HDSR
HARVARD DATA SCIENCE REVIEW

Volume 0 Issue 0
DOI: 00.000
ISSN: 2644-2353

Statistical Paradise Lost? Persistent Non-Response Bias in the 2024 US Presidential Election

Anonymous Authors

Abstract. In the 2024 US presidential election, the polls showed a slight edge for Kamala Harris, but Donald Trump won all seven swing states and the popular vote. This discrepancy between polls and actual elections stems from the selection mechanism of the poll surveys, which may not be on par with the gold standard of simple random sampling (SRS). Previous work has shown the existence of non-response bias in the polls for the 2016 US presidential election and measured its magnitude and direction. This led to improved polling practices leading up to the 2024 US presidential election. We applied the same methodology to the 2024 election and found persistent non-response bias of a similar direction. We found that the effective sample sizes of polls suffered >99% percentage reduction from their original sample sizes. We propose a bias-corrected estimator for future elections and evaluated it on the 2024 election, showing that it reduced RMSE by half.

1. Introduction

Given the global significance of the U.S. presidential elections, polling has emerged as a highly sought-after method for forecasting election outcomes, with 1,702 polls conducted for the 2024 election (FiveThirtyEight, 2024). However, polls are estimators of the actual voting results and have errors. Some of these discrepancies between polls and actual elections can be attributed to the selection mechanism of the poll surveys, where the gold standard of simple random sampling (SRS) is often practically infeasible. Previous work has shown the existence of non-response bias in the 2016 US presidential election polls and measured its magnitude and direction (Meng, 2018). Scholars have debated that non-response bias could be particularly pronounced in 2016, as Donald Trump was an unconventional presidential candidate and voting for him could be seen as socially undesirable, an effect known as the "Shy Trump Supporter" hypothesis (Coppock, 2017). After serving a term, Donald Trump lost re-election in 2020, and ran again in 2024. The polling industry had eight years between 2016 and 2024, with a presidential election in-between, to improve on their polling techniques. In the 2024 US presidential election, pollsters predicted a slight edge for Democratic candidate Kamala Harris, and an aggregation of 25 polls predicting a 1.6% Democratic lead (270toWin, 2024). Instead, Trump won all seven swing states with a 1.5% lead in two-party popular vote share (Federal Election Commission, 2025), contrary to the forecast of the pollsters, as in 2016.

In this paper, we applied the same methodology from Meng, 2018 to the 2024 US presidential election and found that a non-response bias of similar direction occurred in the polls (Section 2.1). As in 2016, we found that errors are magnified under large population sizes, with larger errors for larger states (Section 2.2). Furthermore, we found evidence of the Big Data Paradox, where bigger samples tend to escape conventional confidence intervals (Section 2.3). We also computed the effective sample sizes of the polls: the SRS-equivalent sample size that would yield the same mean squared error (MSE) (Section 2.4). We found that the polls for the biggest states suffered >99% percentage reduction from their original sample sizes. As the 2024 election polls showed that it is stubbornly difficult to remove non-response bias from the polls directly, a computed data defect correlation can serve as a helpful prior for future elections where we expect a similar prevalence of non-response behavior, thereby improving polling accuracy. We propose a bias-corrected estimator in Section 2.5, using historical data defect correlations to improve future polling estimates. We evaluated it on the 2024 election, finding that it reduced RMSE by half (Section 3.5).

2. Methods

Following the setup of Meng, 2018, we conduct our analysis on the Common Content Dataset from the Cooperative Election Study (CES) (Schaffner et al., 2025). The data includes a nationally representative sample of 60,000 American adults interviewed before and after the 2024 elections. The CES study employs a sample-matching methodology, which samples target samples from a target population, then matches an opt-in survey respondent who has the closest attributes to a person in the target sample.

2.1. **Computing the data defect correlation.** The actual error between the sample mean and the population mean can be fundamentally decomposed into three factors through an equation we term as the "Meng identity" (Equation 2.3 in Meng, 2018):

$$\bar{G}_n - \bar{G}_N = \underbrace{\rho_{R,G}}_{\text{Data Quality}} \times \underbrace{\sqrt{\frac{1-f}{f}}}_{\text{Data Quantity}} \times \underbrace{\sigma_G}_{\text{Problem Difficulty}} \tag{2.1}$$

where

$\bar{G}_n$ is the sample mean, e.g. the share of respondents who prefers Trump

$\bar{G}_N$ is the population mean, e.g. the vote share of Trump

$\rho_{R,G}$ is the **data defect correlation** e.g. the correlation between responding to the survey and voting for Trump

f is the ratio of sample size to population size

σ_G is the inherent standard deviation of G.

Before the election, we only have estimate $\bar{G}_n$. After the election, we will have observed $\bar{G}_N, f, \sigma_G$. We can then proceed to calculate the data defect correlation.

$$\rho_{R,G} = \frac{\overline{G}_n - \bar{G}_N}{\sigma_G \sqrt{\frac{1-f}{f}}} \tag{2.2}$$

In the results (Section 3.1), we plot a histogram of the data defect correlation across all states for both Trump and Harris. We assume that the data defect correlation is i.i.d. across states, so we expect a scenario with negligible non-response bias to have a distribution around 0, while a scenario with constant non-zero non-response bias to have a distribution centered away from zero, as found in Figure 5 of Meng, 2018.

2.2. **Testing for the Law of Large Populations.** One way to quantify the error under non-response bias is to report it in units of error under an otherwise perfect SRS. $\bar{G}_n$ under SRS is unbiased, so its mean-squared error (MSE) is the same as its variance:

$$\mathbb{V}_{\text{SRS}}(\bar{G}_n) = \frac{1-f}{n}\frac{N}{N-1}\sigma_G^2, \tag{2.3}$$

Combining the SRS error in Eq 2.3 with Meng identity in Eq 2.1 , we obtain

$$Z_{n,N} \equiv \frac{\bar{G}_n - \bar{G}_N}{\sqrt{V_{\text{SRS}}(\bar{G}_n)}} = \sqrt{N-1}\,\rho_{R,G}. \tag{2.4}$$

The dependence of $Z_{n,N}$ on $\sqrt{N}$ under non-response bias is known as the Law of Large Populations in Meng, 2018:

> **Law of Large Populations (LLP)** Among studies sharing the same (fixed) average data defect correlation $\mathbb{E}_R(\rho_{R,G}) \neq 0$, the (stochastic) error of $\bar{G}_n$, relative to its benchmark under SRS, grows with the population size N at the rate of $\sqrt{N}$.

To test for this relationship, we have to find evidence of a square root dependence on N. Taking logs on both sides of Eq 2.4 yields

$$\log|Z_{n,N}| = \log|\rho_{R,G}| + 0.5\,\log\big(N-1\big) \approx \log|\rho_{R,G}| + 0.5\,\log N. \tag{2.5}$$

Using the post-election error to compute $Z_{n,N}$ through the left-hand side of Eq 2.4 and plotting it against $\log N$, we expect to see a gradient of 0.5 if $\rho_{R,G}$ stays roughly constant.

2.3. **Lower Confidence in Bigger Data with the Big Data Paradox.** We begin by constructing the confidence intervals for $\bar{G}_n$ by assuming a Normal distribution for the error, and finding Z_n, the z-score. Note that the Z_n here is distinct from $Z_{n,N}$ in Section 2.2.

$$Z_n = \frac{\bar{G}_n - \bar{G}_N}{\sqrt{\bar{G}_n(1-\bar{G}_n)/n}} = \frac{\sqrt{n}\,\sqrt{D_O}\,\rho_{R,G}}{\sqrt{1 - D_O\,\rho_{R,G}^2} - \sqrt{D_O}\,\rho_{R,G}\left(\sqrt{\frac{\bar{G}_N}{1-\bar{G}_N}} - \sqrt{\frac{1-\bar{G}_N}{\bar{G}_N}}\right)}. \tag{2.6}$$

where $D_O = \frac{1-f}{f}$. For tight races ($\bar{G}_N \approx 0.5$), the second term in the denominator is close to zero, yielding

$$Z_n \approx \sqrt{n} \cdot \frac{\sqrt{D_O}\rho_{R,G}}{\sqrt{1 - D_O\rho_{R,G}^2}} \tag{2.7}$$

In this case, if we hold the sample ratio f constant, and thus D_O constant, and when $\rho_{R,G}^2$ is non-zero and does not depend on n, we have that Z_n scales by $\sqrt{n}$. This $n^{1/2}$ relationship is in stark comparison to the usual confidence interval that scales with $n^{-1/2}$ under SRS. Consequently, when the selection mechanism differs from that of SRS and there is a correlation between selection and the type of response, we predict that large samples will fall outside typical confidence intervals. This is known as the *Big Data Paradox* in Meng, 2018.

The Big Data Paradox: The bigger the data, the surer we fool ourselves.

2.4. **Computing the effective sample sizes.** Another way to quantify the error is to calculate the *effective sample size*, which is the sample size under SRS that will result in the same observed error in expectation. To solve for the effective sample size, we equate the mean-squared error (MSE) of $\bar{G}_n$ to the error under SRS using an effective sample size n_{eff}:

$$\mathbb{E}_R[\bar{G}_n - \bar{G}_N]^2 = \mathbb{V}_{SRS}(\bar{G}_{n_{\text{eff}}}) \tag{2.8}$$

Using Jensen's inequality and the assumption that $\mathbb{E}_R[\rho_{R,G}^2] \leq \frac{f}{1-f}$, Meng, 2018 showed that

$$n_{\text{eff}} \leq \frac{1}{\mathbb{E}_R[\rho_{R,G}^2]} \frac{f}{1-f} \tag{2.9}$$

Using the range of sample ratios (0.0002 to 0.0006) in CES, the assumption becomes $\mathbb{E}_R[\rho_{R,G}^2] \leq \frac{0.0002}{1-0.0002} \approx 0.0002$. In the results, we will show that all data defect correlations computed satisfy $\rho_{R,G}^2 \leq 0.0001$, and the assumption holds comfortably. In Section 3.4, we compute the upper bound for the effective sample sizes of each state and compute the percentage reductions in sample sizes.

2.5. **Saving the Polls with Prior-Informed Bias Correction.** We propose a new estimator of vote share by applying bias correction to the the sample average. Using the Meng identity (Eq 2.1), we can theoretically construct an unbiased estimator

$$\bar{G}'_n = \bar{G}_n - \rho_{R,G} \times \sqrt{\frac{1-f}{f}} \times \hat{\sigma}_G \tag{2.10}$$

This is ideal, but we do not have access to $\rho_{R,G}, f, \sigma_G$ before the election. We estimate each of the unknown quantities in turn. We estimate $\rho_{R,G}$ using the average data defect

correlation in 2016, $\bar{\rho}_{R,G,2016}$. We estimate f by using the turnout of each state in 2016, t_{2016}, and apply that to the voting-eligible population in 2024, $N_{\text{VEP},2024}$.

$$\hat{f}_{2024} = \frac{n}{N_{\text{VEP},2024} \cdot t_{2016}} \tag{2.11}$$

Finally, we estimate the problem difficulty σ_G by taking the sample standard deviation $\hat{\sigma}_{G,2024}$ from the polls. This yields our estimator in Equation 2.12.

$$\bar{G}_n^* = \bar{G}_n - \bar{\rho}_{R,G,2016} \times \sqrt{\frac{1 - \hat{f}_{2024}}{\hat{f}_{2024}}} \times \hat{\sigma}_{G,2024} \tag{2.12}$$

Note that this estimator can be computed pre-election, and more sophisticated estimators for each of the components ρ, f, σ could be possible. We evaluated the performance of this estimator in Section 3.5. We used the datasets from the Election Lab at the University of Florida for the 2016 turnout (McDonald, 2023) and for the voting-eligible population in 2024 (McDonald, 2024).

3. Results

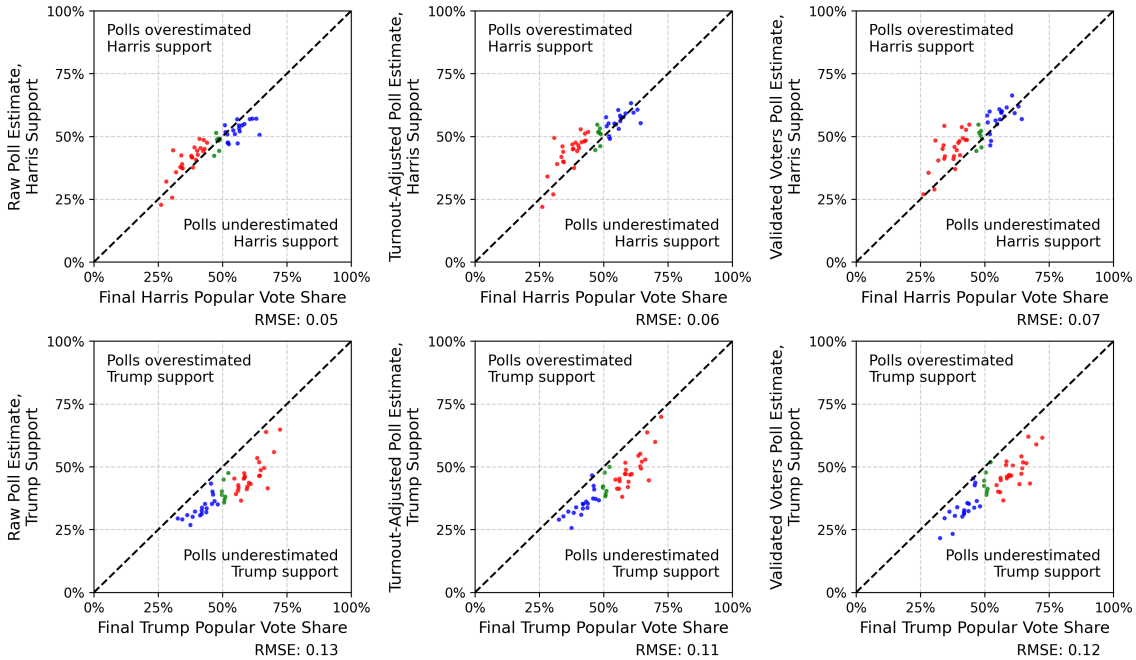


Figure 1. Poll estimate vs actual vote share for Harris and Trump for each state, using raw polls and turnout-adjusted polls. Red states are solid or likely Republican states, blue states are solid or likely Democrat states, green states are tossup states. Classifications are from the 2024 CPR Electoral College Ratings (Cook Political Report, 2024). Details on turnout estimation can be found in Section E in the appendix.

We first recreate Figure 4 from Meng, 2018 for the 2024 election in Figure 1. We see that for Harris the poll estimates are roughly in line with the actual vote share, with a RMSE of 0.05, 0.06, 0.07 for raw polls, turnout-adjusted polls and validated voter polls respectively. For Trump, there is visible underestimation even after accounting for voter validation, with this pattern holding across all states. The RMSE is higher at 0.13, 0.11 and 0.12. This points towards the existence of non-response bias from Trump voters that is more serious than non-response bias from Harris voters. In other words, voters who vote for Trump are less likely to be selected to answer the survey.

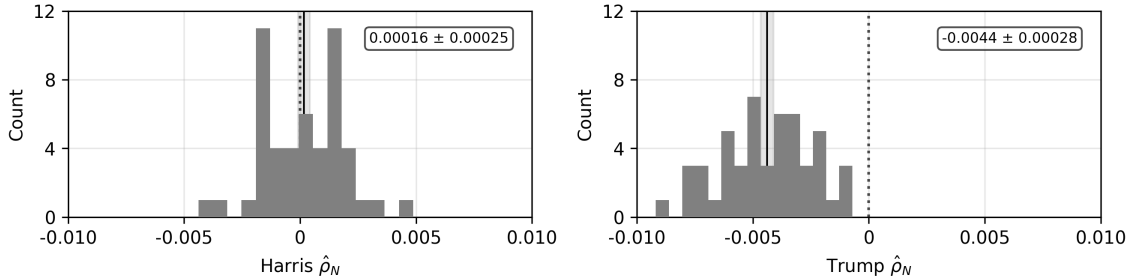
3.1. **Computing the data defect correlation.** Using Eq 2.2, we plot the binned histogram of the data defect correlation for each state in Figure 2 using three subsets of data: all respondents, likely voters, and validated voters. Likely voters are same as those used for turnout-adjusted estimates in Figure 1. Using all respondents, there appear to be no data defect correlation for Harris, while Trump has a data defect correlation centered around -0.0044 (similar to the -0.0045 as observed in Meng, 2018). However, after correcting for turnout (either estimating it pre-election or using validated voters post-election), the data defect correlations for Harris moved away from non-zero to a positive 0.0013 and for Trump the data defect correlations decreased in magnitude to -0.003. In all three cases, the selection mechanism in CCES manifested non-response bias for Trump voters.

3.2. **Testing for the Law of Large Populations.** In Figure 3, we plot $|Z_{n,N}|$, the relative error $\bar{G}_n - \bar{G}_N$ in units of the SRS error $\mathbb{V}_{SRS}(\bar{G}_n)$, to the number of total votes recorded in each state (including votes for each candidate and other alternative candidates), with log axes. We also fitted the regression of Eq 2.5 to the plot. We report the gradient coefficient of $\log N$ and its standard error in the boxes.

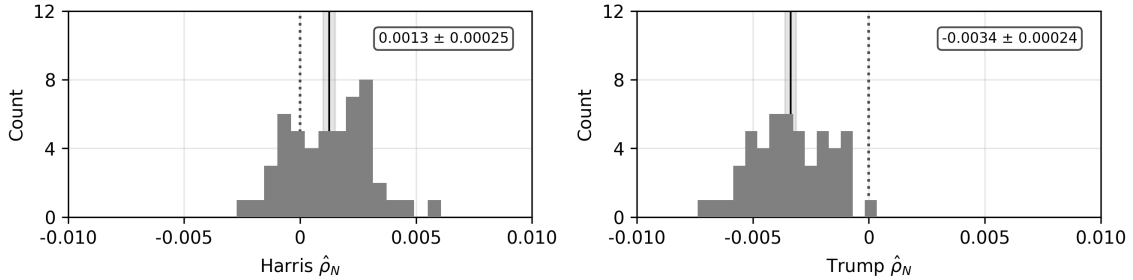
We analyze the case for all respondents and for likely voters. For Harris, we found that $|Z_{n,N}|$ stays roughly constant regardless of the number of voters in each state, which echoes the findings for 2016 in Meng, 2018. The standard error for the gradient for Harris puts zero within range, which is expected for a SRS where $Z_{n,N}$ stays relatively constant over N. For Trump, we see a visible increase in $|Z_{n,N}|$ as the actualized voter population size grows. We found the gradient coefficient considered alongside the standard error to be well within 0.5 and away from zero. Recalling the construction of Eq 2.5, the clear linear trend in the right plot of Fig 3 with a gradient of 0.5 is indicative of a non-zero data defect correlation that does not depend on N. This shows that the estimators $\bar{G}_n$ for Trump's vote share have errors that scale with the state population size $\sqrt{N}$. A gradient close to 0.5 (0.44 with a standard error of 0.09) was also reported in Figure 6 of Meng, 2018 for Trump.

3.3. **Lower Confidence in Bigger Data with the Big Data Paradox.** In Section 2.3, we assumed constant sample ratios to describe a dependence of the standardized error Z_n on $\sqrt{n}$. Although the sample ratios in CES are not constant across states, they do not vary as widely as much as the sample sizes for each state, as Figure 6 shows. We thus approximate that the sample ratios are constant across states to check for the $n^{1/2}$ dependence described

(A) Using all respondents



(B) Using likely voters (pre-election)



(C) Using validated voters (post-election)

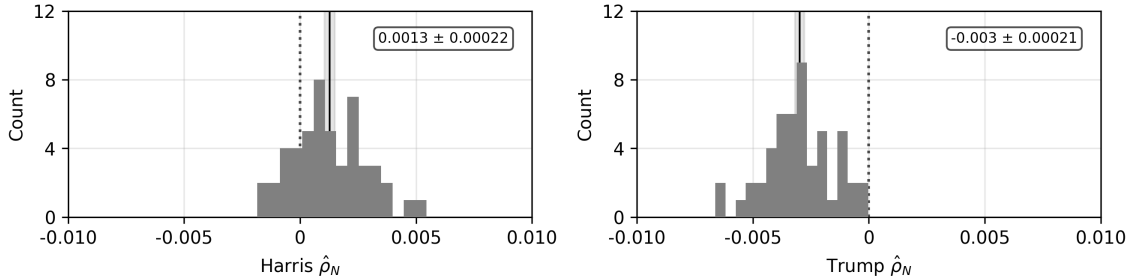
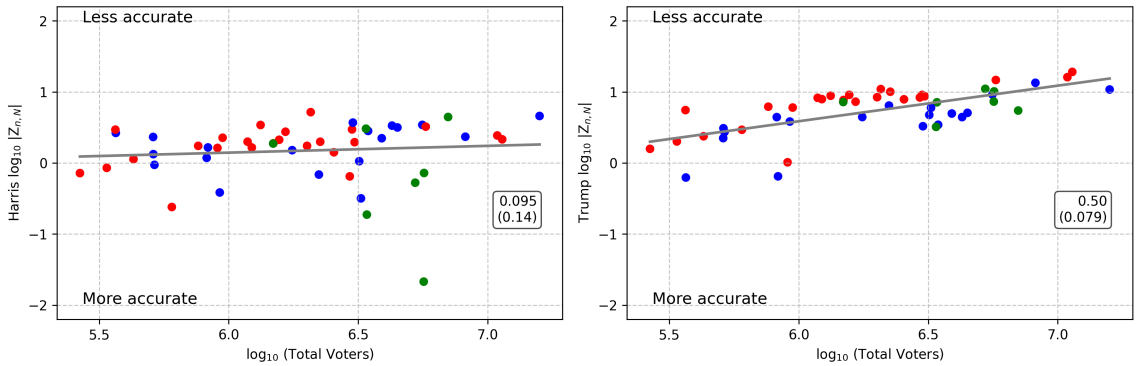


Figure 2. Histogram of data defect correlation for Harris and Trump voters for each state in the 2024 US presidential election. The boxes in the top right of the plots show "mean ±2 standard error". Zero data defect correlation is marked in dotted vertical lines. The means are marked in solid vertical lines. ±2 standard errors are shaded in light gray.

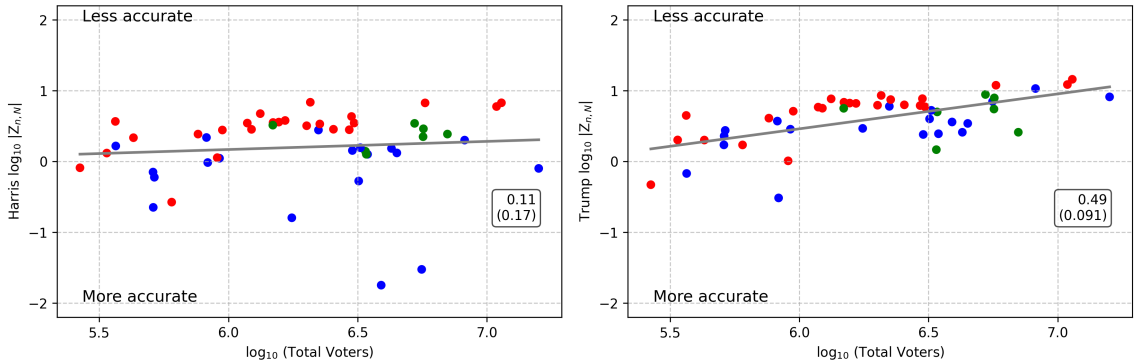
in Section 2.3. The full table of sample sizes and sample ratios for each state can be found at Table 3 in the appendix.

In Figure 4, we plot Z_n, the error standardized by its error, to the sample size n, with log on the x-axis. Note that we chose to plot against sample size n instead of population size N to better showcase the dependence on n. We shade the confidence intervals of ±2 standard deviation under typical Normal distribution assumptions in gray. We found that for 2024 the error for Trump tends to exit the confidence interval as the sample size grows,

(A) Using all respondents



(B) Using likely voters (pre-election)



(C) Using validated voters (post-election)

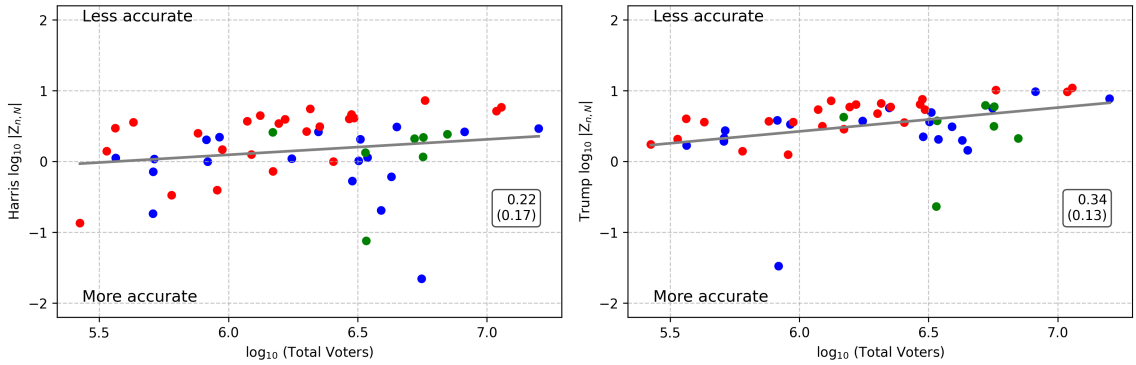


Figure 3. Error in SRS-units $Z_{n,N}$ to the number of total votes in each state.

reaching more than 10 standard deviations for the biggest states like California and Texas. The smaller states of Vermont and Wyoming are within the confidence intervals.

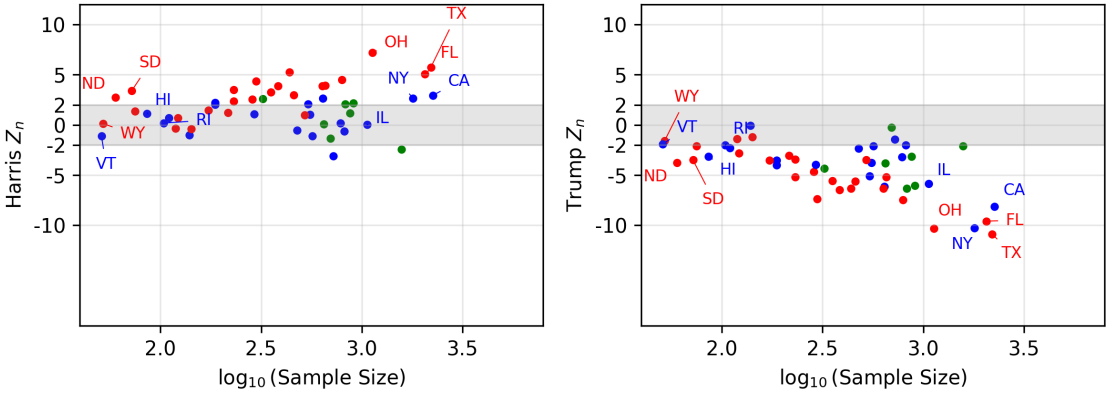


Figure 4. Standardized error Z_n to sample size for each state, using validated voters.

3.4. **Computing the effective sample sizes.** Using the methods in Section 2.4, we compute the effective sample size for each state, which is the sample size that yields the same error under SRS as the errors observed with the current selection mechanism in CES. We include all respondents for this section, not just validated voters. We report the results for the five biggest states by actualized voters in Table 1 and for the five smallest states in Table 2. We report the full results for all states in Table 4 in the appendix.

Table 1. Effective sample sizes and percentage reduction for predicting Trump vote shares in CES 2024 for the five biggest states

State	Sample size	Total votes	Effective sample size	Percentage reduction
Texas	4222	11,380,105	19	99.55%
California	4035	15,862,536	13	99.67%
Florida	3543	10,893,548	17	99.53%
New York	3205	8,199,062	20	99.37%
Pennsylvania	2503	7,034,206	18	99.27%

Table 2. Effective sample sizes and percentage reduction for predicting Trump vote shares in CES 2024 for the five smallest states

State	Sample size	Total votes	Effective sample size	Percentage reduction
Wyoming	88	266,353	17	80.60%
Alaska	117	338,177	18	84.72%
North Dakota	99	365,059	14	85.85%
Vermont	85	366,389	12	85.90%
South Dakota	131	428,922	16	87.95%

We found that the effective sample sizes are between 12 and 30 (inclusive), even though the actual sample sizes are between 85 and 4,222. The biggest percentage reductions occur

for the biggest states, and the five biggest states all have percentage reductions above 99%. Meanwhile, the five smallest states have percentage reductions below 90%, and are also the five states with the smallest percentage reductions. Examining Eq 2.9, it is expected that effective sample size stays constant regardless of actual sample size when the sample ratio stays constant (which is roughly the case here, see Figure 6 and Table 3). Since the effective sample size is roughly constant across states in the CES dataset, it follows that the percentage reduction is the biggest in the biggest states.

3.5. **Saving the Polls with Bias Correction.** We computed the bias-corrected estimates for 2024 using Eq 2.12 and data from 2016 and plotted it against the actualized vote share in Figure 5. There is no longer any visible and consistent underestimation of Trump's vote share across states. The RMSE is 0.05, which is much lower than that originally of Trump (0.13), and similar to that of Harris in Figure 1. Note that the good performance of the new estimator $\bar{G}_n^*$ is driven by the similarity between the average data defect correlation for 2016 and 2024, which was -0.0044 and -0.0045 respectively. This average data defect correlation might vary in utility to future elections, where the data defect correlation might shift in distribution.

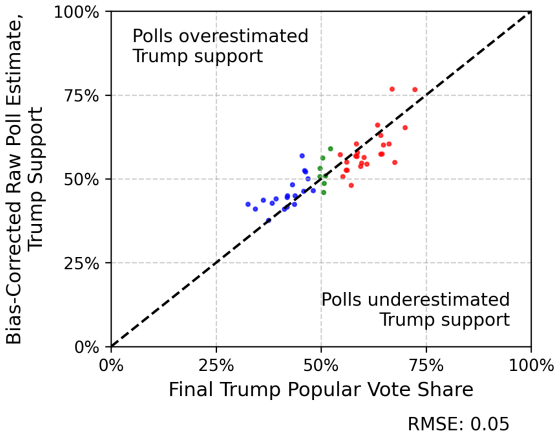


Figure 5. Bias-corrected raw poll estimate of Trump's vote share in 2024

4. Discussion

Before the election, using either the all respondents subset or the likely voters subset revealed a strong non-response bias from Trump voters, with the effect diminishing slightly when analyzing validated voters post-election. This raises questions on how we can improve polling techniques so that non-response bias could be alleviated in pre-election analysis. We offered one such suggestion in Section 2.5 but we believe this can be further improved upon.

We did not consider the errors coming from voters changing their mind between answering the survey and voting in the election, and assumes that survey respondents answer truthfully. Out of 50,396 survey respondents who were asked the question "Which candidate

for President of the United States do you prefer?", 2,999 respondents responded "I'm not sure". Even if we assign *all* these responses to count towards Trump's estimate, we still get a consistent underestimation of Trump's share across all states. We report this result in Figure 9 in the appendix. This shows that there is still non-response bias in the selection mechanism of CES even after accounting for dishonest responses.

One idea to counteract the data defect correlation is to increase the data quantity, as hinted at in the Meng identity (Eq 2.1), since data quantity is one of the factors of errors alongside data quality. However, as argued in Section 3 in Meng, 2018, this is not feasible. We present a new example using the 2024 election. Suppose we would like the effective sample size in Pennsylvania, a swing state which had 7 million votes counted in 2024 and where CES collected 2,503 survey responses, to reach 1,000. Using Eq 2.9, which gives a theoretical upper bound to the effective sample size, we require a sample ratio of at least 0.019, which is approximately 130,000 Pennsylvanian respondents. The entire CES survey only has $60,000$ respondents across all states. To move forward and improve polling estimates, we couldn't rely on using data quantity to fully overcome the issues from data quality, but we have to directly address non-response bias itself.

Further work includes computing the data defect correlation for other nation-wide state-level surveys, including that of 270toWin, 2024, and compare it to CES.

Appendix A. Details on sample sizes and sample ratios

Table 3 shows the sample ratio, sample size and total votes in each state for the 2024 US presidential election. Sample ratios and sample sizes are from CES Schaffner et al., 2025. The total votes are from (FEC), 2024.

Table 3. Sample ratio, sample size and total votes in each state for the 2024 US presidential election

State	Sample ratio	Sample size	Total votes
Alabama	0.00039	882	2,256,352
Alaska	0.00035	117	338,177
Arizona	0.00034	1,162	3,389,319
Arkansas	0.00042	492	1,182,676
California	0.00025	4,035	15,862,536
Colorado	0.00027	867	3,190,873
Connecticut	0.00030	536	1,758,429
Delaware	0.00039	202	511,697
Florida	0.00033	3,543	10,893,548
Georgia	0.00031	1,608	5,250,047
Hawaii	0.00030	153	516,701
Idaho	0.00027	246	904,812
Illinois	0.00035	1,936	5,592,368

Indiana	0.00037	1,078	2,933,770
Iowa	0.00033	544	1,656,849
Kansas	0.00034	449	1,327,591
Kentucky	0.00041	844	2,073,309
Louisiana	0.00031	615	2,006,975
Maine	0.00027	224	831,375
Maryland	0.00029	868	3,015,650
Massachusetts	0.00027	917	3,453,369
Michigan	0.00026	1,470	5,662,504
Minnesota	0.00025	799	3,240,916
Mississippi	0.00039	475	1,228,008
Missouri	0.00041	1,215	2,993,596
Montana	0.00029	173	602,963
Nebraska	0.00030	285	947,159
Nevada	0.00040	590	1,484,840
New Hampshire	0.00034	283	822,116
New Jersey	0.00032	1,362	4,272,725
New Mexico	0.00035	321	923,403
New York	0.00039	3,205	8,199,062
North Carolina	0.00027	1,537	5,679,647
North Dakota	0.00027	99	365,059
Ohio	0.00036	2,083	5,765,017
Oklahoma	0.00043	672	1,566,173
Oregon	0.00040	891	2,229,467
Pennsylvania	0.00036	2,503	7,034,206
Rhode Island	0.00035	178	510,659
South Carolina	0.00035	895	2,548,140
South Dakota	0.00031	131	428,922
Tennessee	0.00037	1,133	3,063,942
Texas	0.00037	4,222	11,380,105
Utah	0.00029	426	1,487,951
Vermont	0.00023	85	366,389
Virginia	0.00026	1,177	4,482,576
Washington	0.00030	1,166	3,898,835
West Virginia	0.00053	403	762,390
Wisconsin	0.00032	1,077	3,415,213
Wyoming	0.00033	88	266,353

Figure 6 shows that sample ratios in CES do not vary as widely as much as the sample sizes for each state.

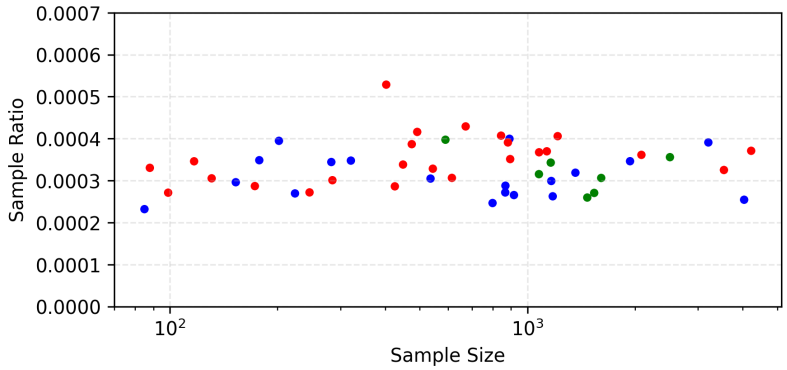


Figure 6. Sample ratio to sample size of Schaffner et al., 2025. The x-axis uses a log scale.

Appendix B. Details on effective sample sizes

Table 4 shows the effective sample sizes and percentage reduction for CES (Schaffner et al., 2025) in each state for the 2024 US presidential election using the methods outlined in Section 2.4.

Table 4. Sample size, total votes, effective sample size and percentage reduction of samples in each state for the 2024 US presidential election

State	Sample size	Total votes	Effective sample size	Percentage reduction
Alabama	882	2,256,352	20	97.71%
Alaska	117	338,177	18	84.72%
Arizona	1,162	3,389,319	18	98.48%
Arkansas	492	1,182,676	21	95.63%
California	4,035	15,862,536	13	99.67%
Colorado	867	3,190,873	14	98.38%
Connecticut	536	1,758,429	16	97.06%
Delaware	202	511,697	20	89.90%
Florida	3,543	10,893,548	17	99.53%
Georgia	1,608	5,250,047	16	99.02%
Hawaii	153	516,701	15	90.00%
Idaho	246	904,812	14	94.29%
Illinois	1,936	5,592,368	18	99.08%
Indiana	1,078	2,933,770	19	98.24%
Iowa	544	1,656,849	17	96.88%
Kansas	449	1,327,591	17	96.11%
Kentucky	844	2,073,309	21	97.51%
Louisiana	615	2,006,975	16	97.43%
Maine	224	831,375	14	93.79%
Maryland	868	3,015,650	15	98.29%

Massachusetts	917	3,453,369	14	98.50%
Michigan	1,470	5,662,504	13	99.09%
Minnesota	799	3,240,916	13	98.41%
Mississippi	475	1,228,008	20	95.79%
Missouri	1,215	2,993,596	21	98.27%
Montana	173	602,963	15	91.43%
Nebraska	285	947,159	16	94.54%
Nevada	590	1,484,840	21	96.52%
New Hampshire	283	822,116	18	93.71%
New Jersey	1,362	4,272,725	16	98.79%
New Mexico	321	923,403	18	94.40%
New York	3,205	8,199,062	20	99.37%
North Carolina	1,537	5,679,647	14	99.09%
North Dakota	99	365,059	14	85.85%
Ohio	2,083	5,765,017	19	99.10%
Oklahoma	672	1,566,173	22	96.70%
Oregon	891	2,229,467	21	97.68%
Pennsylvania	2,503	7,034,206	18	99.27%
Rhode Island	178	510,659	18	89.88%
South Carolina	895	2,548,140	18	97.97%
South Dakota	131	428,922	16	87.95%
Tennessee	1,133	3,063,942	19	98.31%
Texas	4,222	11,380,105	19	99.55%
Utah	426	1,487,951	15	96.53%
Vermont	85	366,389	12	85.90%
Virginia	1,177	4,482,576	14	98.85%
Washington	1,166	3,898,835	15	98.67%
West Virginia	403	762,390	27	93.22%
Wisconsin	1,077	3,415,213	16	98.49%
Wyoming	88	266,353	17	80.60%

Appendix C. Plotting Z_n with all respondents and likely voters

We replot Figure 4 using all respondents and likely voters. The same pattern holds where Z_n tends to escape the confidence interval for larger samples.

Appendix D. Accounting for dishonest responses

In Figure 9, we replotted Figure 1 for Trump but delegating "I'm not sure" responses as part of Trump's estimate. This decreased the RMSE and gave a better fit, but did not

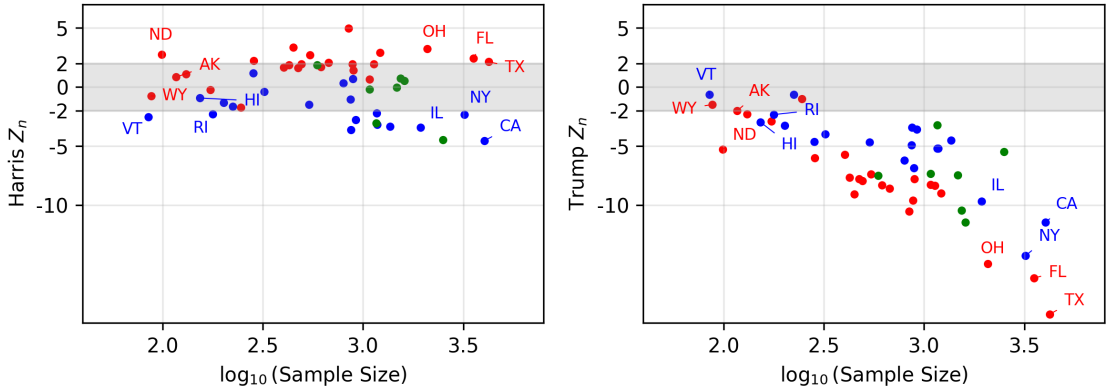


Figure 7. Standardized error Z_n to sample size for each state, using all respondents.

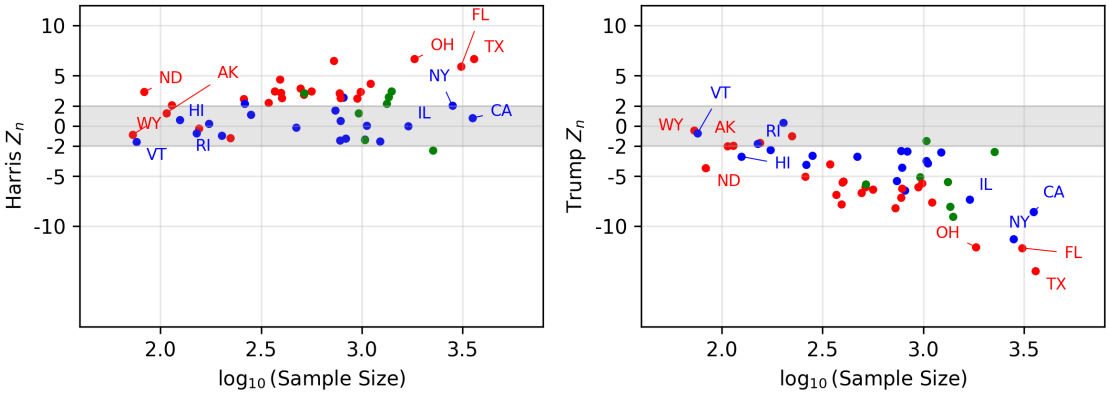


Figure 8. Standardized error Z_n to sample size for each state, using likely voters.

match the performance of the polls for Harris in Figure 1 or that of the bias-corrected estimator in Figure 5.

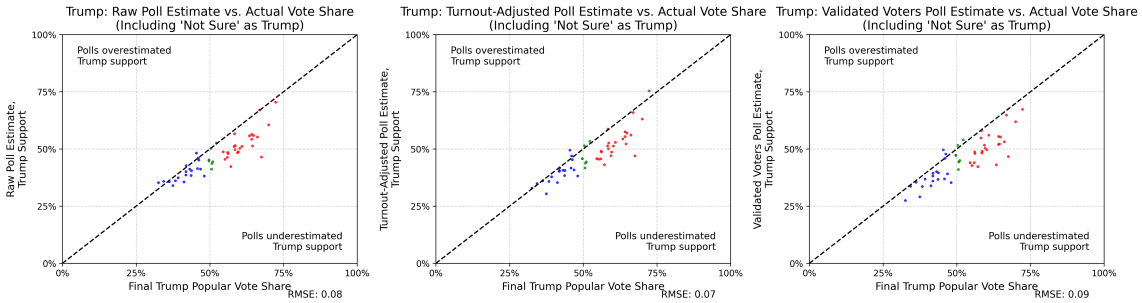


Figure 9. Poll estimate vs actual vote share for Trump for each state, including "I'm not sure" as part of the Trump estimate

Appendix E. Details on estimating turnout

To estimate turnout, we coarsen the data on the voting intention data in CES dataset and consider those who expressed any voting intention (i.e. one of "Yes", "Probably", "Voted") as likely voters, and discard the responses that responded "No" or "Undecided".

$$\bar{G}_{n,\text{turnout-adjusted}} = \frac{\text{Number of likely voters who supports the candidate}}{\text{Total number of likely voters}} \tag{E.1}$$

A more sophisticated approach will be to apply weighting to each of "Yes", "Probably", "Voted", but it is unclear how much weight should be assigned to each class of voters. Resampling from those who chose "Probably" or "Undecided" has been considered, but again there is no clear principle on what probabilities to assign to these classes.

References

270toWin. (2024). 2024 Presidential Election Polls. https://www.270towin.com/2024-presidential-election-polls/

Cook Political Report. (2024). 2024 CPR President Race Ratings. https://www.cookpolitical.com/ratings/presidential-race-ratings

Coppock, A. (2017). Did Shy Trump Supporters Bias the 2016 Polls? Evidence from a Nationally-representative List Experiment. *Statistics, Politics and Policy*, *8*(1), 29–40. https://doi.org/10.1515/spp-2016-0005

(FEC), F. E. C. (2024). Election results and voting information. Retrieved March 25, 2025, from https://www.fec.gov/introduction-campaign-finance/election-results-and-voting-information/

Federal Election Commission. (2025, January). *2024 Presidential Election Results* (Report). Federal Election Commission. https://www.fec.gov/resources/cms-content/documents/2024presgeresults.pdf

FiveThirtyEight. (2024). State of the Polls 2024. Retrieved March 25, 2025, from https://github.com/fivethirtyeight/data/tree/master/state-of-the-polls-2024

McDonald, M. (2023). 2016 General Election Turnout Rates (v1.0). https://election.lab.ufl.edu/dataset/2016-general-election-turnout-rates/

McDonald, M. (2024). 2024 General Election Turnout Rates (v0.3). https://election.lab.ufl.edu/dataset/2024-general-election-turnout-rates-v0-3/

Meng, X.-L. (2018). Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election. *The Annals of Applied Statistics*, *12*(2). https://doi.org/10.1214/18-AOAS1161SF

Schaffner, B., Shih, M., Ansolabehere, S., & Pope, J. (2025, April). Cooperative Election Study Common Content, 2024. https://doi.org/10.7910/DVN/X11EP6