distance from reality, is 3.5%.

Taking this data, a one-way independent ANOVA test will be run with the error as the dependent variable and bias (separated into three categories) will be the independent variable. Since state level polling is much more likely to introduce outliers and large errors, this test will be run once with national polling and again with state level polling so that error can be consistently measured. By doing this, it becomes visible what impact each of the three categories of bias have on the average of polling error. A statistically significant difference between intentional bias and the other categories will work to eliminate the null hypothesis and thus support the idea that intentional bias plays the largest part in error. If bias is consistent no matter the intentionality, we would see consistent results between intentional and unintentional bias but a difference from unbiased polls. If bias does not contribute at all to error, we would not see any statistically significant difference between the three categories and would then have to conclude that bias is not the reason for larger polling error in the 2020 Presidential election. Also, if this effect is shown to be different between state and national polling we could then draw separate conclusions about each one which could require further research in the future.

Results:

NATIONAL

The results of the ANOVA of polls from the 2020 Presidential election show a clear relationship between the level of bias from the polling firm and the average error these polls have at the national level. Also shown below are the descriptives for the national poll cases. This establishes the mean error for the sample of polls based on the three categories of pollsters. While the category of unbiased polls had a mean error of 3.2%, something within the expected margin of error for most national polls, the average error for the other categories is higher. Unintentional bias sits at an average of 4.6% error while intentional bias comes in at a hefty 5.6% error, beyond the typical margin for national polls.

ANOVA

Error

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	39.753	2	19.877	3.340	.046
Within Groups	232.080	39	5.951		
Total	271.833	41			

Descriptives

Error

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Unbiased	16	3.188	1.7017	.4254	2.281	4.094	.5	6.5
Unintentional Bias	15	4.633	3.1593	.8157	2.884	6.383	.5	9.5
Intentional Bias	11	5.591	2.2115	.6668	4.105	7.077	2.5	9.5
Total	42	4.333	2.5749	.3973	3.531	5.136	.5	9.5

These numbers alone do not tell us much about the relationship between the error and if it is significant or simply random difference. However, the comparison shown below illustrates where significant differences in the mean error exist. The initial ANOVA shows a significant difference between groups with p = .046. While looking at the lower and upper bounds in the descriptives gives us some hints as to where the relationship exists, running a Tukey post hoc test will break down the significance by each categories' relationship with one another. The follow up Tukey post hoc reveals that the significance exists in the difference between intentionally biased polls and unbiased polls (p = .042). There are no other significant relationships between other categories. This result allows us to reject the null hypothesis that there is no relationship between the category which a poll is placed in and the average error those national polls generated in the 2020 Presidential election.

Multiple Comparisons

Dependent Variable: Error

Tukey HSD

(I) Bias Ranking	(J) Bias Ranking	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Unbiased	Unintentional Bias	-1.4458	.8767	.238	-3.582	.690
	Intentional Bias	-2.4034*	.9555	.042	-4.731	-.076
Unintentional Bias	Unbiased	1.4458	.8767	.238	-.690	3.582
	Intentional Bias	-.9576	.9683	.588	-3.317	1.402
Intentional Bias	Unbiased	2.4034*	.9555	.042	.076	4.731
	Unintentional Bias	.9576	.9683	.588	-1.402	3.317

*. The mean difference is significant at the 0.05 level.

STATE

Based on this result, one might expect that a similar relationship would exist in state polling as well. However, that was not found to be the case. As is shown below, there was actually a reverse in which polls showed the most error. Unbiased and unintentionally biased state polls had similarly above average error in this election while it was actually intentionally biased polls that obtained the lowest score for mean bias. Intentionally biased polls performed .9% better than unbiased polls at the state level as opposed to performing 2.4% worse on average at the national level. Importantly, the ANOVA shows no significant relationship between the groups with a p = .482.

ANOVA

Error

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	11.066	2	5.533	.737	.482
Within Groups	675.942	90	7.510		
Total	687.008	92			

Descriptives

Error

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Unbiased	59	4.544	2.8847	.3756	3.792	5.296	.3	12.2
Unintentional Bias	16	4.550	2.6351	.6588	3.146	5.954	1.8	12.2
Intentional Bias	18	3.672	2.2898	.5397	2.534	4.811	.3	9.2
Total	93	4.376	2.7327	.2834	3.814	4.939	.3	12.2

I ran a Tukey post hoc test on these variables as well. While it is already clear that no relationship exists between the groups, being consistent with the use of a Tukey post hoc may still give us a greater window into why there is no significant relationship.. The follow up analysis, as expected, also fails to show any relationship between any of the categories at the state level. This means that the null hypothesis cannot be rejected at the state level in the same way that it was for national polling. There is no significant difference in means for state level polling so no clear relationship can be shown at the state level. For unbiased and unintentionally biased polling, the significance of their differences is p = 1.000. This is an emphatic inability to reject the null hypothesis as it relates to methodology and error at the state level. While there was more significance between intentionally biased polling and the other categories, it is nowhere as significant as the ANOVA showed at the national level. All of the null hypotheses for error

remain standing as none can be rejected.

Multiple Comparisons

Dependent Variable: Error

Tukey HSD

(I) Bias Ranking	(J) Bias Ranking	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Unbiased	Unintentional Bias	-.0059	.7725	1.000	-1.847	1.835
	Intentional Bias	.8718	.7379	.467	-.887	2.630
Unintentional Bias	Unbiased	.0059	.7725	1.000	-1.835	1.847
	Intentional Bias	.8778	.9416	.621	-1.366	3.122
Intentional Bias	Unbiased	-.8718	.7379	.467	-2.630	.887
	Unintentional Bias	-.8778	.9416	.621	-3.122	1.366

Conclusions:

Based on the results of the two tests, the null hypothesis was rejected at the national level but not at the state level. What this tells us is that there is some kind of relationship between the level of bias in polling and the error generated by those polls. While the difference between unbiased polls and unintentionally biased polls was not significant, there was still a visible difference in the mean values that trended in the same direction as the actually significant result. In thinking of polling aggregates, the inclusion of intentionally biased polls would have a negative effect on the aggregate accuracy by increasing the mean error of the population of polls. This means that for national polls in 2020, the best option was to at the very least remove intentionally biased polls and to at least consider if unintentionally biased polls ought to be included.

The state level results were quite interesting in that they failed to parallel the national polling. While it might have sounded logical to assume intentionally biased polls would perform worse in an election, that simply was not the case for state level polling in the 2020 Presidential election. There was no statistically significant difference between the means for error between any category. In fact, intentionally biased polls performed better on average than the two other categories which were basically identical. This means there has to be some key point of divergence between national and state polling in this election. It could be that there was more systematic error across the board in all sets of state polling. While national polls only missed the mark on average by less than 3%, some states like Ohio, Wisconsin, and Iowa saw much larger misses over 6%. It makes sense that intentionally biased polls were more accurate in some of these states than they were nationally since many of the intentionally biased polls were biased in favor of President Trump. While these polls performed poorly in the national popular vote where general error was low, they fared far better in high error state averages. Since the polls with better methodology missed the mark by so much in these states, polls which intentionally skewed similar data in favor of Trump would have looked closer to reality even if that was not their intent.

Nate Silver's (2020) after action report on polling in the election argued that a mix of systemic and idiosyncratic factors led to state polling being more error prone in certain states than the nation as a whole. For example, the Midwest saw greater error overall than the South

and Southwest which could contribute to swing state polling error. Since more polls were conducted in many of these swing states, their error looks outsized beyond the national error that took place. On top of that, individual factors dealing with COVID-19 could have altered likely polling respondents on a state-by-state basis which would have shown up as greater error at the state level than nationally. Due to increased travel restrictions in some states versus others, the demographics of those sampled could change on a state-by-state basis. With some workers being both at home and available to answer calls while other essential workers could not be home, there could be a real divergence between live-call polls and other methods. Thus, certain types of workers could be over or under sampled depending on each state's regulations on workers during the pandemic. For example, Georgia polls may have been more accurate than Michigan polls because the sample available in Georgia suffered less from changes due to the pandemic due to fewer COVID-19 restrictions on work and travel. As for the error generally, Silver stated that, while pollsters attempted to correct for the mistakes of the 2016 Presidential election, pollsters did not correct enough for existing mistakes and spiking turnout caused some turnout projections to misrepresent the population.

With these conclusions in mind, future research ought to be done to discover exactly why national and state polling differed in their mean error in the 2020 Presidential election. Work could be done to evaluate those states which produced the greatest error on average (Wisconsin, Ohio, ect.) to see which factors affected their polling and not those of states with lesser error (Georgia, Arizona, ect.). Also, the emergence of low-propensity voters in this election could have had a greater effect on more elastic states than less elastic ones. Comparing elasticity of the electorate in each state along with their error in 2020 could provide some insights here. States with greater or lesser COVID-19 exposure or restrictions could be different from the norm as well and that could be examined in future work.

Another key aspect of these findings that needs explanation is exactly why pollsters would intentionally attempt to miss reality. At the national level, intentionally biased pollsters saw errors greater than 5% on average. It is not simply the case that these pollsters are just bad at their job; many of the pollsters which were contracted to produce results for candidates or political organizations also later conducted independent surveys which were closer to reality and displayed fewer methodological issues. More qualitative work could be done to discover why a polling firm with the full knowledge of how to conduct a proper poll would skew their data in such a way. Is the bias introduced because the client asked for it to represent their ideal reality or do polling firms skew the numbers without being asked as a way to earn more favor and business with a particular group? No matter what the answer, these inaccurate polls cause averages to be less accurate and increase mean error at the national level. As Berry et al. (2012) states, polling accuracy matters most when elections are close. In our polarized environment where almost every national election is decided by narrow margins, eliminating error where we can is critical to enhancing the polling industry's reputation with the American people.

A darker conclusion can also be drawn from this data. With such a stark difference between the results at the national and state level, what factor could possibly cause something like this? One likely culprit is push polling. To the average voter, push polling can appear like a normal, scientific poll. A person or online prompt from an official sounding source asks questions about their political stances. However, these polls are twisted perversions of the typical practice (Connelly 2014). A push poll is filled with negative information about a particular candidate or movement framed to look just like a regular polling question. Take this Public Policy Polling question from the 2020 North Carolina election as an example:

" Some people, including conservative Fox News host Tucker Carlson, have said Richard Burr's decision to sell his stocks is illegal insider trading and that he should resign immediately. Burr says, in his defense, that he relied solely on public media reports to guide his decision making regarding the sale of stocks. Having heard this, let me ask you again: Do you think Senator Burr should resign, or not?"

The purpose of a question like this is not to reach an objective understanding of the public's opinion on an issue; instead, these push poll questions are designed to change the minds of voters in a particular direction. Approximately 2% of previously undecided voters shifted their position, almost exclusively against Senator Burr, after the information was provided. An especially important point is the appeal to Republican voters through the use of Tucker Carlson, a nationally prominent conservative voice. This push caused a net 3% change in favor of Senator Burr's resignation among just Republican voters, creating a plurality of Republicans in favor of resignation (Public Policy Polling 2020). The goal of this poll was to persuade more voters to want Senator Burr's retirement, not to accurately measure existing sentiment.

This information, paired with the results that intentionally biased polling was no less accurate than other methods at the state, draws us closer to an ominous conclusion. Could it be possible that intentionally biased firms were able to change reality by using push polling questions? This conclusion would certainly explain why there was a difference in the national and state level results. Intentionally biased polling is far more common at the state level and the populations there are much easier to target than a national audience. Considering these polls are not a member of any national council or organization meant to protect voters from these kinds of questions, the opportunity is there. It is also important to consider that intentionally biased firms are more likely to keep their polls hidden from the public since they are almost exclusively commissioned by an outside group or candidate. This means that we are left unaware of both the questions they may ask and the true quantity of polling they conduct during an election cycle. This veil solidifies Michalos' (1991) worst fears about what polling can be used to accomplish. If intentionally biased pollsters have found a means by which they can alter the reality on the ground in important states then there is a dangerous possibility of this behavior escalating in the future. More research ought to be done in this area before the 2024 Presidential election if we can hope to better understand the true cause of the disparity between state and national polling. Otherwise, this unethical practice could continue to poison polling in the United States for the foreseeable future.

Work Cited

Berry, M., & Bickers, K. (2012). Forecasting the 2012 Presidential Election with State-Level Economic Indicators. *PS: Political Science and Politics, 45*(4), 669-674.

Campbell, W. Joseph. (2020). *Lost in a Gallup: Polling Failure in U.S. Presidential Elections*. University of California Press.

Cohn, Nate (2020). "What Went Wrong With Polling? Some Early Theories". New York Times.

Connelly, M. (2014, June 18). "Push Polls, Defined". https://www.nytimes.com/2014/06/19/upshot/push-polls-defined.html

Crespi, I. (1988). *Pre-Election Polling: Sources of Accuracy and Error*. Russell Sage Foundation.

Edwards-Levy, Ariel, Natalie Jackson, and Janie Velencia (2017). "Polling in the 2016 Election". *Trumped: The 2016 Election that Broke All the Rules* 152-166. Rowman & Littlefield.

FiveThirtyEight. (2020, May 19). Pollster ratings. https://projects.fivethirtyeight.com/pollster-ratings/

Martin, E., Traugott, M., & Kennedy, C. (2005). A Review and Proposal for a New Measure of Poll Accuracy. *The Public Opinion Quarterly, 69*(3), 342-369.

Michalos, A. (1991). Ethical Considerations regarding Public Opinion Polling during Election Campaigns. *Journal of Business Ethics, 10*(6), 403-422.

Northcott, R. (2015). Opinion Polling and Election Predictions. *Philosophy of Science, 82*(5), 1260-1271.

Public Policy Polling (2020). Majority Think Burr Should RESIGN; Tillis Unpopular Too. https://www.publicpolicypolling.com/polls/majority-think-burr-should-resign-tillis-unpopular-too/

RealClearPolitics. Election 2020 - General Election: Trump VS. Biden. https://www.realclearpolitics.com/epolls/2020/president/us/general_election_trump_vs_biden-6247.html

Rothschild, D. M., & Wolfers, J. (2012). Forecasting Elections: Voter Intentions Versus Expectations. *SSRN Electronic Journal*.

Silver, Nate. (2020). "The Polls Weren't Great. But That's Pretty Normal". https://fivethirtyeight.com/features/the-polls-werent-great-but-thats-pretty-normal/

Walsh, E., Dolfin, S., & DiNardo, J. (2009). Lies, Damn Lies, and Pre-Election Polling. *The American Economic Review, 99*(2), 316-322.