Forecasting the 2024 Election: Tight Margins and Swing State Uncertainty*

Harris Leads Trump by 0.8% as Voters Remain Divided

Mariko Lee   Karen Riani   Cristina Su Lam

November 4, 2024

We forecast the 2024 U.S. presidential winner using state-level poll averages for Kamala Harris and Donald Trump, weighted by poll quality and sample size. Applying a Bayesian regression model, we find a slight 0.8% overall lead for Harris over Trump and Harris leading in all 7 swing states. We weigh the polls with aims to increase the influence of the most reliable data in shaping the forecast— limiting biased results and showing a clearer voter sentiment amidst a particularly competitive election. These results show who may be leading the elections via polls, giving campaigns the estimations they need and reinforcing public trust in the election process, especially as the outcome likely hinges on key swing states.

Table of contents

---

1 Introduction

Amid the chaotic, high-stakes 2024 U.S. presidential election, forecasting models offer useful evaluations into voter sentiment as Americans face a tight race between top candidates Donald Trump and Kamala Harris. With the attempted assassination of Trump, Joe Biden's endorsement of Harris, and recent crises such as hurricanes disrupting voter access in swing states, public opinion remains volatile (ABC News 2024b). Key events like Trump's policy proposals on social security and public safety, Harris's advocacy for reproductive rights, and other rising debates on gun control have further polarized the electorate, making forecasting models essential in offering clarity amid this competitive landscape (CBS News 2024; Human Rights Watch 2022).

Using presidential polling data from FiveThirtyEight, our Bayesian model's estimand is the probability of support for Kamala Harris, weighted by pollster rating and sample size to improve the reliability of aggregated polling data. By refining the aggregation of polls through Bayesian regression, filtered from the reelection day past July 21st, 2024 and weighted by sample size and pollster rating, this study contributes to the growing body of forecast models—including media outlets' own polls-of-polls and various Bayesian methods. Our model estimates only a slim 0.8% lead for Harris over Trump in the overall popular vote—a close margin that highlights the high uncertainty in public sentiment. In line with other forecasts that struggle to declare a clear frontrunner, this narrow lead reflects the volatility of the race and the importance of updated, representative polling data (The New York Times 2024c).

This uncertainty emphasizes the significance of swing states in election forecasts—particularly states with nearly equal support for each party and a history of alternating between Democratic and Republican candidates in recent elections (Bloomberg 2024). Existing research identifies seven critical swing states, with Pennsylvania standing out as a tipping point in more than 1 in 5 simulations (Associated Press 2024). With 19 electoral votes, Pennsylvania could prove decisive in the event of a close race. Our estimates reveal Harris' lead in each swing state, however they are slim leads. The competitive nature of these swing states shows the need for continued data updates to maintain forecast accuracy in a closely contested race.

By prioritizing higher-quality polls and focusing on data collected after Harris's campaign announcement, the model seeks to mitigate biases and improve forecast reliability by using a polls-of-polls method over individual pollster data. This approach not only clarifies voter sentiment amid competing forecasts but also demonstrates the value of weighting variables such as pollster rating and sample size in enhancing prediction accuracy. While some states lack sufficient polling data—introducing potential limitations that could be improved with more data and model complexity—the overall findings provide strategic forecasts for decision-makers and reinforce public confidence in an election where no candidate holds a decisive lead. Appendix A provides a detailed assessment on The Washington Post, one of the polling methodologies used by prominent sources, including sampling approaches, weighting, and non-response handling strategies to ensure robust data quality.

The rest of this paper is structured as follows: Section 2 presents the data sources and methodology, Section 3 presents the forecasting model, then Section 4 followed by Section 5, which discusses the results and their implications. Section 6 concludes with a discussion of the limitations of the study and suggestions for future research, followed by Section 7 with The Washington Post's pollster methodology, and Section 8 a budgeted idealized methodology.

2 Data Analysis

2.1 Overview

To forecast the 2024 U.S. presidential election, this analysis uses state-level polling data for candidates Kamala Harris and Donald Trump. Our primary data source is FiveThirtyEight's 2024 Presidential Election Forecast Database (FiveThirtyEight 2024a), which aggregates polling data from multiple organizations and provides quality ratings for each poll, helping to identify and weigh the most reliable sources. The dataset covers polls conducted from July 21, 2024—when Kamala Harris announced her reelection bid—through to November 3, 2024.

Data cleaning was conducted in R (R Core Team 2024) using several packages: tidyverse (Wickham et al. 2019), janitor (Firke 2023), and lubridate (Grolemund and Wickham 2023). Duplicate rows were removed, and the dataset was filtered to include only polls for Harris and Trump. Observations with missing values in key variables such as 'sample size,' 'numeric grade,' and 'state' were excluded. A binary variable, 'Chosen Candidate' (1 for Harris, 0 for Trump), was created to indicate support for each candidate. After these adjustments, the dataset was refined from its original 17,765 entries to 2,367 observations. The primary variables of interest include pollster, numeric grade, state, sample size, percentage, and chosen candidate. A sample of the cleaned dataset is displayed in Table 1.

For visualization and analysis, the knitr (Xie 2023), ggplot2 (Wickham 2016), kableExtra (Zhu 2024), map (Becker, Ray Brownrigg. Enhancements by Thomas P Minka, and Deckmyn. 2023) packages were utilized to generate tables, graphs, and maps, presenting the data clearly and effectively.

Table 1: Sample of President Polls Dataset

Pollster	Numeric Grade	State	Sample Size	Percentage	Chosen Candidate
Patriot Polling	1.1	Pennsylvania	903	48.7	1
Patriot Polling	1.1	Pennsylvania	903	49.6	0
Patriot Polling	1.1	Michigan	858	49.2	1
Patriot Polling	1.1	Michigan	858	48.5	0
Patriot Polling	1.1	Wisconsin	835	49.0	1

Table 2 presents summary statistics of the three numerical variables used. The **Percentage Support** variable, representing the proportion of respondents supporting a candidate, has a mean of 47.25% with a standard deviation of 4.41%, ranging from a minimum of 25% to a maximum of 70%. This spread reflects varying levels of support for each candidate across different polls.

The **Sample Size** variable, indicating the number of respondents in each poll, has a mean of 930.92 with a substantial standard deviation of 550.61, indicating variability in poll sizes. The minimum sample size is 301, while the maximum reaches 6,526, highlighting the range in poll reliability, as larger samples tend to yield more accurate reflections of public sentiment.

Lastly, the **Numeric Grade** variable, which rates pollster quality, shows an average score of 2.28 with a standard deviation of 0.63, spanning from 0.5 to 3. This distribution suggests that the dataset includes a mix of polls with varying reliability, with higher-rated polls contributing more heavily to the overall forecast due to weighted adjustments (further discussed in Section 3).

Table 2: Summary Statistics

Statistic	Mean	SD	Min	Max
Percentage Support	47.25	4.41	25	70
Sample Size	930.92	550.61	301	6,526
Numeric Grade	2.28	0.63	0.5	3

2.2 Measurement and Limitations

Converting real-world electoral dynamics into data requires careful measurement of public sentiment towards presidential candidates for the 2024 U.S. presidential election. This dataset is compiled from polling data collected by established pollsters such as Siena/NYT and YouGov, employing diverse methodologies to gauge voter preferences. Key variables captured include the percentage of respondents favoring each candidate, sample sizes, and demographic information about the participants.

The measurement process begins with pollsters designing surveys aimed at accurately reflecting voter attitudes. To ensure a representative sample, stratified random sampling methods are employed, accounting for the electorate's diverse demographics. Poll responses are gathered through various channels, including phone interviews, online surveys, and face-to-face interactions, ultimately quantifying voter support as a percentage for each candidate.

After polling concludes, the dataset undergoes thorough cleaning and processing to rectify missing values and discrepancies. Each entry corresponds to a specific poll, providing details such as the date of the poll, the methodology used, and the percentage of support for each candidate. A critical component is the Numeric Grade, which ranges from 0.5 ("Poor") to

3.0 ("Exceptional"), indicating poll quality and reliability in capturing true public sentiment. Based on FiveThirtyEight's methodology (ABC News 2024a), this grading system prioritizes data from more rigorous and representative polls, thereby enhancing forecast accuracy.

To determine these Numeric Grades, FiveThirtyEight assesses factors like sampling methods, sample sizes, methodological transparency, and historical pollster accuracy. Polls with robust data collection practices and larger representative samples receive higher ratings, reflecting a greater likelihood of reliability. Conversely, lower ratings signify methodological concerns that could affect accuracy.

The dataset transforms complex voter opinions into structured data, assigning reliability metrics through the Numeric Grade to facilitate systematic analysis of public sentiment trends. This methodology aids in making more precise predictions about election outcomes. However, challenges such as sampling bias, survey design limitations, and non-response bias must be acknowledged, as they can impact data accuracy and emphasize the importance of careful interpretation of polling results.

2.3 Outcome Variable

The chosen candidate—whether a poll favors Harris or Trump—serves as the outcome variable in this analysis. Figure 1 shows the distribution of polls between the two candidates, Harris and Trump, revealing a nearly equal split in support, with 49.8% favoring Trump and 50.2% favoring Harris. This close distribution highlights a competitive polling landscape, suggesting that public opinion is almost evenly divided between the candidates.

In Figure 2, the state-level distribution of support further emphasizes this balance, with only slight variations in candidate preference across states. For instance, states like Maryland, Montana, Nevada, Ohio, and Utah show a marginally higher level of support for Harris, representing the Democratic party. Conversely, in states such as Florida, Minnesota, and Pennsylvania, there is slightly greater support for Trump, representing the Republican party. These minor differences indicate regional trends but reinforce the overall theme of a closely contested race.
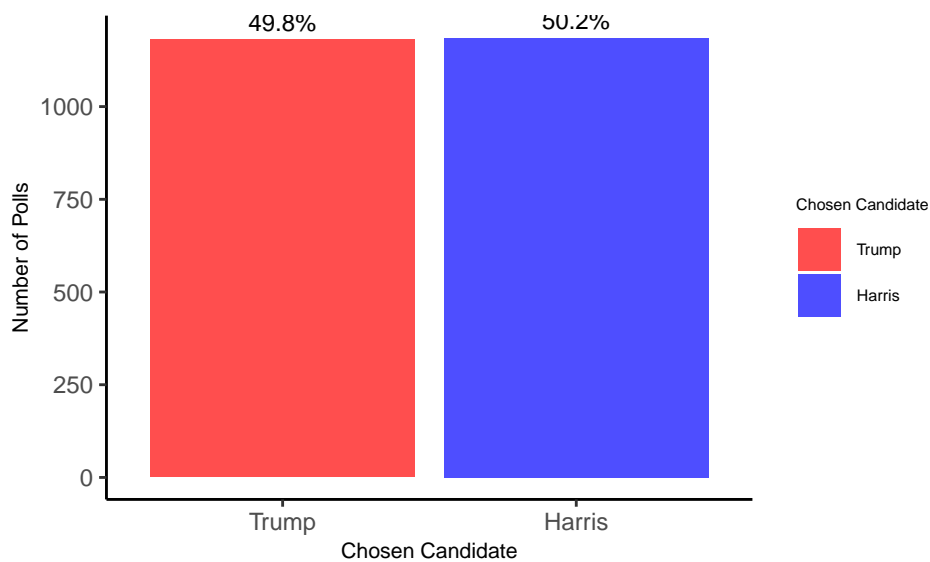
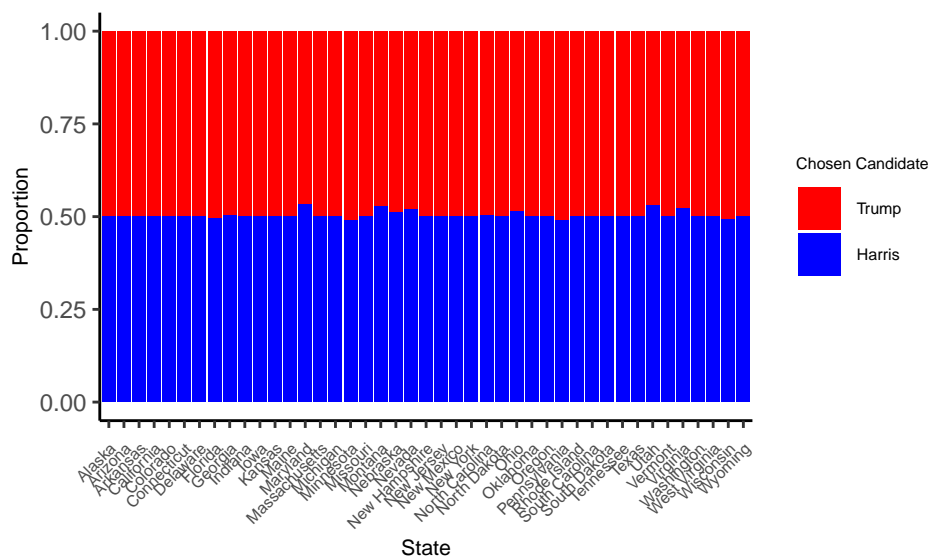Figure 1: Distribution of Polls by Harris vs Trump

Figure 2: Distribution of Candidate Preference by State

2.4 Predictor Variables

2.4.1 Distribution of Sample Size

Figure 3 shows the distribution of sample sizes across polls, with a right-skewed pattern indicating that most polls have smaller sample sizes and a few larger outliers. The majority of polls range between 300 and 1,500 respondents, peaking around 1,000 respondents, while only a few reach the maximum sample size of 6,526, as noted in Table 2. This highlights that larger sample sizes are relatively uncommon in the dataset.

The prevalence of smaller sample sizes suggests greater variability in the precision of poll estimates, as smaller samples generally introduce more uncertainty. In contrast, larger sample sizes provide more stable and reliable reflections of public sentiment. This distribution emphasizes the need to account for sample size in the analysis to ensure that more reliable, larger polls are given appropriate weight, thus enhancing the accuracy of public opinion estimates.
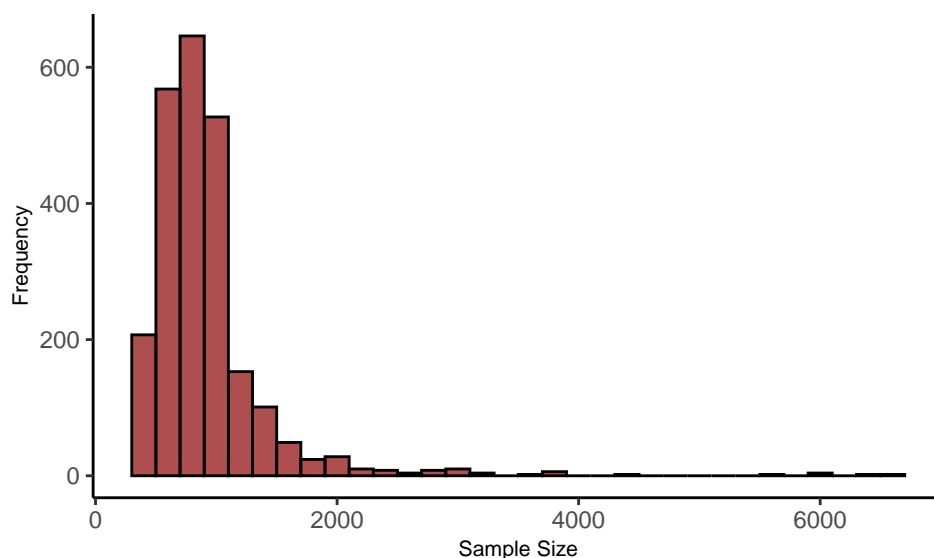
Figure 3: Distribution of Sample Size

2.4.2 Distribution of Percentage of Candidate Support by Chosen Candidate

Figure 4. The density plot shows the distribution of candidate support percentages for Donald Trump and Kamala Harris, with both peaking around the 50% mark, where Harris has a slightly higher concentration near this midpoint. The magenta overlap region highlights this balance, showing no significant advantage for either candidate.

A few outliers in both distributions show support above 60% or below 40%, representing occasional strong deviations, though these instances are rare. This pattern reflects a balanced public sentiment with minor variation between the candidates and points to a highly competitive race. The symmetry and overlap of the distributions reinforce the close competition between Trump and Harris, indicating an almost evenly divided public opinion across the dataset.
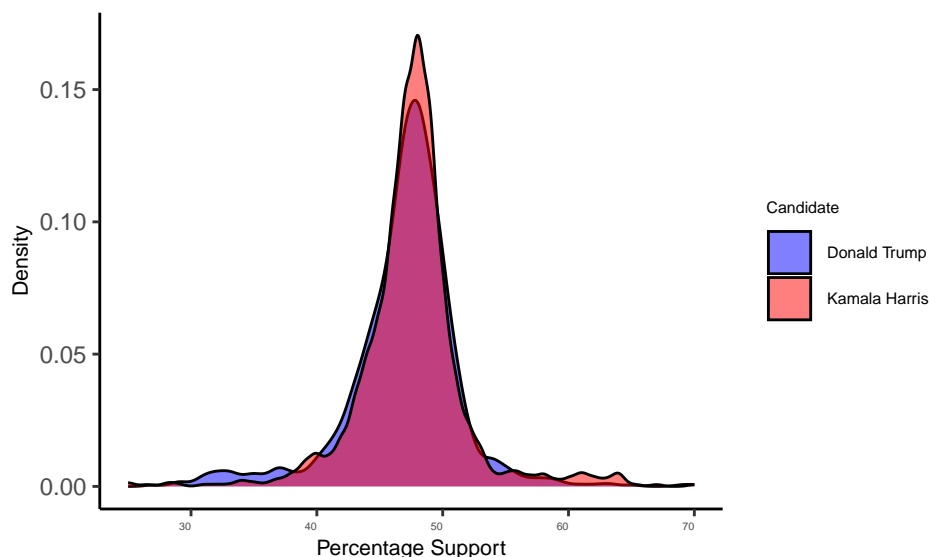

Figure 4: Distribution of Candidate Support by Chosen Candidate

2.4.3 Distribution of Pollster and Numeric Grades

Figure 5 displays the distribution of Numeric Grades assigned to pollsters, as per FiveThirtyEight's grading system, with each grade indicating poll reliability. The x-axis shows the Numeric Grades, ranging from 0.5 to 3.0, while the y-axis represents the frequency of each grade.

The chart reveals that most polls come from higher-rated pollsters, with the highest concentrations at grades 2.0 (298 polls) and 3.0 (386 polls). This distribution suggests that the dataset primarily comprises data from pollsters with moderate to high reliability, while lower-quality sources, such as those with grades 0.5 and 1.0, are less common.

Complementing this, Figure 6 shows the distribution of polls by individual pollsters, including only those with more than 20 counts. In this graph, bar colors indicate each pollster's Numeric Grade, with darker shades of blue representing higher reliability scores. Notably, pollsters such as Siena/NYT and Marquette Law School have the highest Numeric Grade of 3.0, marking

them as highly reliable sources. In contrast, pollsters like Trafalgar Group and Patriot Polling, with scores of 0.7 and 1.1, respectively, are represented as lower-quality sources.

These insights are crucial, as they enable the model to assign weights based on pollster reliability, improving the accuracy of the analysis. By prioritizing data from higher-grade pollsters, the model can more effectively capture public sentiment, minimizing the impact of potential biases or inaccuracies from less reliable sources.
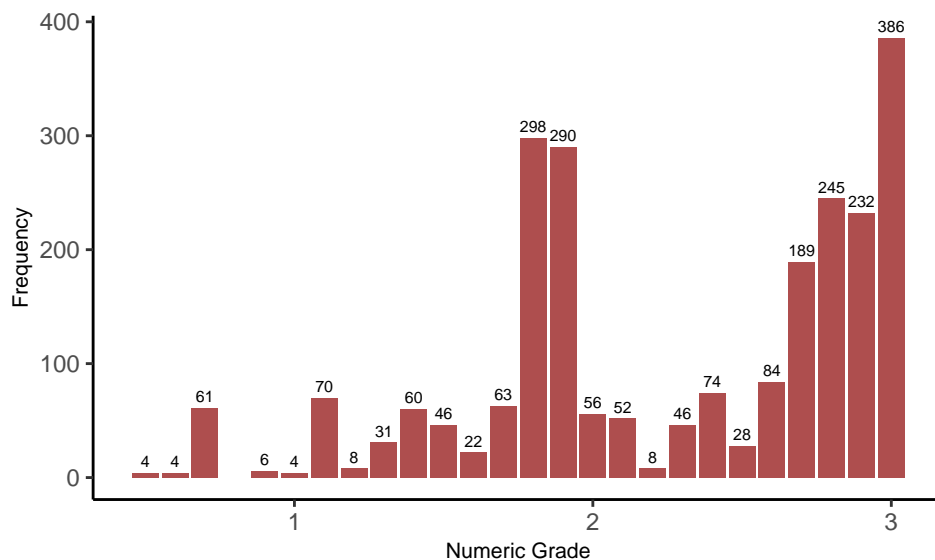

Figure 5: Distribution of Numeric Grade
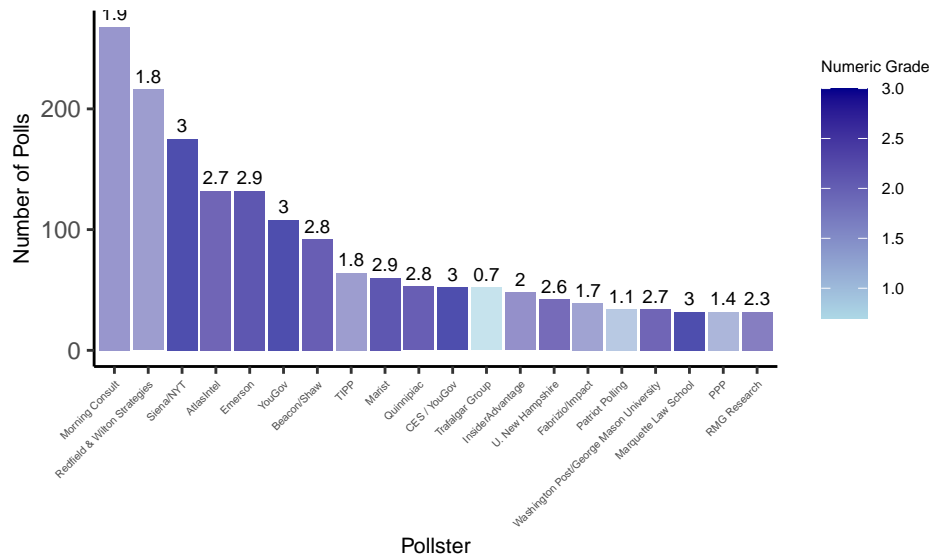
Figure 6: Distribution of Pollsters by Numeric Grade

2.4.4 Distribution of Polls and Candidate Support by State

Figure 7 displays the number of polls conducted in each U.S. state. Pennsylvania has the highest poll count, followed by other key swing states such as Arizona, Georgia, Michigan, and Wisconsin. The concentration of polls in these states likely reflects their strategic importance as battlegrounds where shifts in public opinion can be crucial. In contrast, states like Vermont, Wyoming, and West Virginia have far fewer polls, suggesting less emphasis on states with more predictable or less competitive outcomes.

Figure 8 presents a map indicating the leading candidate by state, with blue for Kamala Harris, red for Donald Trump, and gray for states with insufficient polling data. The map reveals a geographic divide, with Trump leading in many central and southern states, while Harris is favored in parts of the Northeast and West. This distribution aligns with typical U.S. regional political trends, where certain areas have a historical leaning toward one party. The map provides a visual summary of these regional preferences, highlighting how candidate support varies across the country.

Figure 7: Number of Pollsters by State

Figure 8: Candidate with Highest Support Percentage by State
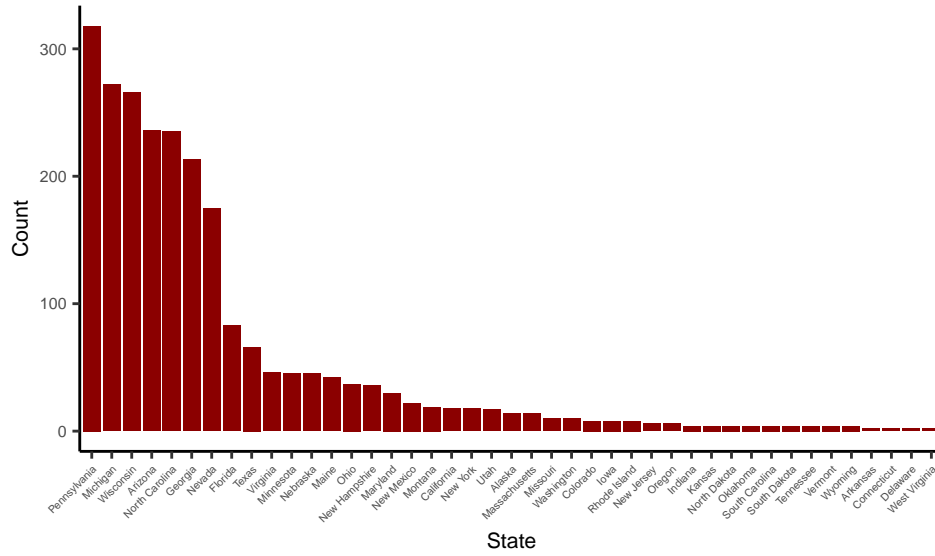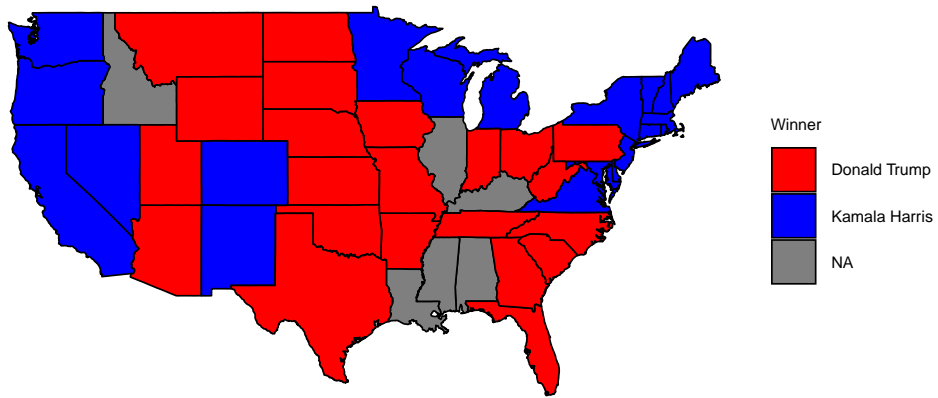
3 Model

3.1 Overview

This section presents a Bayesian model designed to predict voter support for Donald Trump and Kamala Harris during the 2024 U.S. presidential election. The model assesses the level of support for each candidate across 43 states, including key swing states. It outputs the likelihood that a voter will favor either candidate, utilizing variables such as pollster, state, and percentage of candidate support. Additionally, weights were incorporated into the model to account for variations in pollster reliability and sample size.

To conduct this analysis, the R packages rstanarm (Goodrich et al. 2024) and MLmetrics (Ben-Haim et al. 2023) were employed. The rstanarm package streamlines the Bayesian modeling process by providing tools for fitting generalized linear models using a Bayesian framework, while MLmetrics supplies various metrics for assessing model performance, including accuracy, F1 score, and area under the curve (AUC).

3.2 Model Structure

- Mathematical Representation

$$\mathrm{logit}(y_i) = \beta_0 + \beta_1 \times \mathrm{Pct}_i + u_{\mathrm{pollster}} + u_{\mathrm{state}} \tag{1}$$

$$\tag{2}$$

- Component Explanation
- y_i: Binary outcome variable indicating whether a respondent supports Harris (1) or not (0).
- β_0:Intercept term representing the baseline log-odds of support for Harris when the percentage support is zero.
- β_1: Coefficient for the percentage support for Harris (Pct), reflecting how changes in support percentage affect the log-odds of selecting Harris.
- u_{pollster}:Random effect for the specific pollster, accounting for variability in support across different pollsters.
- u_{state}: Random effect for the state in which the poll is conducted, capturing regional variations in support.

3.3 Variable Selection

In constructing the Bayesian model, careful consideration was given to the selection of predictor variables to ensure that the analysis accurately reflects the complexities of voter support.

- Percentage Support (Pct): This variable is integral to the model as it captures the nuanced impact of varying levels of support for the candidates. By utilizing percentage support as a continuous measure, the model facilitates a more detailed examination of how incremental changes in support influence the likelihood of selecting Harris over Trump. This approach enhances the understanding of voter behavior compared to categorizing support into discrete groups, which could obscure critical trends and insights.

- Random Effects: Treating pollster and state effects as random variables accounts for unobserved heterogeneity in the data. Different pollsters may use varied methodologies, resulting in fluctuations in reported support levels that fixed effects alone cannot capture. Additionally, geographical differences can influence voter preferences, highlighting the need to model state-level variability. By incorporating random effects for both pollster and state, the model effectively addresses this variability, enhancing its robustness and validity. This approach recognizes that the factors influencing voter support can differ significantly depending on the polling context.

3.4 Weights and Priors

The weights in this model are calculated using the formula:

$$\text{weights} = \text{Numeric Grade} \times \left(\frac{\text{Sample Size}}{\text{Mean(Sample Size)}} \right) \tag{3}$$

This approach ensures that more reliable polls—indicated by a higher Numeric Grade—carry greater influence in the analysis. By incorporating the sample size relative to the average sample size, this weighting system acknowledges that larger, more representative samples tend to yield more accurate reflections of public sentiment. This method effectively balances the impact of various polls based on their reliability and representation, leading to a more robust estimation of voter support.

Regarding the priors, we define them as follows:

$$\text{priors} \sim \text{Normal}(0.5, 0.1) \tag{4}$$

This prior is centered at 0.5 with a standard deviation of 0.1, reflecting the belief that, in the absence of strong data, the probabilities of support for both candidates are expected to be balanced. This choice of prior is particularly advantageous in situations where data is limited or when there is uncertainty regarding the expected distribution of voter support. By employing a normal prior, we establish a baseline assumption that allows the model to remain flexible and effectively integrate new data as it becomes available.

3.5 Model Validation and Justification

3.5.1 Model Validation Methods

The model validation process involved out-of-sample testing, where the dataset was split into training and testing sets in a 70-30 ratio. Both the Bayesian and Logistic models utilized the same features, with the key difference being that the logistic model did not include random effects for pollster and state. As seen in Table 3. The performance of each model was evaluated using three key metrics: F1 Score, Area Under the Curve (AUC), and Root Mean Square Error (RMSE).

Table 3: Comparison of Model Performance Metrics

Model	F1 Score	AUC	RMSE
Logistic Model	0.4217	0.4261	0.5551
Bayesian Model	0.5343	0.5301	0.5096

Although the Bayesian model still shows room for improvement—evidenced by F1 Score and AUC values not close to 1 and a RMSE indicating some prediction error—it outperforms the logistic model across all metrics. This suggests that the approach, which incorporates random effects to account for variability from different pollsters and states, offers a more nuanced understanding of voter support.

3.5.2 Diagnostics

To validate the Bayesian framework, several diagnostic checks were conducted to assess robustness and ensure that underlying assumptions were satisfied. Key among these were convergence assessments and posterior predictive checks. These checks provided a visual representation of how well the predictions align with the observed data. The graph reveals a close overlap between the observed outcomes (y) and the replicated predictions (y_{rep}), demonstrating effective capture of the distribution of voter support for Harris. This strong alignment suggests that the approach not only fits the training data well but also generalizes effectively to new observations. The flexibility of this framework allows for the incorporation of prior information and consideration of variability due to pollster and state effects, enhancing overall robustness (Refer to the Section 9, Table 5 for a visual illustration).

3.5.3 Model Justification

The choice of the Bayesian model is justified by its capacity to account for the random effects of pollsters and states, providing a richer understanding of public sentiment. This model allows

for nuanced insights that a simpler logistic regression might miss, particularly in a context where polling data is subject to variation due to different survey methods and demographic factors. Ultimately, while both models exhibit imperfections, the Bayesian model's superior performance metrics support its selection as the final model for predicting voter support in the 2024 U.S. presidential election.

4 Results

The analysis of the national predicted popular vote percentages indicates that Kamala Harris is projected to receive 50.44% of the vote, while Donald Trump is expected to garner 49.56%. These figures, illustrated in Figure 9 and calculated by averaging the mean predictions across all states, reveal a narrow lead of just 0.88% for Harris, highlighting a highly competitive race. This reveals a lack of dominance in candidacy and general uncertainty of pollster outcomes, suggesting that even this slight lead may not be substantial.
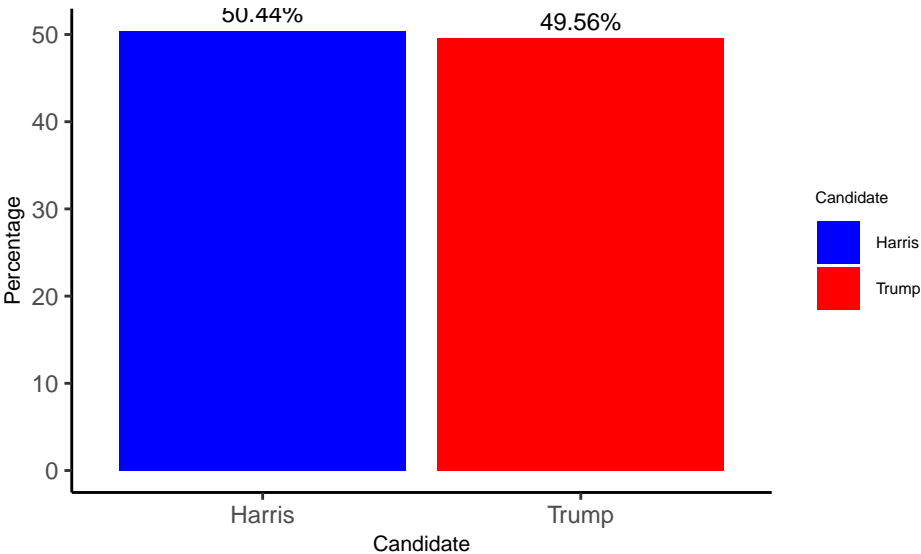
Figure 9: National Popular Vote Prediction

Additionally, Figure 10 presents a heat map of Harris's predicted probability by state, illustrating her lead over the majority of states but revealing that this advantage is generally weak, with most not peaking over a 1% advantage. Most states show only a slight preference for Harris over Trump, with the margin of error spanning both candidates in many cases. Specifically, out of the 43 states with available data, 25 states lean toward Harris, yet none of these states shows a substantial lead over more than a few percent over Trump. This pattern additionally shows the substantial closeness of the race, as evidenced by the margin of error, where

overlapping intervals highlight the slim advantage and limited certainty in Harris's predicted support.
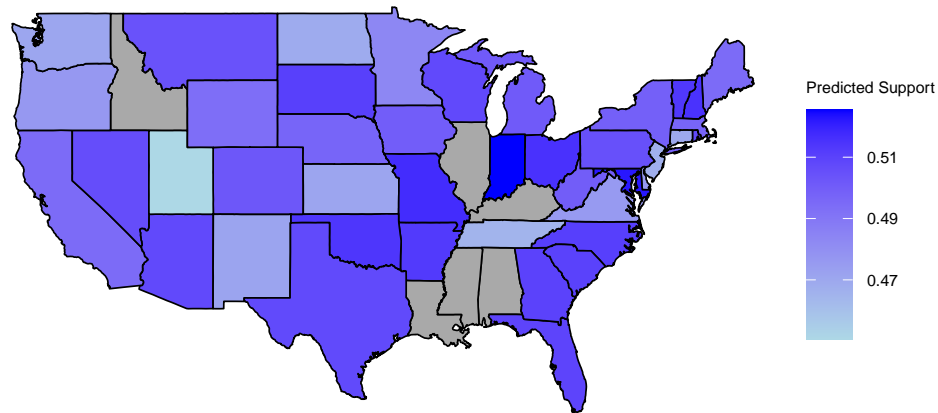
Figure 10: Heatmap of Predicted Support for Harris by State

Figure 11 zooms in on the seven swing states, each of which shows a slight lean toward Harris. The graph includes a red dashed line at the 50% threshold, and although Harris is estimated to lead in these swing states, the margin is extremely tight, with most probabilities hovering just above 50%. Only Michigan and North Carolina show a slight lead of just over 1%, reinforcing the sensitivity of these predictions. This pattern suggests that even small shifts in voter sentiment could tip these critical states in either direction.

Figure 11: Predicted Probability of Harris Winning by State with Swing State Highlight and 50% Threshold

Looking closer at the margin of error, in Figure 12 we can see Harris's predicted probabilities by state. States to the right of the dashed line (above 0.5) lean towards Harris, while those to the left lean towards Trump. In swing states like North Carolina, Michigan, and Pennsylvania, Harris's minimal lead is coupled with error bars that overlap this threshold. This suggests a high level of uncertainty—small shifts could easily change the lead in these states. Harris has a slight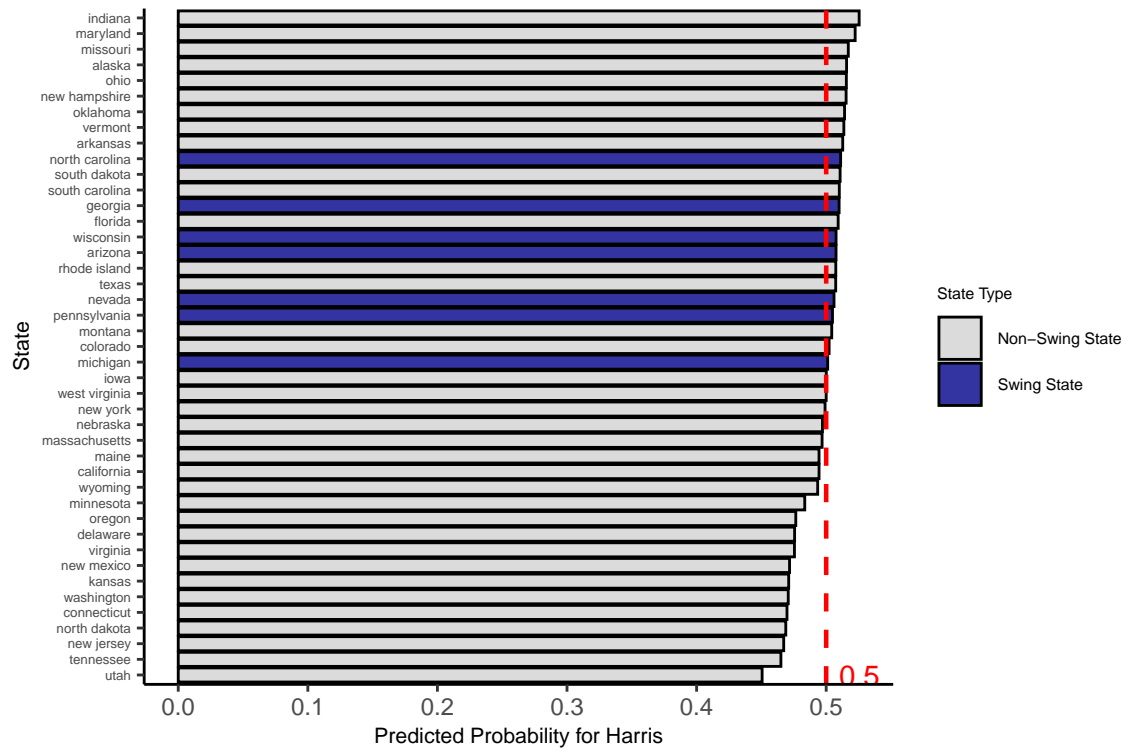 edge, but in many states, this lead is within the margin of error, thus showing that the election remains too close to call.
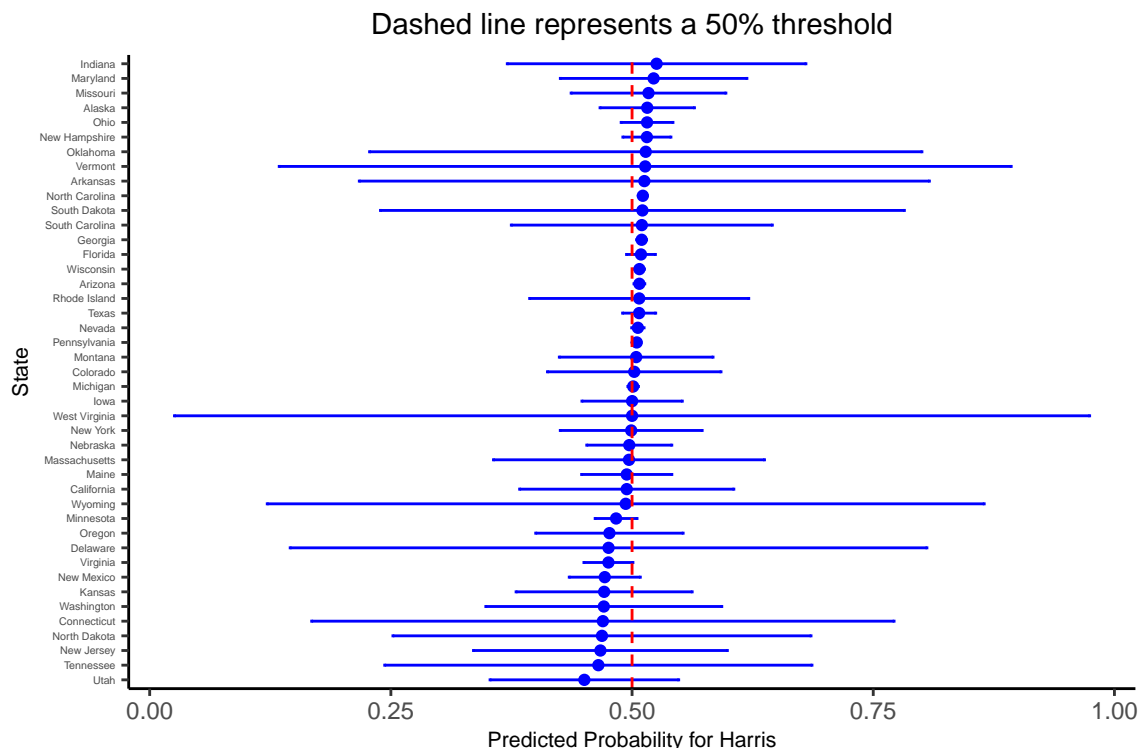
Figure 12: State-Level Predicted Probability for Harris with Margin of Error

Table 4 provides a summary of the model's coefficients and fit statistics, capturing key model parameters like the intercepts, predictor estimates, and evaluation metrics. The negative intercept of -2.07 represents the baseline log odds for Harris, establishing the prediction's starting point in the absence of other predictors. With all intercepts being equal, there is initially a lower probability for Harris's support without accounting for additional factors of percentage (pct) of support or pollster influence. The non-significant coefficient for sample size indicates that it does not directly affect prediction but serves as part of the poll weighting. In contrast, the pct coefficient of 0.04 suggests a slight positive effect on the odds for Harris, meaning an increase in her polling percentage slightly raises her predicted support odds. Low variance estimates for pollster and state effects imply limited inter-group variability, indicating that neither adds substantial random variation to the results.

Regarding model evaluation, the WAIC of 7072.4 and log likelihood of -3518.4 do not offer much information alone however a higher likelihood and a lower WAIC is preferred. The RMSE of 0.5 points to moderate accuracy, which aligns with the close and uncertain nature of the election. Additionally, a low ICC value suggests limited random variance contributed by pollster and state, furthermore showing the model's constrained precision in forecasting a definitive winner.

Table 4: Summary of Bayesian Model

Parameter	Estimate
Intercept	-3.779
Percentage Support	0.081
Pollster Variability	0.001
State Variability	0.001
Number of Observations	2367.000
R^2	0.027
Marginal R^2	0.026
Intraclass Correlation (ICC)	0.000
Log-Likelihood	-3849.195
WAIC	7728.200
RMSE	0.500

5 Discussion

In a highly competitive election, accurately forecasting outcomes requires more than just raw polling data—it demands careful aggregation and weighting to capture true voter sentiment. With this in mind, our model estimates a narrow 0.8% lead for Kamala Harris in the national popular vote, resulting in an aggregated support of approximately 50.4% across states.

While this suggests a slight edge for Harris, the lead easily falls within a $\pm 3\%$ margin that reflects substantial uncertainty (Andrew Mercer 2018). When a candidate's lead is within this margin, the difference may not be statistically significant, meaning the true preference could feasibly favor either candidate. In this context, a 0.8% lead falls well within this range and could be due to sampling noise rather than a clear voter preference. In close races, even minor shifts in voter sentiment or turnout in key states could alter the outcome, showing the need for monitoring closely divided swing states and precise state-level analysis.

Our model, through poll aggregation, weighted by pollster reliability and sample size, refines polling data into a more realistic prediction, helping offset biases in individual polls. This is because the weighting higher-quality polls reduces the influence of less reliable data seen in unweighted or individual poll estimates, providing a clearer view of voter sentiment. This approach, especially in closely contested swing states, increased the model's sensitivity to

shifts in these pivotal regions, reflecting on their power to overturn the elections while having a historically uncertain favourable candidate.

In the 7 swing states– Arizona, Wisconsin, North Carolina, Pennsylvania, Georgia, Michigan and Nevada, our model shows Harris leading in each, with only 2 swing states showing a lead greater than 1%. Furthermore, the small margins seen in Figure 12 show swing state leads ithin their respective margin of errors, pushing the uncertainty of the estimated lead. Other forecasts like FiveThirtyEight's and The NYTimes' offer similar leads for Harris (FiveThirtyEight 2024a; The New York Times 2024a). When independent models using varied methodologies predict a similar outcome, it can suggest a trend that might hold under different sampling assumptions or adjustments. This convergence across forecasts further emphasizes the tightness of the race however, they can enhance confidence in a potential Harris lead, although it remains small.

The results are can additionally be hard to anticipate due to various unaccounted biases and inaccurate model adjustments. Trump may simply overturn his polls despite Harris' slight forecasted leads due to shy voters, where people don't publicly identify as conservative due to the negative social stigmas around it. What is more likely is an overturn due to the non-response bias. Republicans tend to be less responsive when it comes to polls which can lead to an unrepresentative dataset despite reweighed models altered to account for poor levels of distributed surveys. As there is no guarantee that the altered models work, it is not a confident show of prediction. Harris may also experience a large overturn due to similar inaccurately adjusted estimates in attempts to mitigate seemingly nonrepresentative data to an accurate prediction. However there the election may face incorrect election results due to small surveying or sampling errors, which can be devastating due to the tightness of the race (The New York Times 2024b).

6 Limitations and Next Steps

6.1 Limitations

This analysis identifies several limitations that may impact the reliability of the predictions. First, the logistic model shows high levels of multicollinearity between state and pollster variables. This multicollinearity may inflate standard errors, complicating the interpretation of coefficients and impacting the model's accuracy. Addressing this issue in future analyses could enhance interpretability and offer a clearer comparison to the Bayesian and other models.

Additionally, environmental disruptions, such as recent hurricanes in key swing states, may introduce biases. These events affect both public sentiment and access to polling, potentially skewing results. While the state-level margin of error partially reflects this uncertainty, it also shows the model's sensitivity to external societal influences that are challenging to control. Maine and Nebraska, which allocate some electoral votes by congressional district, are

simplified in this analysis by grouping each state into a single unit. While this approach captures broader state trends, it could underrepresent district-specific voting patterns, and future studies could address district-level data.

The model is fairly simple in comparison to other forecasts and does not currently adjust for other traditional variables or bias (i.e. partisianship) as polling sources, which could lead to an unbalanced or less accurate view of candidate support (Cambridge University Press 2024). Incorporating adjustments for pollster partisanship would improve accuracy by reflecting a more comprehensive cross-section of public opinion.

Another simplification in this analysis is the treatment of Maine and Nebraska, where data were aggregated at the state level instead of by congressional district. While this approach provides an overall view of candidate support in these states, it overlooks district-specific dynamics that could impact electoral vote allocation. This simplification streamlines the model but may underrepresent nuanced trends within these states.

Our method of weighting the polls does not have any constraints and could lead to the over-weighting of certain data groups. If one sample size was much bigger than the majority, its weight would be extreme as well, potentially overcompensating for its affects.

Finally, data gaps remain a significant factor, as polling data for several states were unavailable. Although these states are not swing states, the lack of data introduces uncertainty and potentially limits the model's representativeness. This sparsity in effect with Harris' recent campaign declaration, contributes to an overall margin of error.

6.2 Next Steps

Future models could benefit from greater complexity to address these limitations. Incorporating additional influential factors, such as poll methodology (e.g., online versus in-person surveys), and pollster partisanship, would add depth to the analysis. Multi-level regression and post-stratification approaches, commonly used in polls-of-polls models, have shown to be effective in forecasting elections by capturing uncertainty more comprehensively (Alexander 2023). Increasing model complexity with these factors could improve forecast reliability, particularly in an election as competitive as this one. Integrating historical state-level polling data from previous elections could help fill data gaps and provide stronger context for recent trends. This is especially relevant given Biden's late endorsement of Harris, where past polling data could help address the deficit of recent, state-level Democratic polling information. While this model focuses on the national popular vote, future work could benefit from projections of electoral votes, as these hold greater relevance in determining the election outcome. In practice, reducing the influence of large polls is essential to avoid disproportionately weighting them. Implementing a cap on sample sizes, along with winsorizing extreme values, would prevent larger surveys from skewing results. Similarly, adjusting the weights of pollsters that release multiple polls in a short time frame would balance their impact, ensuring no single pollster overwhelms the dataset's representation (FiveThirtyEight 2024b).

7 Appendix A: Methodology Analysis of The Washington Post Polling

With an evaluation of sampling methodology, recruitment, handling non-response, and questionnaire design, this appendix offers an analysis of the polling methodology used by The Washington Post in collaboration with ABC News. The objective is to analyze these approaches' strengths and weaknesses and determine how they affect polling accuracy.

7.1 Population, Frame, and Sample

The Washington Post, in partnership with ABC News, employs a combination of text-to-web polls and random digit dialing (RDD) for landlines and mobile phones to reach a large and representative sample of American adults and registered voters (The Washington Post 2024a).

With an evaluation of sampling methodology, recruitment, handling non-response, and questionnaire design, this appendix analyzes the polling methodology used by The Washington Post in collaboration with ABC News. The objective is to assess the strengths and weaknesses of these approaches and understand their impact on polling accuracy.

7.2 Population, Frame, and Sample

The target population for The Washington Post and ABC News polls includes U.S. adults and registered voters. The sampling frame, or list from which the sample is drawn, includes landlines, mobile phones, and internet users, covering a broad demographic range. Using a combination of text-to-web polls and random digit dialing (RDD) for landlines and mobile phones, they aim to reach a probability sample where each individual in the population has a known chance of selection. This minimizes sampling bias and improves the representativeness of the sample (The Washington Post 2024a).

The Washington Post's polling averages use only national and state-level polls that comply with strict quality and transparency criteria. These surveys were chosen because they employ suitable stratification and weighting strategies in addition to random sample approaches (The Washington Post 2024b). To represent key demographics such as age, race, gender, and education, the samples are weighted (The Washington Post 2024a) (The Washington Post 2024b).

7.3 Sample Recruitment

Data collection includes live phone interviews and text-to-web surveys, ensuring diverse demographic representation. In a typical 2024 poll, sample collection consisted of 21% text-to-web invites, 15% landlines, and 64% mobile phones (The Washington Post 2024a). This technique facilitates access to younger and minority voters, who may not be well represented in conventional landline-based surveys.

ABC News also uses probability-based recruiting through the Ipos KnowledgePanel by using address-based sampling from the Delivery Sequence File of the US Postal Service. Since internet connections and equipment are offered at no cost, this guarantees that even households without internet connections or digital devices are involved (ABC News 2024a).

7.4 Sampling Approaches and Trade-offs

The Washington Post uses stratified random sampling to ensure that important demographics are represented proportionately to their voter base. Stratified sampling increases the likelihood of a balanced representation, while weighting addresses over- or under-representation of particular groups (The Washington Post 2024a) (The Washington Post 2024b).

When state-level polling data is scarce, The Washington Post incorporates historical voting records from the last two presidential elections (The Washington Post 2024b). This adjustment helps estimate current preferences but depends on past trends, which may miss recent shifts in voter sentiment (The Washington Post 2024a).

7.5 Non-response Handling

Non-response can introduce non-response bias, where the views of non-respondents may differ from respondents, potentially skewing results. The Washington Post addresses this issue using response weighting based on age, race, and education to align the final sample with the overall population distribution (The Washington Post 2023).

ABC News also addressed non-response bias by applying post-stratification adjustments and sending email reminders to non-respondents. In addition, The Washington Post and ABC News ensure that their samples are weighted to account for any anomalies in non-response (ABC News 2024a) (The Washington Post 2023).

Despite these initiatives, non-response bias is still a concern, especially for populations that are less inclined to take part in surveys, including younger or less politically active people (The Washington Post 2023).

7.6 Questionnaire Design

The Washington Post creates its surveys with neutrality and clarity to avoid influencing responses. Questions are rotated to minimize question order bias, and multiple-choice options, including "No opinion," are provided (The Washington Post 2024a). Randomizing questions prevents any unintentional influence on how respondents interpret and answer follow-up questions (The Washington Post 2023).

ABC News follows similar principles, offering surveys in both Spanish and English to ensure inclusivity. Leading questions are purposefully omitted from the questionnaires to better capture genuine public opinion across language barriers (ABC News 2024a).

7.7 Strengths and Weaknesses of the Methodology

Strength:
Sampling Method: The combination of RDD, text-to-web polls, and live phone interviews enable access to a wide demographic, including younger and harder-to-reach voters (The Washington Post 2024a) (The Washington Post 2023).
Post-stratification Weighting: Effective weighting adjusts for demographic imbalances, enhancing poll accuracy and representativeness (The Washington Post 2024a) (ABC News 2024a).
Transparent Approach: Using only high-quality polls and transparency in methodology increases the credibility of The Washington Post's polling data (The Washington Post 2023) (The Washington Post 2024b).

Weaknesses:
Non-response Bias: Despite response weighting, non-response remains an issue, particularly with groups less inclined to participate, such as younger individuals (The Washington Post 2023) (The Washington Post 2024a).
Dependency on Historical Data: In states with limited polling, depending on past election data may not accurately reflect recent shifts in voter preferences (The Washington Post 2024b).

7.8 Conclusion

The Washington Post and ABC News polling methodologies offer a framework for measuring public opinion in the 2024 U.S. presidential election. Employing various sampling strategies, stratification, and weighting improves sample representation. However, challenges such as non-response bias and reliance on historical data in under-polled states need continued adjustments to maintain polling reliability.

8 Appendix B: Idealized Survey & Methodology - $100K Budget

8.1 Overview

With a budget of $100K, this appendix outlines the survey methodology designed to predict the 2024 U.S. Presidential Election. The goal is to collect representative, high-quality data using recruiting, poll aggregation, and selective sample methods. This approach aims to minimize errors commonly encountered in survey research and to produce reliable insights.

8.2 Sampling Approach

We will use stratified random sampling to ensure fair representation across key demographic and geographic subgroups. This approach reduces bias and provides a more accurate view of voter preferences.

Stratification Criteria:
- Age Groups
- Gender
- Education Levels
- Geographic Regions
- Political Affiliation

Sample Size Goal:
- 10,000 respondents across states and demographics to achieve **high statistical power** with a margin of error below $\pm 1\%$

Trade-offs:
- While stratified sampling improves representativeness, it requires precise demographic data, which can increase costs.

Missing Data:
- Specific demographics (e.g., younger males) may have lower response rates. To address this, we will use **post-stratification weighting** and **data imputation**.

8.3 Recruitment Strategy

To ensure a broad demographic reach, we will combine **online recruitment** and **random-digit dialing (RDD)** for telephone surveys. This approach targets both urban and rural populations, increasing inclusivity.

Online Recruitment:
- Target ads on Google, Facebook, and Twitter to engage younger voters and urban populations.
- Budget Allocation: $25,000

Random-Digit Dialing (RDD):
- Phone surveys to engage older and rural voters with limited internet access.
- Budget Allocation: $30,000

Incentives:
- $5 gift cards offered to participants to increase response rates. - Budget Allocation: $20,000

Non-Response Handling:
- Increase incentives for underrepresented groups and send multiple follow-up reminders.

8.4 Data Validation

To ensure data quality, several validation techniques will be implemented:

Survey Logic Check:
- Identifies and flags contradictory responses (e.g., respondents under 18 claiming voter registration).

Attention Check:
- Includes questions to confirm active engagement (e.g., "Select 'Confirm' to proceed")

Post-Stratification Weighting:
- Adjusts the sample for over- and under-enumeration and weighting to reflect the demographic composition of the U.S. population.

Mode and Measurement Errors:
- Enumerators will be trained to reduce inconsistencies between online and phone responses, and survey questions will be straightforward to minimize misreporting.

8.5 Poll Aggregation Methodology

A **poll-of-polls aggregation** method will be applied to reduce bias and balance fluctuation between individual polls. **Weighting Criteria:**
- **Sample Size:** larger samples receive higher weight for greater reliability.
- **Recency:** More recent polls are given higher weight to capture modern voter sentiment.
- **Pollster Rating:** Polls from highly rated pollsters receive higher weights to reduce the impact of bias.

8.6 Survey Implementation

The survey will be created and administered through Google Forms for efficient data collection and secure storage. The main survey sections and example questions are outlined below.

Access the survey: [Google Form](#)

Survey Overview
Title: 2024 US Presidential Election Poll
Purpose: To collect public opinions and forecast election outcomes.
Estimated Time: Less than 5 minutes
Confidentiality: Responses are anonymous and used solely for research

Consent to Participate - I confirm that I understand the purpose of the survey, and I agree to participate.

1. What is your age group?

- Under 18
- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65-74
- 75 and older

2. What is your gender identity?

- Male
- Female
- Non-binary
- Prefer not to disclose

3. What is the highest level of education you have completed?

- Less than high school
- High school diploma or GED
- Some college, no degree
- Associate's degree
- Bachelor's degree
- Master's degree
- Doctorate or professional degree (e.g., MD, JD, PhD)

4. Are you currently registered to vote in the United States?

- Yes
- No, but I am eligible and plan to register before the election
- No, and I am not planning to register
- No, I am not eligible to vote in the United States

5. What is your current political affiliation?

- Republican
- Democrat
- Independent
- Libertarian
- Green Party
- Prefer not to disclose

6. How would you describe your political views?

- Very conservative
- Somewhat conservative
- Moderate
- Somewhat liberal
- Very liberal
- Prefer not to disclose

7. Who do you currently support in the 2024 U.S. presidential election?

- Donald Trump (Republican)
- Kamala Harris (Democrat)
- Undecided
- Other

8. How likely are you to vote in the 2024 presidential election?

- Extremely likely
- Very likely
- Moderately likely
- Somewhat likely
- Not at all likely

8.7 Budget Breakdown

- **Online Recruitment** : $25,000
- **RDD Recruitment**: $30,000
- **Incentives for Participants**: $20,000
- **Data Processing & Validation**: $15,000

- **Miscellaneous Expenses**: $10,000

Total: $100,000

8.8 Conclusion

This survey methodology uses stratified sampling, multi-channel recruitment, and data validation procedures to ensure accurate forecasting of the 2024 US Presidential Election. Through poll-of-polls aggregation, it offers a stable and reliable prediction by mitigating fluctuations in individual polls. The design aims to balance accuracy, inclusivity, and efficiency within a $100K budget, offering meaningful insights into voter behavior and sentiment.

Table 5: Posterior Predictive Check for Model Fit Assessment
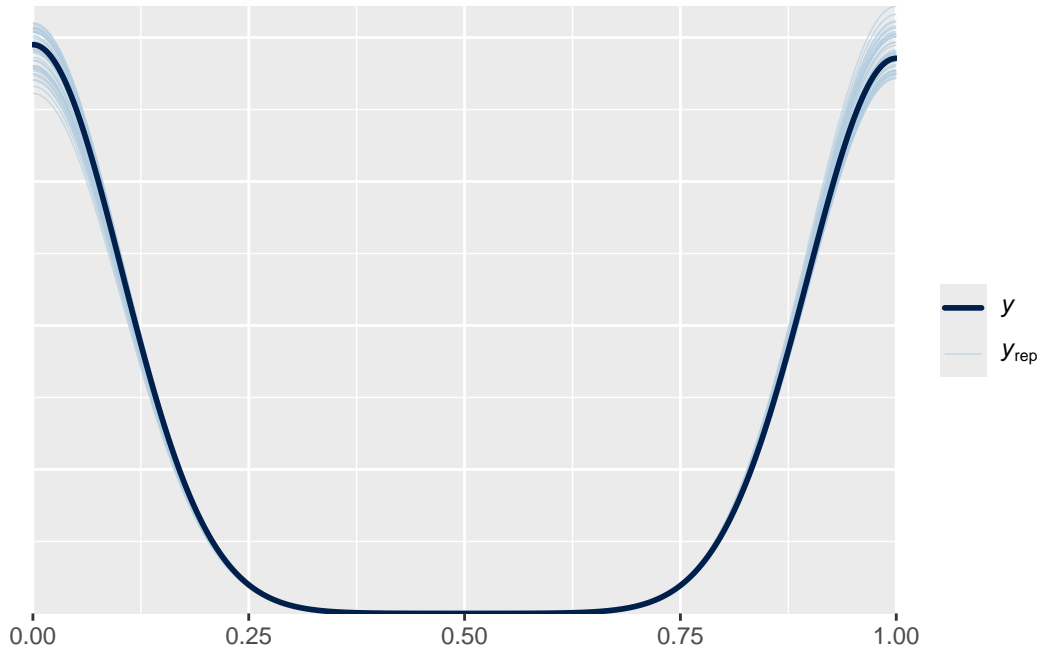
9 Appendix C

- Table 5

References

ABC News. 2024a. "ABC News Polling Methodology Standards." https://abcnews.go.com/US/PollVault/abc-news-polling-methodology-standards/story?id=145373.

———. 2024b. "Hurricanes Helene and Milton Impact on 2024 Election." https://abcnews.go.com/538/hurricanes-helene-milton-affect-2024-election/story?id=114783517.

Alexander, Rohan. 2023. "Telling Stories with Data." Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Andrew Mercer. 2018. "Methods 101: Margin of Error." https://www.pewresearch.org/fact-tank/2018/08/17/methods-101-margin-of-error/.

Associated Press. 2024. "Election Polls and Forecasts." https://apnews.com/article/election-polls-kamala-harris-donald-trump-b497bd1015d35ed563e085c8db075802.

Becker, Richard A., Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka, and Alex Deckmyn. 2023. *Maps: Draw Geographical Maps.* https://CRAN.R-project.org/package=maps.

Ben-Haim, Yaniv et al. 2023. *MLmetrics: Machine Learning Evaluation Metrics.* https://CRAN.R-project.org/package=MLmetrics.

Bloomberg. 2024. "Trump and Harris Tied in Key States." https://www.bloomberg.com/news/features/2024-10-23/new-poll-has-trump-harris-tied-in-key-states-just-12-days-to-election.

Cambridge University Press. 2024. "Information, Incentives, and Goals in Election Forecasts." *Judgment and Decision Making.* https://www.cambridge.org/core/journals/judgment-and-decision-making/article/information-incentives-and-goals-in-election-forecasts/CDED59D35F6599BF9D31348A79B196F9.

CBS News. 2024. "Trump-Harris Campaign Promises in the 2024 Election." https://www.cbsnews.com/news/trump-harris-campaign-promises-2024-election/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

FiveThirtyEight. 2024a. *FiveThirtyEight: 2024 US Presidential Election Polls.* https://projects.fivethirtyeight.com/polls/president-general/2024/national/.

———. 2024b. "How Our Polling Averages Work." https://fivethirtyeight.com/methodology/how-our-polling-averages-work/.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Grolemund, Garrett, and Hadley Wickham. 2023. *Lubridate: Make Dealing with Dates a Little Easier.* https://CRAN.R-project.org/package=lubridate.

Human Rights Watch. 2022. "U.s. Supreme Court Overturns Roe v. Wade." https://www.hrw.org/news/2022/06/24/us-supreme-court-topples-roe-v-wade-blow-rights.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

The New York Times. 2024a. "2024 Presidential Election Polls." https://www.nytimes.com/interactive/2024/us/elections/polls-president.html.

———. 2024b. "Opinion | Why This Year's Election Polls May Be Less Predictive Than Usual." https://www.nytimes.com/2024/10/23/opinion/election-polls-results-trump-harris.html.

———. 2024c. "Trump-Harris Polls and the Tight 2024 Election." https://www.nytimes.com/2024/10/21/upshot/trump-harris-polls-election.html.

The Washington Post. 2023. "Polling Methods and Standards (2023)." https://www.washingtonpost.com/documents/f409a864-d58c-430a-8adc-ee6a8a344854.pdf?itid=lk_inline_manual_6.

———. 2024a. "ABC News/the Washington Post Polling Methodology (2024)." https://www.washingtonpost.com/documents/b4c28200-18a7-412e-a4da-4e8cf9b91fd5.pdf.

———. 2024b. "Presidential Polling Averages Methodology." https://www.washingtonpost.com/elections/2024/06/26/presidential-polling-averages-methodology/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, and Garrett Grolemund et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-project.org/package=kableExtra.