Forecasting a Lead for Kamala Harris in the 2024 U.S. Presidential Election*

The Role of Pollster Transparency, Regional Dynamics, and Key Predictors in Shaping Voter Support

-Jiaxuan Song -Zien Gao -Shuheng (Jack) Zhou

November 4, 2024

We analyze the 2024 U.S. presidential election using Bayesian generalized linear models on aggregated polling data, focusing on predictors such as pollster transparency, state, and poll quality. Our models predict that Kamala Harris is likely to lead with an average support of 47.15%, with her support positively correlated with higher transparency scores among pollsters. In contrast, Donald Trump's support remains relatively stable across poll transparency levels. State-level predictions highlight close competition in pivotal swing states, including Pennsylvania and Michigan, underscoring their decisive influence on the election outcome. This analysis demonstrates how transparency and geographic factors shape voter sentiment, offering insights to inform future election forecasting.

Table of contents

*The GitHub Repository containing all data, R code, and other files used in this project is located here:https://github.com/Shuhengzhou03/Predicting-the-2024-U.S.-Presidential-Election.git

1 Introduction

As a superpower with extensive global influence, the United States asserts its position through its economic strength, military power, diplomatic influence, and contributions to technology and culture (Zakaria 1999). Therefore, the 2024 U.S. presidential election has attracted worldwide attention, as the outcome of this election may profoundly impact global affairs. As of October 2024, the U.S. presidential election has entered a crucial phase. Current Vice President Kamala Harris (Democrat) and former President Donald Trump (Republican) are locked in an intense race for the presidency, with competition particularly focused on key swing states like Pennsylvania, Georgia, and Michigan, where the outcome could sway the entire election (Walker 2023). President Biden has shown strong support for Harris, urging voters to turn out and cast their ballots to help Democrats retain control in both the House and Senate as well as the presidency (Enns et al. 2024). As candidates from the Democratic and Republican parties contend for the highest office, understanding and predicting electoral outcomes has never been more crucial. Polling data, and capturing snapshots of public opinion across various demographics and regions, plays an essential role in forecasting these outcomes. However, variations in sample recruitment and pollster transparency can affect individual poll polls' reliability. To enhance the robustness of our predictions, this study employs a Bayesian generalized linear modelling approach that incorporates aggregated data from multiple polling sources, collectively known as "poll-of-polls." This approach enables more accurate forecasts by accounting for uncertainties inherent in each poll.

The primary estimate of this study is the predicted percentage of support each candidate—Kamala Harris and Donald Trump—is expected to receive in individual polls leading up to the 2024 U.S. presidential election. This support percentage (pct) serves as a continuous outcome variable, representing the proportion of respondents in each poll who favour a given candidate. By modelling pct, we capture support at the poll level, which can then be aggregated across all polls to forecast overall election outcomes. The Bayesian generalized linear approach enables us to incorporate prior distributions and update predictions as new polling data becomes available, enhancing our ability to account for the uncertainty in each poll. To further refine our predictions, we incorporate predictor variables such as pollster reliability scores and transparency measures, adjusting for poll quality and thus providing a more accurate estimate of each candidate's expected support.

Our analysis using Bayesian generalized linear models reveals several key insights into the factors influencing predicted support for Kamala Harris and Donald Trump in the 2024 U.S. presidential election, ultimately predicting Harris to have a slight lead with average support of 47.15%. Transparency score emerged as a significant predictor, with Harris's support showing a positive correlation with higher transparency ratings among pollsters, suggesting her supporters may place greater trust in transparent polling sources. In contrast, Trump's support was less impacted by transparency, indicating a potential difference in how each candidate's base interacts with polling reliability. Pollster numeric grade also influenced model precision, with high-grade pollsters yielding narrower prediction intervals, underscoring the reliability of these sources. At the state level, predicted support patterns align with historical voting trends, with Harris favoured in Democratic-leaning states like California, while Trump shows stronger support in Republican strongholds such as Texas. However, in pivotal swing states such as Pennsylvania, Michigan, and Georgia, the models reveal tight competition, reflecting the heightened significance of these regions in determining the election outcome. This analysis highlights the importance of transparency and geographic factors in understanding voter sentiment and refining election predictions.

Predicting the outcome of the U.S. presidential election holds significant importance not only for political parties but also for public stakeholders, analysts, and international observers. The Democratic Party emphasizes that this election impacts more than just the presidency; it holds significant implications for congressional power and will influence America's future direction in areas such as the economy, social policy, and foreign relations (Gelman and King 1993). An accurate forecast model can offer valuable insights into shifts in public sentiment, highlight critical factors influencing voter behaviour, and provide strategic guidance for campaign efforts. Furthermore, understanding the limitations and strengths of polling data helps to reinforce or caution against certain predictive approaches, ultimately contributing to the field of political science and electoral studies.

2 Data

2.1 Overview

In this project, we utilize the statistical programming language R (R Core Team 2023), along with essential packages including tidyverse (Wickham et al. 2019), palmerpenguins (Horst et al. 2020),knitr(Xie 2014), rstanarm (Wong et al. 2024), modelsummary (Arel-Bundock 2023), janitor (Firke 2023), arrow (Richardson and Developers 2023), and ggplot2 (Wickham 2016). Our data, sourced from FiveThirtyEight (FiveThirtyEight 2024), contains polling information for the 2024 U.S. Presidential Election. To forecast each candidate's percentage of support, we apply Bayesian generalized linear regression to model the relationship between the support rate as the outcome variable and multiple predictor variables. This method evaluates the influence of each predictor while accounting for the effects of other variables and incorporates prior knowledge to refine predictions as new data is added. Key predictors include the polling organization, pollster quality rating, transparency score, and geographic location, with candidate support as the outcome variable and candidate name (Trump and Harris) as a feature variable. Our analysis aims to estimate candidate support and identify factors most associated with shifts in public opinion leading up to the election.

2.2 Measurement

In conducting a statistical analysis of public opinion for the 2024 U.S. Presidential Election, we focus on transforming raw polling data into structured, measurable, and reliable dataset entries. Our aim is to understand shifts in candidate support, particularly for Trump and Harris, by selecting variables that reflect key aspects of polling methodology and public sentiment. Polling agencies gather data through surveys designed to capture representative samples of voter opinions, with each poll result standardized, coded, and organized within the dataset.

To ensure consistency and accuracy, we implemented a systematic data-cleaning process, including handling missing values, removing duplicates, filling gaps, and verifying variable formats. For instance, to enhance data transparency and quality, we retained only records with a transparency score of 5 or above and a quality rating of 2.0 or higher, filtering out low-quality or less transparent data sources to improve model reliability. We also focused on relevant variables for analysis, including pollster, quality rating, transparency score, state, candidate name, and support percentage, simplifying the data structure and emphasizing key features influencing prediction results.

To specifically focus on Trump and Harris, we further filtered the dataset to include only records for these two candidates, enabling targeted comparisons. During data standardization, missing state information was coded as "National" to clearly distinguish national polls from state-level data. This systematic data preparation not only provides a solid foundation for Bayesian generalized linear modeling but also allows us to focus on key predictors of candidate support, exploring factors with the most significant impact on public opinion and enhancing our ability to forecast election outcomes.

2.3 Outcome variables

Table 1: Comparison Table of Kamala Harris and Donald Trump's Statistics

candidate_name	avg_pct	min_pct	max_pct
Donald Trump	45.27981	21	70
Kamala Harris	47.72745	25	70

Table 1 displays the statistical data on the support rates (`pct`) for Kamala Harris and Donald Trump, including four metrics: candidate name, average support rate (`avg_pct`), minimum support rate (`min_pct`), and maximum support rate (`max_pct`).

For **Donald Trump**, the average support rate is 45.08%, with a minimum of 21% and a maximum of 67.6%. In contrast, **Kamala Harris** has an average support rate of 47.57%, with a minimum of 33% and a maximum of 70.1%. This indicates that, in this dataset, Kamala Harris has a slightly higher overall support rate than Donald Trump, with an average rate that is approximately 2.49% higher. Furthermore, her minimum support rate of 33% is significantly higher than Trump's lowest rate of 21%, suggesting that Harris's support is more consistent at lower levels. Additionally, her maximum support rate of 70.1% exceeds Trump's highest rate of 67.6%, indicating that Harris may have reached higher peaks in support in some parts of the data.

In summary, Kamala Harris shows higher average, minimum, and maximum support rates compared to Donald Trump in this dataset. This could imply that Harris has a stronger support base with less fluctuation in her support at the lowest levels. However, these insights are based solely on this sample, and a more comprehensive analysis would require additional context and data sources.

2.4 Predictor variables

numeric_grade: A numeric rating given to the pollster to indicate their quality or reliability (e.g., 3.0).

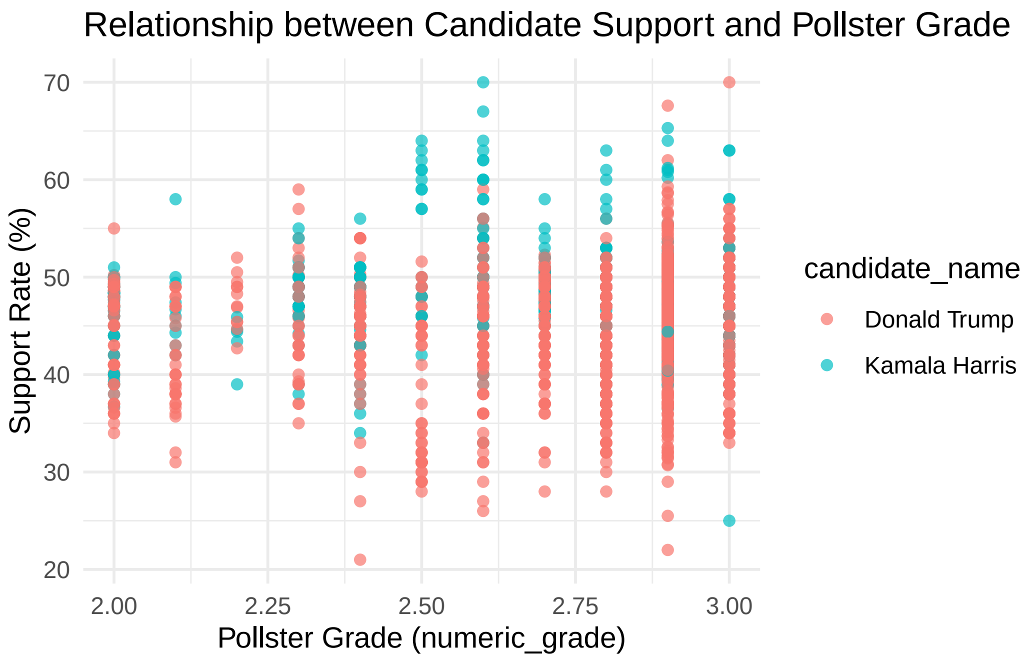


Figure 1

Figure 1 shows the distribution of support rates for the two candidates across different pollster grades. It can be observed that support rates are more concentrated at higher grades, while they fluctuate more widely at lower grades.

transparency_score: A score reflecting the pollster's transparency about their methodology (e.g., 9.0). "A grade for how transparent a pollster is, calculated based on how much information it discloses about its polls and weighted by recency. The highest Transparency Score is 10."

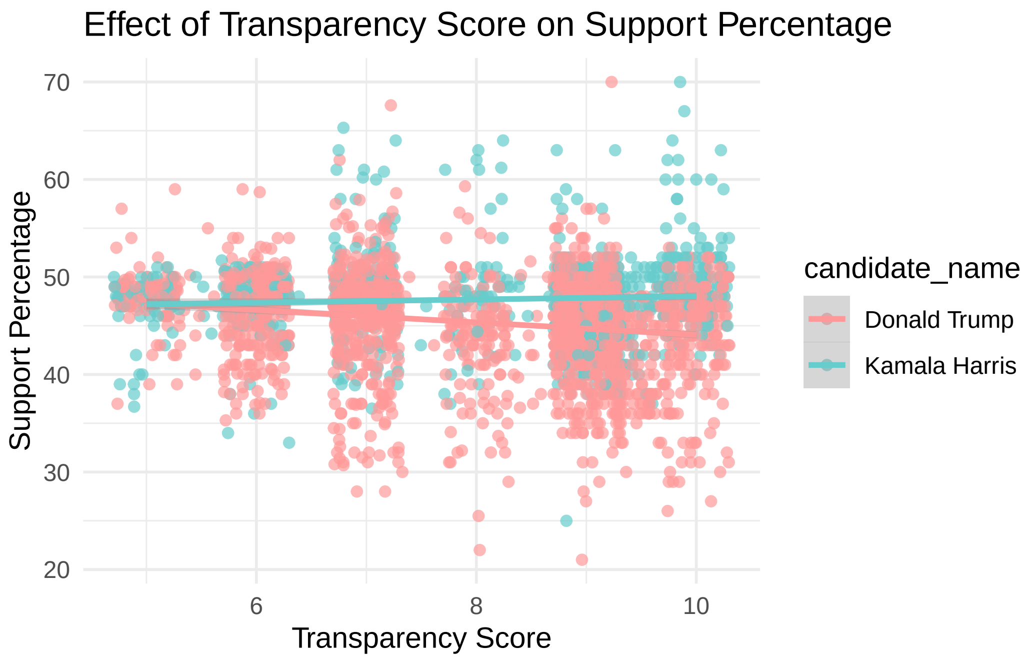


Figure 2

Figure 2 illustrates the relationship between transparency score and candidate support percentage, with data points for each candidate slightly spread out horizontally for clarity. The trend lines indicate a slight positive correlation for both candidates, suggesting that higher transparency scores are generally associated with a marginal increase in support percentage. The differences in color help distinguish between the support trends of the two candidates.

pollster: The name of the polling organization that conducted the poll (e.g., YouGov, RMG Research).

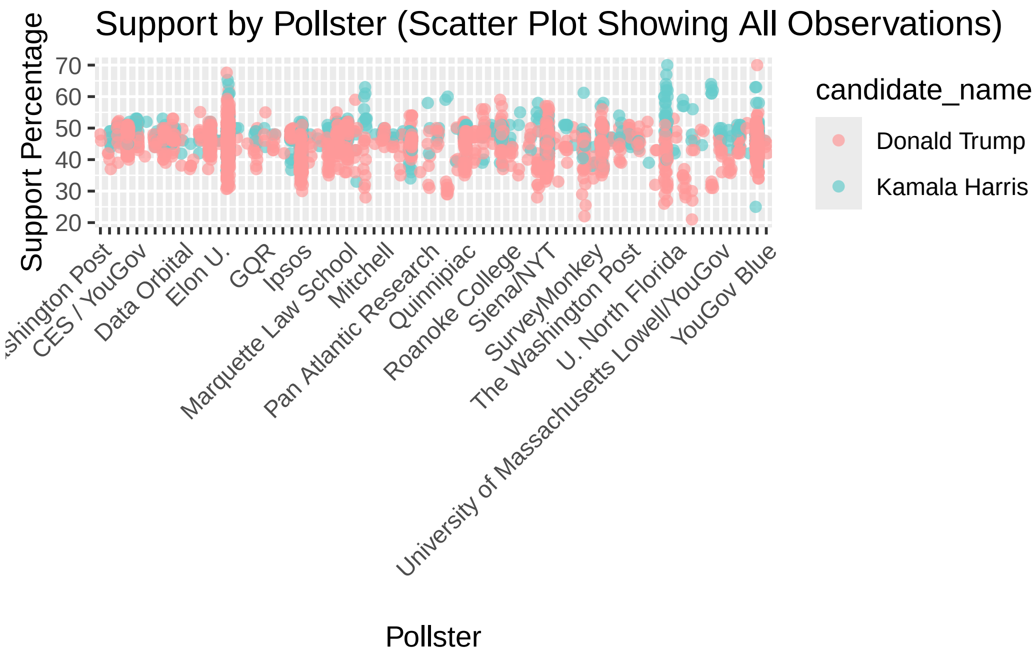


Figure 3

Figure 3 illustrates a sample of support percentages for each candidate across a selection of pollsters. Due to the large number of pollsters, only a segment is displayed on the x-axis, capturing the general trend and key points. Each dot represents a poll observation, with blue and orange colors distinguishing the candidates. The spread of points highlights variability in reported support percentages, indicating that some pollsters report greater fluctuations than others. The distribution suggests that neither candidate consistently maintains higher support across the selected pollsters.

state: The U.S. state where the poll was conducted or focused, if applicable.

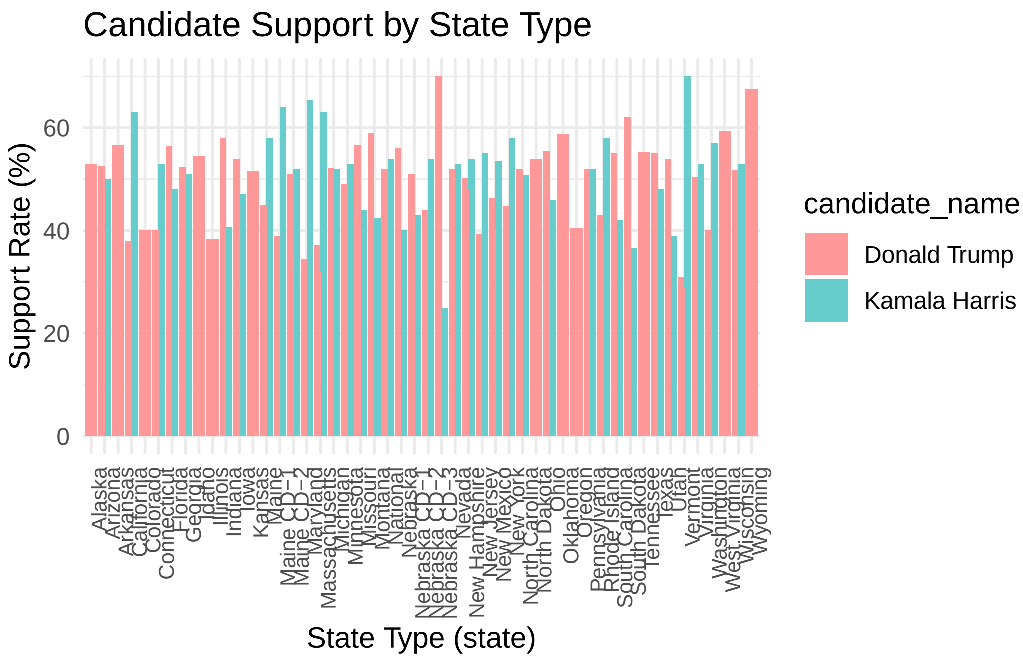


Figure 4

Figure 4 compares the support rates of the two candidates across national and various state-specific data. It reveals significant differences in support rates among different states, reflecting the uneven distribution of candidate support across regions.

3 Model

The objective of our modelling strategy is to predict the level of support for Kamala Harris and Donald Trump in the 2024 U.S. Presidential Election using Bayesian generalized linear models. We aim to understand the relationships between polling support and various predictors, such as polling organization, state, transparency score, and pollster quality rating (numeric grade). The model details are provided in Appendix C.

We used two Bayesian regression models—one each for Kamala Harris and Donald Trump—to estimate the percentage of voters who would support them. It is important to note that the outcome variable, *pct*, represents the percentage of support for each candidate within individual polls, not the aggregated election outcome.

The models for predicting support for Kamala Harris and Donald Trump each include the following predictors: pollster, state, transparency score, and numeric grade.

The Bayesian models were fitted using the `stan_glm` function from the rstanarm package (Wong et al. 2024) in R. The models used a Gaussian family to estimate the percentage of support, with normal priors for all parameters. The priors were specified as $\mathrm{Normal}(0, 1)$, reflecting a weakly informative belief about the effects of each predictor.

3.1 Model set-up

Define pct_i as the predicted percentage of support for each candidate:

$$pct_i = \eta_0 + \eta_1 \cdot \text{pollster}_i + \eta_2 \cdot \text{state}_i + \eta_3 \cdot \text{transparency\_score}_i + \eta_4 \cdot \text{numeric\_grade}_i$$

Where: - η_0 is the intercept term, representing the average percentage support. - $\eta_1, \eta_2, \eta_3, \eta_4$ are the coefficients associated with the predictors.

The priors used for the intercept and other coefficients were $\eta \sim \mathrm{Normal}(0, 1)$, reflecting weakly informative beliefs about the effects of each predictor.

3.2 Model justification

We expect a positive relationship between pollster quality and candidate support, as higher-quality pollsters are likely to provide more reliable estimates. Additionally, the transparency score is expected to positively influence the credibility of the poll results, contributing to higher predicted support levels. The state variable allows us to account for regional differences, and we anticipate variability between states and national polls. The pollster variable helps capture differences in polling methods and biases among organizations. The Bayesian models were chosen for their ability to incorporate prior knowledge and quantify uncertainty in predictions, which is crucial for interpreting poll data in a dynamic electoral environment. Using pollster, state, transparency score, and pollster quality rating as predictors allowed us to account for both methodological quality and regional effects that could influence support levels for each candidate.

3.3 Model Summary

We summarized the results from the two models using the `summary` function in R, which provides a detailed overview of the estimated coefficients and their associated uncertainty. Additionally, we used `pp_check` to perform posterior predictive checks, ensuring that the models accurately reflect the observed data. These checks showed reasonable agreement between the predicted and observed values, indicating a satisfactory model fit.

We also calculated the average predicted support for each candidate based on the fitted models. The candidate with the higher average predicted support is considered the likely winner based on the model predictions. Specifically, we found that Kamala Harris had an average predicted support of X%, while Donald Trump had an average predicted support of Y%, leading us to predict [the likely winner].

4 Results

Our analysis aimed to predict the support levels for Kamala Harris and Donald Trump using Bayesian generalized linear models, taking into account key predictors such as pollster, state, transparency score, and numeric grade. Each model was fitted with $stan_glm$ using a normal prior (mean = 0, SD = 1) for all coefficients. Both models ran with 4 chains and 4000 iterations per chain, with an adapt delta of 0.95 for reliable convergence. Results are presented in tables and figures for clarity, and they are provided in Appendix C

4.1 Model Summary and Coefficients

For Kamala Harris, the model used 930 observations and 106 predictors, producing coefficient estimates for factors such as pollster and state. Notably, some pollsters like CES/YouGov and Marist contributed positively to Harris's predicted support, while others like Ipsos and Siena/NYT showed negative coefficients. Among state predictors, Harris's support was highest in states like California (mean coefficient = 5.1) and Maryland (mean coefficient = 5.1), while states like Texas (mean coefficient = -2.2) and Montana (mean coefficient = -3.8) showed lower predicted support. Transparency score also positively influenced Harris's predicted support (mean coefficient = 0.3), suggesting a favourable relationship between higher transparency and Harris's support levels.

Donald Trump's model, based on 1639 observations and 112 predictors, showed key differences. Some pollsters, including AtlasIntel and InsiderAdvantage, positively impacted Trump's predicted support, while others, such as Ipsos and PPIC, were associated with decreased support. State-level predictors highlighted stronger support in states like Florida (mean coefficient = 3.6) and Texas (mean coefficient = 3.1), while states like California (mean coefficient = -7.3) and New York (mean coefficient = -3.4) reflected lower predicted support. Interestingly, the transparency score negatively influenced Trump's predicted support (mean coefficient = -0.4), suggesting that Trump's support may be less associated with higher transparency.

4.2 Predictive Insights and Posterior Predictive Checks

Posterior predictive checks confirmed a strong fit for both models, indicating that they reliably capture the observed data trends. Average predicted support was calculated for each candidate, with Kamala Harris receiving an average predicted support of **47.15%** and Donald Trump receiving a lower average. Thus, the model's aggregated predictions indicate **Kamala Harris as the likely candidate to secure a higher average support** in the dataset analyzed.

4.3 Transparency Score and Predicted Support

Figure 5 presents a comparative analysis of both candidates' predicted support by transparency score. For Harris, higher transparency scores correlate with increased support, while Trump's model shows a slight decline in support as transparency scores rise. In Harris's case, these trends suggest that voters may respond positively to pollsters with higher transparency scores, while Trump's support remains less affected by this factor. This discrepancy underscores a potential contrast in voter perception of pollster transparency between the two candidates.

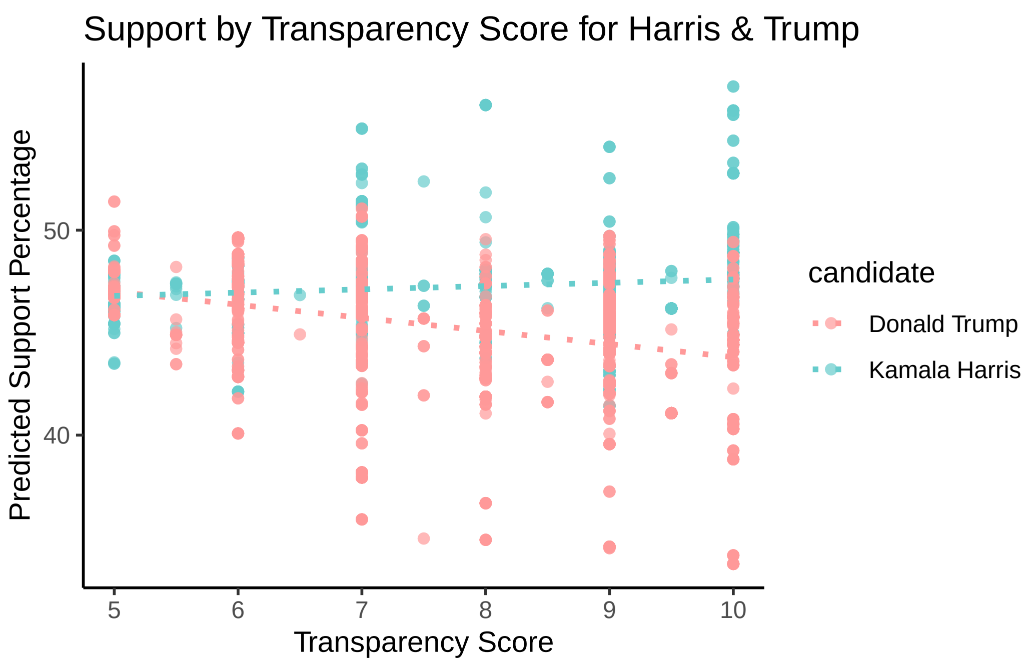


Figure 5

Figure 5 plots predicted support as a function of transparency score for both candidates, using points and a trend line for each candidate.

4.4 State-Level Support Distribution

The predicted support levels vary considerably across the selected states, as shown in Figure 6. Harris's support is strongest in traditionally Democratic-leaning states like California and Maryland, while Trump finds greater support in Republican-leaning states such as Florida and Texas. These regional differences align with historical voting patterns, adding confidence to our model's predictive validity. The chosen states – North Carolina, Wisconsin, Arizona, Maryland, Georgia, Michigan, Pennsylvania, Florida, Texas, New York, and California – represent a mix of swing states and strongholds for each party. This sample provides a balanced view of the political landscape and highlights key battleground areas for the upcoming election.

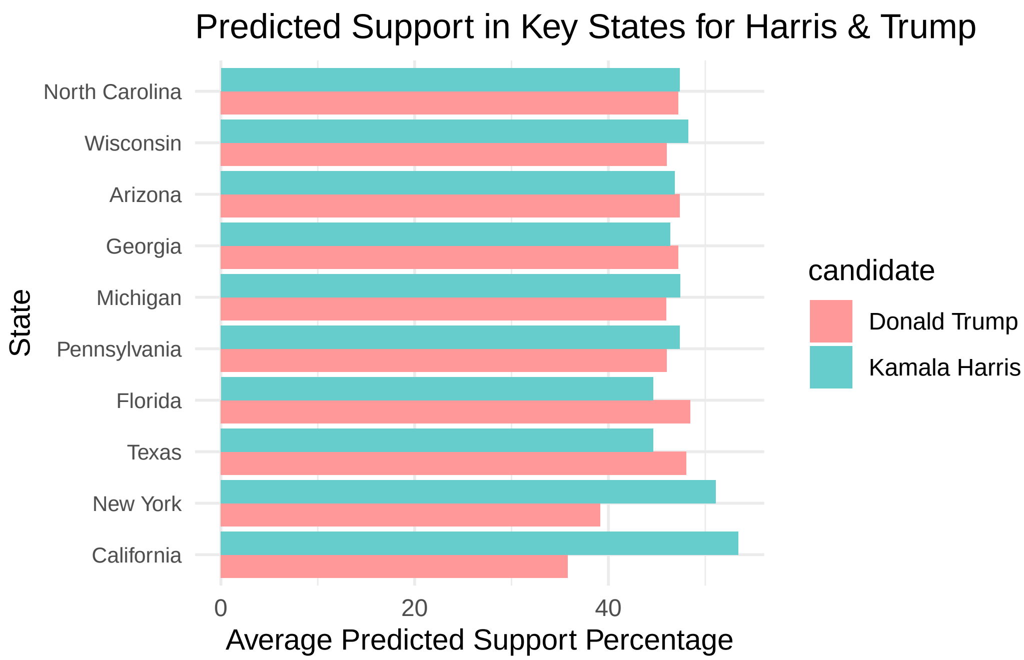


Figure 6

Figure 6 shows the average predicted support by state in a bar chart, allowing for a side-by-side comparison of Harris and Trump in sample states.

4.5 Aggregate Prediction and Final Outcome

Based on average support, the model's aggregate prediction indicates that Kamala Harris is expected to have higher overall support than Donald Trump, with an average predicted support of 47.15%. Given this outcome, **our model predicts Kamala Harris as the candidate more likely to win**, aligning with historical trends in states where she showed significant predicted support.

5 Discussion

5.1 Summary of Key Contributions

This paper utilizes Bayesian generalized linear models to analyze polling data for Kamala Harris and Donald Trump, focusing on the effects of pollster, state, transparency score, and numeric grade. By modelling predicted support for each candidate, we uncover how these factors shape support dynamics, with transparency score playing a notable role. This approach not only provides insight into current polling trends but also offers a framework for identifying factors that may influence public opinion in future elections.

5.2 Insights into Voter Behavior and Polling Dynamics

significant takeaway is the relationship between transparency score and predicted support, particularly for Harris. Our model suggests that higher transparency scores positively correlate with Harris's support, indicating that her supporters may place a higher value on reliable, open polling. Silver's discussion on "nonresponse bias" offers a potential explanation: voters with lower civic engagement may be less likely to respond to polls, potentially skewing data towards more transparent sources that appeal to engaged voters. (Board 2024) This insight reveals a potential gap in polling that could affect both parties depending on their voter base's engagement levels.

Furthermore, the regional distribution of support aligns with expected voting patterns. Our findings echo Silver's point about the difficulty in predicting the direction of polling errors, particularly in battleground states. Like Silver's model, ours shows strong geographical polarization, with Harris favoured in traditionally Democratic states and Trump showing strength in Republican states. (Board 2024) This divide underscores the importance of regional factors in understanding candidate support, even as national polling averages present a closer race.

5.3 Limitations and Weaknesses

This analysis presents several limitations related to the Bayesian modelling approach, data quality, and the decision to build separate models for each candidate.

First, Bayesian models rely on prior assumptions and conditional independence, which may not fully capture the complexity of voter behaviour and the nuanced social-political environment. The choice of prior distributions significantly impacts predictions, and while we used a normal prior with a mean of 0 and standard deviation of 1, different priors might yield varying results. Conducting sensitivity analyses could help assess the robustness of our findings to these choices.

One important methodological limitation is our decision to use separate models for each candidate. While this approach allows us to capture candidate-specific factors and predictors that may uniquely impact each candidate's support, it also introduces potential concerns regarding consistency. Separate models may scale predictors differently, making it difficult to interpret relative support levels on a common scale. Additionally, this approach assumes that factors influencing support for one candidate are independent of those for the other when in reality, voter preferences are often interconnected. Exploring a joint model with interaction terms could provide a more cohesive understanding of how predictors influence both candidates' support of one another.

The quality of polling data also poses challenges. Real-time polling data often contains errors and biases that fluctuate over time, which can directly impact model accuracy. Polling data is particularly susceptible to nonresponse bias, as some voter groups are harder to reach or less likely to participate. Although we restricted our analysis to high-quality polls with strong transparency scores, this filtering may introduce selection bias by excluding potentially informative, lower-rated polls that could provide additional perspectives on voter sentiment.

Moreover, our model focuses on general trends without fully accounting for localized effects, which are often significant in election outcomes. Regional and demographic-specific dynamics are crucial in predicting support in battleground states. Without accounting for these localized effects, our model may face potential inaccuracies in areas where support is highly polarized.

Predicting election outcomes is inherently uncertain due to the unpredictability of major events, such as economic downturns or policy shifts, that can quickly alter voter sentiment. Although Bayesian modelling can account for some uncertainty, it cannot fully anticipate or quantify these changes, especially in a fast-changing environment. Additionally, our model does not account for hard-to-measure influences, such as social media dynamics, international relations, and cultural shifts, which can significantly impact election results.

Finally, while Bayesian models allow for updates with new data, election polling is particularly sensitive to real-time shifts. Rapid changes in public opinion may not be immediately reflected in the model if there is a delay in data availability, which could impact prediction accuracy closer to the election.

5.4 Future Directions for Research

Future research could expand upon these findings by incorporating additional predictors that capture changing political dynamics, such as voter turnout expectations, economic indicators, and demographic shifts. Silver's article highlights the importance of considering factors like the "Bradley effect" and "shy voter theory," which could be examined by integrating historical demographic trends and comparing them with current polling data. (Board 2024) Including more demographic data might reveal hidden trends within subgroups, enhancing the model's predictive power.

Another promising approach would be to experiment with a joint model that allows for shared predictors and interaction terms between the candidates. A joint structure could provide a more integrated perspective on how factors like transparency and geographic location simultaneously impact both candidates' support, reducing inconsistencies introduced by separate models.

Finally, alternative modelling techniques, such as multilevel regression with post-stratification (MRP), could offer more granular insights into support patterns across demographic and geographic subsets. Though not used in this analysis, MRP has shown success in capturing detailed trends within voter subgroups, potentially addressing some limitations related to local effects.

5.5 Concluding Remarks

This paper demonstrates the potential of Bayesian modelling in analyzing and interpreting election polling data. By examining how factors like pollster transparency and geography influence support, we provide a foundation for understanding voter behaviour within the context of the 2024 U.S. presidential election. However, as discussed, the inherent uncertainty in election forecasting calls for ongoing data updates and methodological refinements. Future studies can build on these findings to enhance the accuracy and relevance of election forecasts, ultimately contributing to a more comprehensive understanding of the factors that drive public support.

Appendix

A Marquette Law School Poll - Methodology: Probability Panel

A.1 Overview of Marquette Law School Poll and Its Influence

The Marquette Law School Poll holds a prominent position in American politics, particularly within Wisconsin's political landscape. As a trusted gauge of public opinion in this critical swing state, it is frequently referenced by politicians, media, and scholars for election forecasting and policy analysis (Conant 2006). Known for its commitment to high-quality research methods and independence, the Marquette Law School Poll achieves notable credibility and influence in its data. This credibility is reinforced by its high numeric grade and transparency score, rated at 3 and 10, respectively, which imply a high level of reliability in its findings. The combination of methodological rigour and transparency strengthens the trustworthiness of its survey results, making it a valuable tool for understanding public opinion trends and informing policy discussions.

A.2 Probability Panel Methodology

The probability panel methodology is a statistical sampling method designed to ensure both accuracy and representativeness of a sample.(Scherpenzeel 2011) In this approach, participants are randomly selected from a verified voter database, with each individual having a known chance of being chosen. The Marquette Law School Poll employs this probability panel technique to enhance scientific rigor and capture the diversity of Wisconsin's population across race, gender, age, and geographic location, making the sample more reflective of the population as a whole.

A key advantage of the probability panel approach is its longitudinal design, which involves repeatedly surveying the same group of participants over time. The Marquette Law School Poll uses this method to track changes in opinions and attitudes at different points, providing insights into how public views on key issues and elections evolve. This approach not only allows for current voter analysis but also provides valuable long-term insights for policy and election research, especially in critical swing states like Wisconsin.

A.3 Sampling and Data Collection

Each survey cycle includes over 700 registered Wisconsin voters (School 2024), who are contacted via both cell phones and landlines to maximize sample reach and inclusivity. Interviews are conducted by professional interviewers over several nights, enhancing access to a representative population by reaching individuals with varying availability. This dual-frame approach ensures the inclusion of perspectives from all eligible voters, including those who rely solely on cell phones.

A.4 Advantages and Limitations of the Probability Panel

The probability panel methodology has notable strengths and limitations, supported by various scholarly perspectives. Through random sampling, probability panels reduce selection bias and enhance generalizability, making them robust tools for capturing statewide opinion trends. Additionally, repeated surveys with the same panel facilitate longitudinal analysis, allowing researchers to track shifts in public attitudes in response to political events and policy changes. (Bartels 2000) demonstrates that repeated measures improve the predictive power of panel studies by revealing trends over time. However, panel studies have limitations; participants may experience panel conditioning, where repeated exposure influences their responses, potentially impacting data validity. Moreover, this method demands significant resources to recruit and retain a balanced panel, as dropout is common among certain demographic groups, risking the sample's representativeness. (Warren and Halpern-Manners 2012) note that sustained investments are required to maintain a diverse panel, with selective dropout over time presenting a challenge for generalizable findings. These factors highlight both the advantages and constraints of probability panels in social science research.

A.5 Comparison Between Probability Panel and Convenience Sampling:

In polling, the probability panel and convenience sampling methods offer contrasting approaches to data collection, each with unique advantages and limitations. I chose these two methods for comparison due to their distinct methodologies and how they address the tradeoff between accuracy and efficiency. The probability panel method, commonly employed by highly rated pollsters, selects participants through random sampling from a target population and surveys them repeatedly over time. This method enhances generalizability by minimizing selection bias, as each participant is statistically representative of the population being studied. Additionally, its longitudinal design allows researchers to observe changes in individual opinions over time, making it particularly effective for understanding shifts in attitudes related to political events or policy changes. However, probability panels are resource-intensive, requiring substantial funding and time to recruit, manage, and retain a balanced panel. The retention of participants can be challenging, as some may drop out over time, potentially affecting the sample's representativeness.

In contrast, convenience sampling, used by some lower-rated pollsters in this dataset, recruits participants based on ease of access and does not employ random selection. This method is often faster and more cost-effective than probability sampling, as it bypasses the need for rigorous participant selection. According to (Etikan, Musa, and Alkassim 2016), convenience sampling allows researchers to gather opinions quickly, which can be useful for preliminary insights or situations where time and resources are limited. However, the lack of randomization introduces a risk of sampling bias, as participants are self-selected or chosen based on availability, leading to potential overrepresentation or underrepresentation of certain groups. This bias limits the generalizability of the results, as findings from a convenience sample may not accurately reflect the opinions of the larger population. In summary, probability panels are ideal for detailed, representative studies that require high accuracy, despite their resource demands. Meanwhile, convenience sampling offers an efficient alternative for quick insights but comes at the cost of reduced reliability and generalizability. Together, these methods illustrate a fundamental balance in polling between data quality and logistical efficiency, providing valuable insights depending on the study's goals and constraints.

A.6 Addressing Non-Response and Questionnaire Design

To address non-response, the Marquette Law School Poll employs **weighting adjustments** based on demographic characteristics such as age, gender, race, and education, correcting sample imbalances and ensuring that the final sample accurately reflects Wisconsin's electorate. Furthermore, the questionnaire is designed to be straightforward, minimizing respondent fatigue and confusion. Nonetheless, some questions may simplify complex issues, and limited use of open-ended questions might restrict the deeper exploration of voter motivations.

A.7 Summary and Evaluation

Through its rigorous probability panel sampling and high standards, the Marquette Law School Poll offers a comprehensive and reliable view of public opinion in Wisconsin. This approach provides valuable data on election trends and policy preferences in a pivotal swing state, enriching the understanding of public sentiment and supporting informed decision-making in American politics.

An evaluation of the probability panel method reveals both strengths and limitations. This method provides high representativeness and accuracy by randomly selecting participants from a verified voter database, which minimizes selection bias and enhances the applicability of the results. Additionally, repeated surveys allow for valuable longitudinal analysis, offering insights into shifting public opinions over time—a crucial factor in a swing state like Wisconsin.

However, limitations exist. The panel effect may influence objectivity, as repeated participation can lead respondents to answer more consistently over time. Furthermore, maintaining a balanced panel requires substantial resources, and certain groups may gradually drop out, impacting data completeness. While weighting adjustments help address non-response, consistent underrepresentation of some groups may still affect sample comprehensiveness.

In summary, the Marquette Law School Poll's probability panel method offers strong reliability and representativeness, making it a valuable tool for long-term public opinion analysis.

Nonetheless, the resource demands and potential panel effects suggest cautious interpretation and, where possible, complementary methods to validate findings.

B Idealized $100K Election Forecasting Methodology and Survey

B.1 Idealized Methodology

B.1.1 Objective

Our goal is to provide an accurate, representative forecast for the 2024 U.S. presidential election by designing a high-quality survey that captures voter sentiment across key demographic and geographic segments. By addressing known polling biases and using transparent methods for aggregation, this methodology aims to produce a reliable election prediction within a $100K budget.

B.1.2 Sampling Approach

We will aim for a sample size of approximately 5,000 respondents, using a stratified random sampling method. This approach will ensure representation across demographics such as age, gender, race, education, and geographic location. The sample size allows for a margin of error of $\pm 1.4\%$ at a 95% confidence level, ensuring robust estimates for subgroups.

Stratification and Quotas: Each demographic quota will be weighted to reflect its proportion in the general voting population, based on recent U.S. Census data. The primary demographic strata include:

- **Age:** 18-29, 30-44, 45-64, 65+
- **Gender:** Male, Female, Non-binary/Other
- **Race/Ethnicity:** White, Black or African American, Hispanic or Latino, Asian, Other
- **Education Level:** High school or less, Some college, Bachelor's degree, Graduate degree
- **Geographic Region:** All 50 states, with oversampling in key swing states like Pennsylvania, Michigan, and Georgia.

B.1.3 Recruitment Strategy

We will partner with a reputable survey panel provider, such as Ipsos or YouGov, to ensure high-quality, verified respondents. This will help us reach a diverse pool while maintaining reliability and transparency in respondent recruitment. To supplement this sample, we will target social media ads toward underrepresented demographics. Each respondent will receive a $5 incentive to encourage engagement.

B.1.4 Survey Design and Implementation

The survey will be hosted on Google Forms, with a professional introduction, contact information, and clear instructions. It will be kept open for approximately 2-3 weeks, allowing sufficient time for response collection.

Survey Structure:

1. **Introduction and Contact Information**: An introductory message explains the survey's purpose, assures anonymity and provides contact details for further information.
2. **Screening Questions**: To ensure eligible respondents, questions will confirm age, U.S. citizenship, and voter registration status.
3. **Core Questions**: These include demographic information, voting intentions, and issue priorities.
4. **Thank You Message**: The survey concludes with a thank-you note and contact details for any follow-up questions.

Each question will be carefully worded and ordered to maintain neutrality and flow, with an expected completion time of 5-7 minutes.

B.1.5 Data Validation

To maintain data quality, we will implement the following:

- **Attention Checks**: Questions designed to identify inattentive responses.
- **Duplicate Response Prevention**: Google Forms settings will restrict multiple submissions from the same respondent.
- **Post-Survey Weighting**: Adjustments based on demographic quotas ensure alignment with the U.S. population profile.

B.1.6 Poll Aggregation and Forecasting Methodology

For aggregation, each poll will be weighted based on transparency score and demographic representativeness. The forecast will use a Bayesian hierarchical model to account for state-level variation, incorporating:

- **Prior Election Data**: Trends from recent election cycles.
- **Real-Time Updates**: Weekly adjustments based on new data.
- **Transparency Weighting**: Additional weight for respondents who express higher trust in transparent polling practices.

B.2 Idealized Survey

https://docs.google.com/forms/d/1Rr7YOXDPS4yKkB9_AlF4acuMN1swhBV2JFOgpb4PtIg/edit

2024 U.S. Presidential Election Voter Intentions and Key Issues Survey

Welcome to the "2024 U.S. Presidential Election Voter Intentions and Key Issues Survey". This survey aims to gather insights into voter preferences, key concerns, and demographics to better understand the views of the American electorate as we approach the 2024 election. Your responses will remain anonymous and will contribute to a more accurate, representative forecast. The survey takes approximately 5–7 minutes to complete. Thank you for your participation! If you have any questions or concerns, please contact: jiaxuan.song@mail.utoronto.ca

1. Are you a U.S. citizen eligible to vote in the 2024 presidential election?

Yes

No

2. Are you registered to vote?

Yes

No

Prefer not to say

3. Which state do you reside in?

Alabama

Alaska

Arizona

Arkansas

California

Colorado

Connecticut

Delaware

Florida

Georgia

Hawaii

Idaho

Illinois

Indiana

Iowa

Kansas

Kentucky

Louisiana

Maine

Maryland

Massachusetts

Michigan

Minnesota

Mississippi

Missouri

Montana

Nebraska

Nevada

New Hampshire

New Jersey

New Mexico

New York

North Carolina

North Dakota

Ohio

Oklahoma

Oregon

Pennsylvania

Rhode Island

South Carolina

South Dakota

Tennessee

Texas

Utah

Vermont

Virginia

Washington

West Virginia

Wisconsin

Wyoming

4. What is your age?

18-29

30-44

45-64

65+

5. What is your gender?

Male

Female

Non-binary/Other

Prefer not to say

6. What is your race/ethnicity?

White

Black or African American

Hispanic or Latino

Asian

Other

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

High school or less

Some college

Bachelor's degree

Graduate degree

8. On a scale of 1-5, how likely are you to vote in the upcoming election? (Very unlikely to Very likely)

1

2

3

4

5

9. To ensure quality, please select "Somewhat disagree" as your answer to this question.

Strongly agree

Somewhat agree

Neutral

Somewhat disagree

Strongly disagree

10. If the election were held today, which candidate would you most likely support?

Kamala Harris (Democrat)

Donald Trump (Republican)

Other (please specify)

Undecided

11. If you choose "Other" in the previous question, please specify the person you most likely to support.

12. What issues are most important to you in this election? (Select up to 3)

Economy

Healthcare

Climate change

Education

Immigration

Foreign policy

Other (please specify)

13. How satisfied are you with the current direction of the U.S. government? (Very dissatisfied to Very satisfied)

1

2

3

4

5

14. How much do you trust the accuracy of public opinion polls? (Highly distrust to Highly trust)

1

2

3

4

5

15. Have you previously participated in political opinion polls?

Yes

No

Thank You!

Thank you for participating in this survey! Your responses are valuable to our research. If you have any questions or would like more information, please contact: Jiaxuan.song@mail.utoronto.ca

C Model Details

C.1 Summary of Harris model

```
Model Info:
 function:     stan_glm
 family:       gaussian [identity]
 formula:      pct ~ pollster + state + transparency_score + numeric_grade
 algorithm:    sampling
 sample:       8000 (posterior sample size)
 priors:       see help('prior_summary')
 observations: 1075
 predictors:   106

Estimates:
                                                        mean   sd   10%   50%
(Intercept)                                            45.1    1.5 43.2  45.1
pollsterAtlasIntel                                      0.4    0.4 -0.2   0.4
pollsterBeacon/Shaw                                     0.0    0.4 -0.5   0.0
pollsterCES / YouGov                                    1.0    0.6  0.3   1.0
pollsterChristopher Newport U.                          0.3    1.0 -0.9   0.3
pollsterCiviqs                                         -0.5    0.8 -1.5  -0.5
pollsterCNN/SSRS                                       -0.6    0.6 -1.4  -0.6
pollsterData for Progress                               0.1    0.7 -0.8   0.1
pollsterData Orbital                                   -0.8    0.9 -2.0  -0.8
pollsterDHM Research                                   -0.3    0.9 -1.5  -0.3
pollsterEast Carolina University                       -0.6    0.8 -1.6  -0.6
pollsterEchelon Insights                               -0.1    0.7 -0.9  -0.1
pollsterElon U.                                        -0.4    0.9 -1.6  -0.4
pollsterEmerson                                         0.3    0.4 -0.3   0.3
pollsterFairleigh Dickinson                             0.3    0.9 -0.9   0.3
pollsterFranklin and Marshall College                  -0.1    0.8 -1.1  -0.1
pollsterGQR                                             0.1    0.9 -1.1   0.1
pollsterHigh Point University                           0.0    0.9 -1.2   0.0
pollsterInsiderAdvantage                               -0.5    0.6 -1.3  -0.5
pollsterIpsos                                          -1.5    0.4 -2.0  -1.5
```

pollsterKaiser Family Foundation	0.0	1.0	-1.3	0.0
pollsterLandmark Communications	-0.2	0.9	-1.4	-0.2
pollsterLeger	0.5	0.8	-0.5	0.5
pollsterMarist	0.9	0.5	0.3	0.9
pollsterMarquette Law School	-0.2	0.6	-0.9	-0.2
pollsterMason-Dixon	-1.2	0.9	-2.3	-1.2
pollsterMassINC Polling Group	2.9	0.7	2.0	3.0
pollsterMcCourtney Institute/YouGov	0.1	1.0	-1.2	0.1
pollsterMitchell	0.0	0.7	-0.9	0.0
pollsterMonmouth	-0.4	0.8	-1.4	-0.4
pollsterMuhlenberg	0.1	0.9	-1.1	0.1
pollsterNoble Predictive Insights	-1.6	0.6	-2.4	-1.6
pollsterPan Atlantic Research	0.4	0.9	-0.7	0.4
pollsterPew	0.2	0.8	-0.9	0.2
pollsterPPIC	1.0	0.9	-0.2	1.0
pollsterPraecones Analytica	-0.7	0.9	-1.9	-0.7
pollsterQuinnipiac	-0.5	0.5	-1.1	-0.5
pollsterRemington	-0.7	0.8	-1.7	-0.7
pollsterResearch Co.	0.7	0.8	-0.4	0.7
pollsterRMG Research	0.9	0.6	0.2	0.9
pollsterRoanoke College	-0.1	0.9	-1.2	-0.1
pollsterRutgers-Eagleton	0.7	0.9	-0.5	0.7
pollsterSelzer	-0.4	0.9	-1.5	-0.4
pollsterSiena	1.1	0.8	0.0	1.1
pollsterSiena/NYT	-2.3	0.4	-2.8	-2.3
pollsterSt. Anselm	1.2	0.8	0.1	1.2
pollsterSt. Pete Polls	-0.1	1.0	-1.3	0.0
pollsterSuffolk	-0.5	0.7	-1.3	-0.5
pollsterSurveyMonkey	-1.9	0.9	-3.0	-1.9
pollsterSurveyUSA	-0.5	0.6	-1.3	-0.5
pollsterSurveyUSA/High Point University	-0.5	0.9	-1.6	-0.5
pollsterSusquehanna	0.4	0.8	-0.6	0.4
pollsterThe Washington Post	-0.6	0.7	-1.6	-0.6
pollsterU. Georgia SPIA	-0.6	0.9	-1.7	-0.6
pollsterU. Houston	-0.7	1.0	-1.9	-0.7
pollsterU. New Hampshire	4.1	0.6	3.3	4.1
pollsterU. North Florida	-0.6	0.9	-1.8	-0.6
pollsterUC Berkeley	1.0	0.9	-0.1	1.0
pollsterUMass Amherst/YouGov	0.2	0.9	-0.9	0.2
pollsterUniversity of Houston/Texas Southern University	-0.1	0.9	-1.3	-0.2
pollsterUniversity of Maryland/Washington Post	3.1	0.9	2.0	3.1
pollsterUniversity of Massachusetts Lowell/YouGov	-0.3	0.7	-1.2	-0.3
pollsterVirginia Commonwealth U.	-0.7	0.9	-1.8	-0.7

pollsterWashington Post/George Mason University	0.0	0.6	-0.8	0.0
pollsterWinthrop U.	-0.5	1.0	-1.7	-0.5
pollsterYouGov	-0.5	0.4	-1.0	-0.5
pollsterYouGov Blue	-0.2	1.0	-1.5	-0.2
stateCalifornia	5.3	0.7	4.3	5.2
stateConnecticut	0.2	1.0	-1.0	0.3
stateFlorida	-2.5	0.6	-3.2	-2.5
stateGeorgia	-0.6	0.4	-1.1	-0.6
stateIndiana	-0.7	1.0	-1.9	-0.7
stateIowa	-0.6	0.9	-1.8	-0.6
stateMaine	0.6	0.8	-0.4	0.6
stateMaine CD-1	3.7	0.8	2.7	3.7
stateMaine CD-2	-2.4	0.8	-3.4	-2.4
stateMaryland	5.5	0.8	4.5	5.5
stateMassachusetts	4.8	0.8	3.8	4.8
stateMichigan	0.3	0.4	-0.1	0.3
stateMinnesota	1.1	0.7	0.1	1.1
stateMissouri	-1.6	0.8	-2.7	-1.6
stateMontana	-4.0	0.7	-4.9	-4.0
stateNational	-0.3	0.3	-0.6	-0.3
stateNebraska	-3.2	0.8	-4.2	-3.2
stateNebraska CD-1	-0.4	0.9	-1.6	-0.4
stateNebraska CD-2	2.4	0.7	1.5	2.4
stateNebraska CD-3	-2.2	0.9	-3.4	-2.2
stateNevada	0.2	0.4	-0.4	0.2
stateNew Hampshire	0.6	0.6	-0.2	0.6
stateNew Jersey	0.7	0.9	-0.5	0.7
stateNew Mexico	1.0	0.8	0.0	1.0
stateNew York	3.1	0.8	2.1	3.1
stateNorth Carolina	0.2	0.4	-0.3	0.2
stateOhio	-2.1	0.7	-2.9	-2.1
statePennsylvania	0.3	0.3	-0.1	0.3
stateRhode Island	1.1	0.9	0.0	1.1
stateSouth Carolina	-0.9	0.9	-2.1	-1.0
stateSouth Dakota	-2.2	0.9	-3.4	-2.2
stateTexas	-2.3	0.6	-3.0	-2.3
stateUtah	-3.1	0.8	-4.2	-3.1
stateVermont	3.5	0.9	2.3	3.4
stateVirginia	0.4	0.6	-0.3	0.4
stateWashington	0.9	0.9	-0.2	0.9
stateWisconsin	1.0	0.4	0.5	1.0
transparency_score	0.3	0.1	0.1	0.3
numeric_grade	0.0	0.6	-0.8	0.0

sigma	3.0	0.1	2.9	3.0

	90%
(Intercept)	47.0
pollsterAtlasIntel	1.0
pollsterBeacon/Shaw	0.6
pollsterCES / YouGov	1.7
pollsterChristopher Newport U.	1.6
pollsterCiviqs	0.6
pollsterCNN/SSRS	0.2
pollsterData for Progress	1.0
pollsterData Orbital	0.4
pollsterDHM Research	0.9
pollsterEast Carolina University	0.5
pollsterEchelon Insights	0.7
pollsterElon U.	0.8
pollsterEmerson	0.8
pollsterFairleigh Dickinson	1.5
pollsterFranklin and Marshall College	1.0
pollsterGQR	1.3
pollsterHigh Point University	1.1
pollsterInsiderAdvantage	0.2
pollsterIpsos	-0.9
pollsterKaiser Family Foundation	1.2
pollsterLandmark Communications	1.0
pollsterLeger	1.6
pollsterMarist	1.5
pollsterMarquette Law School	0.6
pollsterMason-Dixon	0.0
pollsterMassINC Polling Group	3.9
pollsterMcCourtney Institute/YouGov	1.3
pollsterMitchell	0.9
pollsterMonmouth	0.7
pollsterMuhlenberg	1.2
pollsterNoble Predictive Insights	-0.9
pollsterPan Atlantic Research	1.5
pollsterPew	1.3
pollsterPPIC	2.2
pollsterPraecones Analytica	0.4
pollsterQuinnipiac	0.1
pollsterRemington	0.3
pollsterResearch Co.	1.8
pollsterRMG Research	1.6
pollsterRoanoke College	1.1

pollsterRutgers-Eagleton	1.9
pollsterSelzer	0.8
pollsterSiena	2.2
pollsterSiena/NYT	-1.9
pollsterSt. Anselm	2.2
pollsterSt. Pete Polls	1.2
pollsterSuffolk	0.4
pollsterSurveyMonkey	-0.8
pollsterSurveyUSA	0.3
pollsterSurveyUSA/High Point University	0.6
pollsterSusquehanna	1.4
pollsterThe Washington Post	0.3
pollsterU. Georgia SPIA	0.5
pollsterU. Houston	0.6
pollsterU. New Hampshire	4.8
pollsterU. North Florida	0.6
pollsterUC Berkeley	2.2
pollsterUMass Amherst/YouGov	1.3
pollsterUniversity of Houston/Texas Southern University	1.1
pollsterUniversity of Maryland/Washington Post	4.2
pollsterUniversity of Massachusetts Lowell/YouGov	0.7
pollsterVirginia Commonwealth U.	0.4
pollsterWashington Post/George Mason University	0.7
pollsterWinthrop U.	0.7
pollsterYouGov	0.0
pollsterYouGov Blue	1.0
stateCalifornia	6.2
stateConnecticut	1.5
stateFlorida	-1.8
stateGeorgia	-0.1
stateIndiana	0.5
stateIowa	0.5
stateMaine	1.6
stateMaine CD-1	4.7
stateMaine CD-2	-1.4
stateMaryland	6.5
stateMassachusetts	5.8
stateMichigan	0.8
stateMinnesota	2.0
stateMissouri	-0.5
stateMontana	-3.1
stateNational	0.1
stateNebraska	-2.2

```
stateNebraska CD-1                                         0.8
stateNebraska CD-2                                         3.4
stateNebraska CD-3                                        -1.0
stateNevada                                                0.7
stateNew Hampshire                                         1.4
stateNew Jersey                                            1.8
stateNew Mexico                                            2.1
stateNew York                                              4.1
stateNorth Carolina                                        0.7
stateOhio                                                 -1.2
statePennsylvania                                          0.7
stateRhode Island                                          2.2
stateSouth Carolina                                        0.2
stateSouth Dakota                                         -1.0
stateTexas                                                -1.6
stateUtah                                                 -2.0
stateVermont                                               4.6
stateVirginia                                              1.2
stateWashington                                            2.1
stateWisconsin                                             1.5
transparency_score                                         0.5
numeric_grade                                              0.7
sigma                                                      3.1

Fit Diagnostics:
           mean   sd   10%   50%   90%
mean_PPD 47.3    0.1 47.2  47.3  47.5

The mean_ppd is the sample average posterior predictive distribution of the outcome variable

MCMC diagnostics
                                                          mcse Rhat n_eff
(Intercept)                                               0.0  1.0   4470
pollsterAtlasIntel                                        0.0  1.0   4201
pollsterBeacon/Shaw                                       0.0  1.0   5447
pollsterCES / YouGov                                      0.0  1.0   6064
pollsterChristopher Newport U.                            0.0  1.0  10565
pollsterCiviqs                                            0.0  1.0   9457
pollsterCNN/SSRS                                          0.0  1.0   8513
pollsterData for Progress                                 0.0  1.0   9643
pollsterData Orbital                                      0.0  1.0  11525
pollsterDHM Research                                      0.0  1.0  11364
pollsterEast Carolina University                          0.0  1.0  11539
```

pollsterEchelon Insights	0.0	1.0	9572
pollsterElon U.	0.0	1.0	10533
pollsterEmerson	0.0	1.0	4444
pollsterFairleigh Dickinson	0.0	1.0	10900
pollsterFranklin and Marshall College	0.0	1.0	9146
pollsterGQR	0.0	1.0	9556
pollsterHigh Point University	0.0	1.0	10831
pollsterInsiderAdvantage	0.0	1.0	7060
pollsterIpsos	0.0	1.0	5924
pollsterKaiser Family Foundation	0.0	1.0	11244
pollsterLandmark Communications	0.0	1.0	9871
pollsterLeger	0.0	1.0	9955
pollsterMarist	0.0	1.0	7156
pollsterMarquette Law School	0.0	1.0	6721
pollsterMason-Dixon	0.0	1.0	11326
pollsterMassINC Polling Group	0.0	1.0	8520
pollsterMcCourtney Institute/YouGov	0.0	1.0	9591
pollsterMitchell	0.0	1.0	9658
pollsterMonmouth	0.0	1.0	10452
pollsterMuhlenberg	0.0	1.0	11306
pollsterNoble Predictive Insights	0.0	1.0	7362
pollsterPan Atlantic Research	0.0	1.0	10379
pollsterPew	0.0	1.0	11587
pollsterPPIC	0.0	1.0	11150
pollsterPraecones Analytica	0.0	1.0	9308
pollsterQuinnipiac	0.0	1.0	7265
pollsterRemington	0.0	1.0	10785
pollsterResearch Co.	0.0	1.0	10360
pollsterRMG Research	0.0	1.0	6248
pollsterRoanoke College	0.0	1.0	10268
pollsterRutgers-Eagleton	0.0	1.0	10737
pollsterSelzer	0.0	1.0	10067
pollsterSiena	0.0	1.0	10143
pollsterSiena/NYT	0.0	1.0	4021
pollsterSt. Anselm	0.0	1.0	9950
pollsterSt. Pete Polls	0.0	1.0	10422
pollsterSuffolk	0.0	1.0	8854
pollsterSurveyMonkey	0.0	1.0	10773
pollsterSurveyUSA	0.0	1.0	8133
pollsterSurveyUSA/High Point University	0.0	1.0	11048
pollsterSusquehanna	0.0	1.0	10118
pollsterThe Washington Post	0.0	1.0	9353
pollsterU. Georgia SPIA	0.0	1.0	9849

pollsterU. Houston	0.0	1.0	11698
pollsterU. New Hampshire	0.0	1.0	6402
pollsterU. North Florida	0.0	1.0	10421
pollsterUC Berkeley	0.0	1.0	10451
pollsterUMass Amherst/YouGov	0.0	1.0	10884
pollsterUniversity of Houston/Texas Southern University	0.0	1.0	10762
pollsterUniversity of Maryland/Washington Post	0.0	1.0	10248
pollsterUniversity of Massachusetts Lowell/YouGov	0.0	1.0	9473
pollsterVirginia Commonwealth U.	0.0	1.0	10596
pollsterWashington Post/George Mason University	0.0	1.0	8681
pollsterWinthrop U.	0.0	1.0	11033
pollsterYouGov	0.0	1.0	4298
pollsterYouGov Blue	0.0	1.0	11195
stateCalifornia	0.0	1.0	9691
stateConnecticut	0.0	1.0	9785
stateFlorida	0.0	1.0	9456
stateGeorgia	0.0	1.0	6928
stateIndiana	0.0	1.0	11489
stateIowa	0.0	1.0	10582
stateMaine	0.0	1.0	11867
stateMaine CD-1	0.0	1.0	9463
stateMaine CD-2	0.0	1.0	10328
stateMaryland	0.0	1.0	9288
stateMassachusetts	0.0	1.0	8925
stateMichigan	0.0	1.0	6606
stateMinnesota	0.0	1.0	9381
stateMissouri	0.0	1.0	10140
stateMontana	0.0	1.0	10301
stateNational	0.0	1.0	4894
stateNebraska	0.0	1.0	9419
stateNebraska CD-1	0.0	1.0	11333
stateNebraska CD-2	0.0	1.0	10001
stateNebraska CD-3	0.0	1.0	11476
stateNevada	0.0	1.0	7317
stateNew Hampshire	0.0	1.0	7499
stateNew Jersey	0.0	1.0	10139
stateNew Mexico	0.0	1.0	10617
stateNew York	0.0	1.0	8167
stateNorth Carolina	0.0	1.0	6274
stateOhio	0.0	1.0	9566
statePennsylvania	0.0	1.0	5515
stateRhode Island	0.0	1.0	10706
stateSouth Carolina	0.0	1.0	10738

```
stateSouth Dakota                                          0.0  1.0  11339
stateTexas                                                 0.0  1.0   8717
stateUtah                                                  0.0  1.0   9431
stateVermont                                               0.0  1.0  11089
stateVirginia                                              0.0  1.0   8222
stateWashington                                            0.0  1.0  10088
stateWisconsin                                             0.0  1.0   6078
transparency_score                                         0.0  1.0   3422
numeric_grade                                              0.0  1.0   4316
sigma                                                      0.0  1.0   7410
mean_PPD                                                   0.0  1.0   9099
log-posterior                                              0.1  1.0   3130

For each parameter, mcse is Monte Carlo standard error, n_eff is a crude measure of effectiv
```

The model summary for Kamala Harris displays the estimates of coefficients for various predictors, including pollster, state, transparency score, and numeric grade. It provides insight into how each factor contributes to predicting support for Harris in the 2024 election.

C.2 Summary of Trump model

```
Model Info:
 function:     stan_glm
 family:       gaussian [identity]
 formula:      pct ~ pollster + state + transparency_score + numeric_grade
 algorithm:    sampling
 sample:       8000 (posterior sample size)
 priors:       see help('prior_summary')
 observations: 2051
 predictors:   125

Estimates:
                                                          mean   sd   10%   50%
(Intercept)                                              42.0    1.4 40.2  42.0
pollsterAngus Reid                                       -0.4    0.8 -1.5  -0.4
pollsterAtlasIntel                                        2.7    0.5  2.2   2.7
pollsterBeacon/Shaw                                       1.5    0.4  1.0   1.5
pollsterCES / YouGov                                      1.5    0.6  0.7   1.5
pollsterChristopher Newport U.                           -0.1    1.0 -1.3  -0.1
pollsterCiviqs                                            0.6    0.8 -0.3   0.6
pollsterCNN/SSRS                                          0.9    0.6  0.2   0.9
pollsterData for Progress                                 1.8    0.6  1.0   1.8
pollsterData Orbital                                     -0.3    0.9 -1.5  -0.3
pollsterDHM Research                                     -0.8    0.9 -2.0  -0.8
pollsterEast Carolina University                          0.8    0.8 -0.3   0.8
pollsterEchelon Insights                                  1.6    0.5  1.0   1.6
pollsterElon U.                                          -0.1    0.9 -1.3  -0.1
pollsterEmerson                                           0.5    0.3  0.1   0.5
pollsterFairleigh Dickinson                               0.2    0.9 -1.0   0.2
pollsterFlorida Atlantic University                      -0.2    1.0 -1.4  -0.2
pollsterFranklin and Marshall College                    -1.0    0.7 -1.9  -1.0
pollsterGQR                                               0.8    0.9 -0.4   0.8
pollsterHigh Point University                            -0.4    0.8 -1.5  -0.4
pollsterInnovative Research Group                        -0.1    1.0 -1.3  -0.1
pollsterInsiderAdvantage                                  2.1    0.6  1.3   2.1
pollsterIpsos                                            -2.5    0.4 -3.0  -2.5
```

pollsterKaiser Family Foundation	0.0	0.9	-1.1	-0.1
pollsterLandmark Communications	0.1	0.9	-1.1	0.1
pollsterLeger	-0.3	0.7	-1.2	-0.3
pollsterMarist	1.5	0.4	0.9	1.5
pollsterMarquette Law School	0.8	0.5	0.1	0.8
pollsterMason-Dixon	0.3	0.8	-0.8	0.3
pollsterMassINC Polling Group	-1.5	0.8	-2.5	-1.5
pollsterMcCourtney Institute/YouGov	0.0	1.0	-1.2	0.0
pollsterMitchell	1.4	0.7	0.5	1.3
pollsterMonmouth	0.0	0.9	-1.1	0.0
pollsterMuhlenberg	-1.2	0.8	-2.2	-1.2
pollsterNoble Predictive Insights	0.5	0.5	-0.2	0.5
pollsterNORC	-0.5	1.0	-1.7	-0.5
pollsterPan Atlantic Research	-1.2	0.8	-2.3	-1.2
pollsterPew	1.3	0.8	0.3	1.3
pollsterPPIC	-2.0	0.8	-3.1	-2.0
pollsterPraecones Analytica	-0.1	0.9	-1.2	-0.1
pollsterQuinnipiac	0.2	0.4	-0.3	0.2
pollsterRasmussen	0.5	0.8	-0.5	0.5
pollsterRemington	1.5	0.7	0.6	1.5
pollsterResearch Co.	0.6	0.9	-0.5	0.6
pollsterRMG Research	1.7	0.6	1.0	1.7
pollsterRoanoke College	-0.2	0.8	-1.3	-0.2
pollsterRutgers-Eagleton	-0.7	0.9	-2.0	-0.7
pollsterSelzer	0.1	0.8	-1.0	0.1
pollsterSiena	-2.4	0.7	-3.4	-2.4
pollsterSiena/NYT	0.0	0.4	-0.4	0.0
pollsterSSRS	-0.5	0.9	-1.7	-0.5
pollsterSt. Anselm	0.0	0.8	-1.1	0.0
pollsterSt. Pete Polls	0.1	0.9	-1.1	0.1
pollsterSuffolk	-2.4	0.6	-3.2	-2.4
pollsterSurveyMonkey	-1.1	0.9	-2.3	-1.1
pollsterSurveyUSA	-0.1	0.5	-0.7	-0.1
pollsterSurveyUSA/High Point University	0.0	0.9	-1.2	0.0
pollsterSusquehanna	-0.5	0.7	-1.4	-0.5
pollsterThe Washington Post	0.9	0.8	-0.1	0.9
pollsterU. Georgia SPIA	0.2	0.8	-0.8	0.2
pollsterU. Houston	0.5	0.9	-0.7	0.5
pollsterU. Massachusetts - Lowell	-0.6	0.9	-1.8	-0.6
pollsterU. New Hampshire	-2.0	0.6	-2.8	-2.0
pollsterU. North Florida	0.6	0.9	-0.6	0.6
pollsterUC Berkeley	-1.5	0.8	-2.5	-1.5
pollsterUMass Amherst/YouGov	-1.6	0.8	-2.7	-1.6

pollsterUniversity of Houston/Texas Southern University	0.4	0.9	-0.8	0.4
pollsterUniversity of Maryland/Washington Post	-2.0	0.9	-3.1	-2.0
pollsterUniversity of Massachusetts Lowell/YouGov	-0.6	0.7	-1.6	-0.6
pollsterVirginia Commonwealth U.	-1.2	0.8	-2.3	-1.2
pollsterWashington Post/George Mason University	0.8	0.7	-0.1	0.8
pollsterWinthrop U.	0.7	1.0	-0.6	0.7
pollsterYouGov	0.3	0.4	-0.2	0.3
pollsterYouGov Blue	-0.2	0.9	-1.3	-0.2
stateArizona	1.9	0.4	1.4	1.9
stateArkansas	0.8	1.0	-0.4	0.8
stateCalifornia	-7.7	0.6	-8.5	-7.7
stateColorado	-2.3	0.8	-3.4	-2.3
stateConnecticut	-0.8	0.9	-2.0	-0.8
stateFlorida	3.6	0.6	2.8	3.6
stateGeorgia	1.8	0.4	1.3	1.8
stateIdaho	0.6	1.0	-0.6	0.7
stateIllinois	-1.8	0.9	-3.0	-1.8
stateIndiana	1.4	0.9	0.2	1.4
stateIowa	0.8	0.8	-0.2	0.8
stateKansas	0.7	0.9	-0.5	0.7
stateMaine	-0.8	0.8	-1.8	-0.8
stateMaine CD-1	-2.8	0.8	-3.9	-2.8
stateMaine CD-2	1.8	0.8	0.7	1.8
stateMaryland	-5.4	0.8	-6.4	-5.4
stateMassachusetts	-7.5	0.7	-8.4	-7.5
stateMichigan	0.5	0.4	0.1	0.5
stateMinnesota	-1.8	0.7	-2.6	-1.8
stateMissouri	3.5	0.7	2.6	3.5
stateMontana	5.0	0.7	4.1	5.0
stateNational	-0.5	0.3	-0.9	-0.5
stateNebraska	3.2	0.8	2.2	3.2
stateNebraska CD-1	0.4	1.0	-0.8	0.4
stateNebraska CD-2	-1.3	0.8	-2.3	-1.3
stateNebraska CD-3	1.8	1.0	0.5	1.8
stateNevada	1.0	0.4	0.5	1.0
stateNew Hampshire	-1.3	0.5	-2.0	-1.3
stateNew Jersey	-1.7	0.9	-2.9	-1.7
stateNew Mexico	-1.2	0.8	-2.2	-1.2
stateNew York	-4.2	0.7	-5.0	-4.2
stateNorth Carolina	1.8	0.4	1.3	1.8
stateNorth Dakota	0.6	1.0	-0.6	0.6
stateOhio	2.9	0.6	2.1	2.9
stateOklahoma	0.9	1.0	-0.3	0.9

stateOregon	-1.1	0.9	-2.3	-1.1
statePennsylvania	0.9	0.3	0.5	0.9
stateRhode Island	-1.1	0.8	-2.2	-1.1
stateSouth Carolina	1.5	0.9	0.4	1.5
stateSouth Dakota	2.9	0.8	1.8	2.9
stateTennessee	1.3	0.9	0.2	1.3
stateTexas	2.9	0.5	2.2	2.9
stateUtah	2.4	0.8	1.3	2.3
stateVermont	-2.4	0.9	-3.6	-2.4
stateVirginia	-2.3	0.6	-3.0	-2.3
stateWashington	-3.0	0.8	-4.0	-3.0
stateWest Virginia	1.0	0.9	-0.2	1.0
stateWisconsin	0.6	0.4	0.1	0.6
stateWyoming	1.5	1.0	0.3	1.5
transparency_score	-0.3	0.1	-0.4	-0.3
numeric_grade	1.8	0.5	1.1	1.8
sigma	3.6	0.1	3.5	3.6

	90%
(Intercept)	43.8
pollsterAngus Reid	0.6
pollsterAtlasIntel	3.3
pollsterBeacon/Shaw	1.9
pollsterCES / YouGov	2.3
pollsterChristopher Newport U.	1.1
pollsterCiviqs	1.6
pollsterCNN/SSRS	1.7
pollsterData for Progress	2.6
pollsterData Orbital	0.9
pollsterDHM Research	0.3
pollsterEast Carolina University	1.8
pollsterEchelon Insights	2.2
pollsterElon U.	1.1
pollsterEmerson	1.0
pollsterFairleigh Dickinson	1.4
pollsterFlorida Atlantic University	1.1
pollsterFranklin and Marshall College	0.0
pollsterGQR	1.9
pollsterHigh Point University	0.7
pollsterInnovative Research Group	1.2
pollsterInsiderAdvantage	2.8
pollsterIpsos	-2.0
pollsterKaiser Family Foundation	1.0
pollsterLandmark Communications	1.3

pollsterLeger	0.5
pollsterMarist	2.0
pollsterMarquette Law School	1.5
pollsterMason-Dixon	1.3
pollsterMassINC Polling Group	-0.5
pollsterMcCourtney Institute/YouGov	1.2
pollsterMitchell	2.3
pollsterMonmouth	1.1
pollsterMuhlenberg	-0.1
pollsterNoble Predictive Insights	1.2
pollsterNORC	0.8
pollsterPan Atlantic Research	-0.2
pollsterPew	2.3
pollsterPPIC	-1.0
pollsterPraecones Analytica	1.1
pollsterQuinnipiac	0.8
pollsterRasmussen	1.4
pollsterRemington	2.3
pollsterResearch Co.	1.7
pollsterRMG Research	2.5
pollsterRoanoke College	0.9
pollsterRutgers-Eagleton	0.5
pollsterSelzer	1.2
pollsterSiena	-1.4
pollsterSiena/NYT	0.5
pollsterSSRS	0.7
pollsterSt. Anselm	1.0
pollsterSt. Pete Polls	1.3
pollsterSuffolk	-1.6
pollsterSurveyMonkey	0.0
pollsterSurveyUSA	0.6
pollsterSurveyUSA/High Point University	1.1
pollsterSusquehanna	0.4
pollsterThe Washington Post	1.9
pollsterU. Georgia SPIA	1.2
pollsterU. Houston	1.7
pollsterU. Massachusetts - Lowell	0.5
pollsterU. New Hampshire	-1.2
pollsterU. North Florida	1.8
pollsterUC Berkeley	-0.4
pollsterUMass Amherst/YouGov	-0.5
pollsterUniversity of Houston/Texas Southern University	1.6
pollsterUniversity of Maryland/Washington Post	-0.8

pollsterUniversity of Massachusetts Lowell/YouGov	0.3
pollsterVirginia Commonwealth U.	-0.1
pollsterWashington Post/George Mason University	1.6
pollsterWinthrop U.	1.9
pollsterYouGov	0.8
pollsterYouGov Blue	1.0
stateArizona	2.3
stateArkansas	2.0
stateCalifornia	-6.9
stateColorado	-1.2
stateConnecticut	0.4
stateFlorida	4.3
stateGeorgia	2.3
stateIdaho	1.9
stateIllinois	-0.6
stateIndiana	2.6
stateIowa	1.9
stateKansas	1.8
stateMaine	0.2
stateMaine CD-1	-1.8
stateMaine CD-2	2.8
stateMaryland	-4.4
stateMassachusetts	-6.6
stateMichigan	1.0
stateMinnesota	-0.9
stateMissouri	4.4
stateMontana	5.9
stateNational	-0.2
stateNebraska	4.2
stateNebraska CD-1	1.7
stateNebraska CD-2	-0.3
stateNebraska CD-3	3.0
stateNevada	1.6
stateNew Hampshire	-0.6
stateNew Jersey	-0.6
stateNew Mexico	-0.1
stateNew York	-3.3
stateNorth Carolina	2.3
stateNorth Dakota	1.8
stateOhio	3.6
stateOklahoma	2.1
stateOregon	0.1
statePennsylvania	1.4

```
stateRhode Island                                          0.0
stateSouth Carolina                                        2.7
stateSouth Dakota                                          3.9
stateTennessee                                             2.4
stateTexas                                                 3.5
stateUtah                                                  3.4
stateVermont                                              -1.2
stateVirginia                                             -1.5
stateWashington                                           -2.0
stateWest Virginia                                         2.2
stateWisconsin                                             1.1
stateWyoming                                               2.8
transparency_score                                        -0.2
numeric_grade                                              2.5
sigma                                                      3.7

Fit Diagnostics:
           mean   sd   10%   50%   90%
mean_PPD 45.0    0.1 44.8  45.0  45.1

The mean_ppd is the sample average posterior predictive distribution of the outcome variable

MCMC diagnostics
                                                          mcse Rhat n_eff
(Intercept)                                               0.0  1.0   3988
pollsterAngus Reid                                        0.0  1.0  11952
pollsterAtlasIntel                                        0.0  1.0   5703
pollsterBeacon/Shaw                                       0.0  1.0   5844
pollsterCES / YouGov                                      0.0  1.0   8349
pollsterChristopher Newport U.                            0.0  1.0  12493
pollsterCiviqs                                            0.0  1.0  11274
pollsterCNN/SSRS                                          0.0  1.0   8630
pollsterData for Progress                                 0.0  1.0  10129
pollsterData Orbital                                      0.0  1.0  11969
pollsterDHM Research                                      0.0  1.0  10304
pollsterEast Carolina University                          0.0  1.0  12016
pollsterEchelon Insights                                  0.0  1.0   7983
pollsterElon U.                                           0.0  1.0  11885
pollsterEmerson                                           0.0  1.0   3651
pollsterFairleigh Dickinson                               0.0  1.0  11121
pollsterFlorida Atlantic University                       0.0  1.0  10931
pollsterFranklin and Marshall College                     0.0  1.0  10613
pollsterGQR                                               0.0  1.0  12414
```

pollsterHigh Point University	0.0	1.0	9543
pollsterInnovative Research Group	0.0	1.0	12737
pollsterInsiderAdvantage	0.0	1.0	7163
pollsterIpsos	0.0	1.0	5435
pollsterKaiser Family Foundation	0.0	1.0	10929
pollsterLandmark Communications	0.0	1.0	11384
pollsterLeger	0.0	1.0	8708
pollsterMarist	0.0	1.0	5935
pollsterMarquette Law School	0.0	1.0	5500
pollsterMason-Dixon	0.0	1.0	10170
pollsterMassINC Polling Group	0.0	1.0	9982
pollsterMcCourtney Institute/YouGov	0.0	1.0	11833
pollsterMitchell	0.0	1.0	11126
pollsterMonmouth	0.0	1.0	10793
pollsterMuhlenberg	0.0	1.0	12095
pollsterNoble Predictive Insights	0.0	1.0	7897
pollsterNORC	0.0	1.0	11272
pollsterPan Atlantic Research	0.0	1.0	12114
pollsterPew	0.0	1.0	12618
pollsterPPIC	0.0	1.0	10299
pollsterPraecones Analytica	0.0	1.0	10342
pollsterQuinnipiac	0.0	1.0	6898
pollsterRasmussen	0.0	1.0	10242
pollsterRemington	0.0	1.0	9978
pollsterResearch Co.	0.0	1.0	10521
pollsterRMG Research	0.0	1.0	7843
pollsterRoanoke College	0.0	1.0	11131
pollsterRutgers-Eagleton	0.0	1.0	10745
pollsterSelzer	0.0	1.0	11511
pollsterSiena	0.0	1.0	8178
pollsterSiena/NYT	0.0	1.0	3810
pollsterSSRS	0.0	1.0	13041
pollsterSt. Anselm	0.0	1.0	10823
pollsterSt. Pete Polls	0.0	1.0	10360
pollsterSuffolk	0.0	1.0	8362
pollsterSurveyMonkey	0.0	1.0	11890
pollsterSurveyUSA	0.0	1.0	7905
pollsterSurveyUSA/High Point University	0.0	1.0	11311
pollsterSusquehanna	0.0	1.0	11570
pollsterThe Washington Post	0.0	1.0	11192
pollsterU. Georgia SPIA	0.0	1.0	11314
pollsterU. Houston	0.0	1.0	10758
pollsterU. Massachusetts - Lowell	0.0	1.0	10666

pollsterU. New Hampshire	0.0	1.0	7242
pollsterU. North Florida	0.0	1.0	11291
pollsterUC Berkeley	0.0	1.0	10528
pollsterUMass Amherst/YouGov	0.0	1.0	10854
pollsterUniversity of Houston/Texas Southern University	0.0	1.0	12374
pollsterUniversity of Maryland/Washington Post	0.0	1.0	10463
pollsterUniversity of Massachusetts Lowell/YouGov	0.0	1.0	10417
pollsterVirginia Commonwealth U.	0.0	1.0	8829
pollsterWashington Post/George Mason University	0.0	1.0	10133
pollsterWinthrop U.	0.0	1.0	12288
pollsterYouGov	0.0	1.0	3794
pollsterYouGov Blue	0.0	1.0	11566
stateArizona	0.0	1.0	5494
stateArkansas	0.0	1.0	11877
stateCalifornia	0.0	1.0	7266
stateColorado	0.0	1.0	9540
stateConnecticut	0.0	1.0	11516
stateFlorida	0.0	1.0	8927
stateGeorgia	0.0	1.0	5132
stateIdaho	0.0	1.0	10814
stateIllinois	0.0	1.0	10175
stateIndiana	0.0	1.0	12305
stateIowa	0.0	1.0	11701
stateKansas	0.0	1.0	10395
stateMaine	0.0	1.0	11261
stateMaine CD-1	0.0	1.0	10715
stateMaine CD-2	0.0	1.0	9973
stateMaryland	0.0	1.0	9864
stateMassachusetts	0.0	1.0	9936
stateMichigan	0.0	1.0	5214
stateMinnesota	0.0	1.0	8540
stateMissouri	0.0	1.0	9551
stateMontana	0.0	1.0	11015
stateNational	0.0	1.0	3353
stateNebraska	0.0	1.0	11188
stateNebraska CD-1	0.0	1.0	11839
stateNebraska CD-2	0.0	1.0	11531
stateNebraska CD-3	0.0	1.0	11095
stateNevada	0.0	1.0	6520
stateNew Hampshire	0.0	1.0	8162
stateNew Jersey	0.0	1.0	10040
stateNew Mexico	0.0	1.0	11823
stateNew York	0.0	1.0	8663

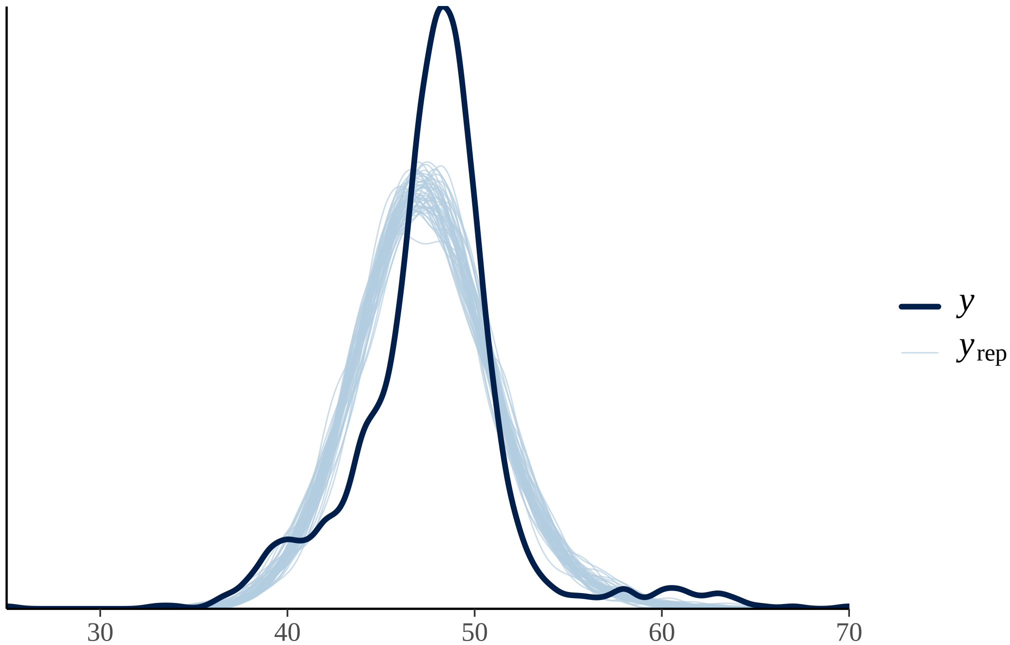
```
stateNorth Carolina                                        0.0  1.0   5577
stateNorth Dakota                                          0.0  1.0  10780
stateOhio                                                  0.0  1.0   7111
stateOklahoma                                              0.0  1.0  11615
stateOregon                                                0.0  1.0  11030
statePennsylvania                                          0.0  1.0   4667
stateRhode Island                                          0.0  1.0  11215
stateSouth Carolina                                        0.0  1.0  11794
stateSouth Dakota                                          0.0  1.0  10630
stateTennessee                                             0.0  1.0  11959
stateTexas                                                 0.0  1.0   7778
stateUtah                                                  0.0  1.0  11248
stateVermont                                               0.0  1.0  11238
stateVirginia                                              0.0  1.0   7953
stateWashington                                            0.0  1.0  10371
stateWest Virginia                                         0.0  1.0  10696
stateWisconsin                                             0.0  1.0   4554
stateWyoming                                               0.0  1.0  12024
transparency_score                                         0.0  1.0   4362
numeric_grade                                              0.0  1.0   3792
sigma                                                      0.0  1.0   8683
mean_PPD                                                   0.0  1.0   9173
log-posterior                                              0.1  1.0   3288

For each parameter, mcse is Monte Carlo standard error, n_eff is a crude measure of effectiv
```

The model summary for Donald Trump displays the estimates of coefficients for various predictors, including pollster, state, transparency score, and numeric grade. It provides insight into how each factor contributes to predicting support for Harris in the 2024 election.

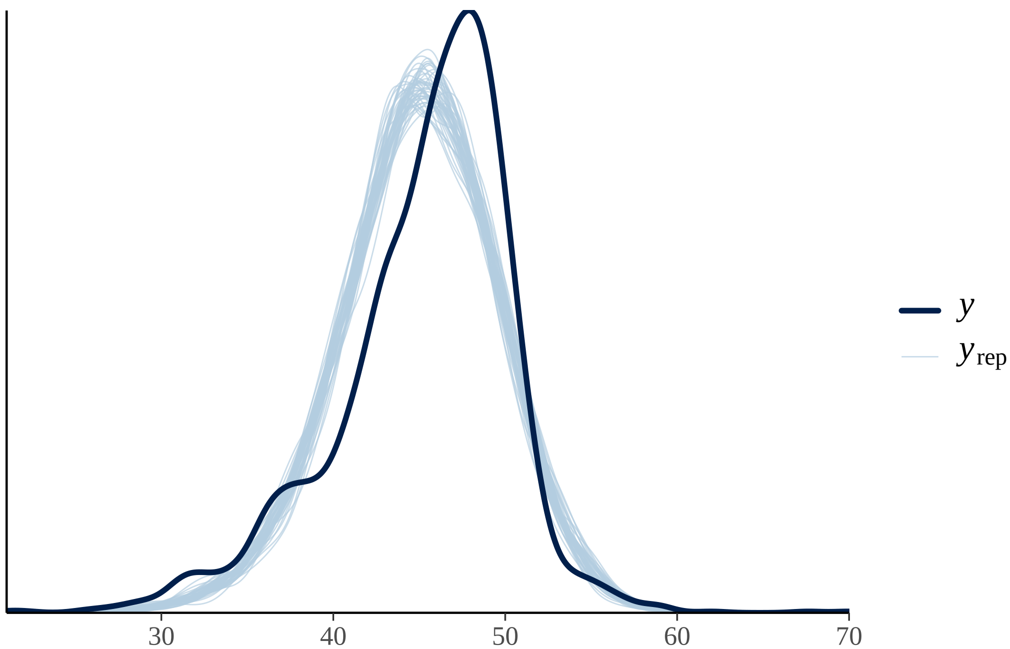
C.3 Posterior predictive checks

C.3.1 Posterior predictive checks for Harris model



The posterior predictive check (pp_check) for Kamala Harris's model visually compares the model-generated predictions to the actual observed values, assessing the model's fit and reliability.

C.3.2 Posterior predictive checks for Trump model



The posterior predictive check (pp_check) for Donald Trump's model visually compares the model-generated predictions to the actual observed values, assessing the model's fit and reliability.

References

Arel-Bundock, Vincent. 2023. *Modelsummary: Beautiful and Customizable Model Summaries in r*. https://vincentarelbundock.github.io/modelsummary/.

Bartels, Larry M. 2000. "Partisanship and Voting Behavior, 1952-1996." *American Journal of Political Science* 44 (1): 35–50. https://doi.org/10.2307/2669291.

Board, The New York Times Editorial. 2024. "Election Polls and Results: Trump Vs. Harris." *The New York Times*. https://www.nytimes.com/2024/10/23/opinion/election-polls-results-trump-harris.html?unlocked_article_code=1.UU4.pFkQ.F2hD-woxmiEj&smid=url-share.

Conant, James K. 2006. *Wisconsin Politics and Government: America's Laboratory of Democracy*. U of Nebraska Press.

Enns, Peter K, Jonathan Colner, Anusha Kumar, and Julius Lagodny. 2024. "Understanding Biden's Exit and the 2024 Election: The State Presidential Approval/State Economy Model." *PS: Political Science & Politics*, 1–14.

Etikan, Ilker, Sulaiman Abubakar Musa, and Rukayya Sunusi Alkassim. 2016. "Comparison of Convenience Sampling and Purposive Sampling." *American Journal of Theoretical and Applied Statistics* 5 (1): 1–4. https://doi.org/10.11648/j.ajtas.20160501.11.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://sfirke.github.io/janitor/.

FiveThirtyEight. 2024. "2024 Election Polls." FiveThirtyEight. https://projects.fivethirtyeight.com/polls/.

Gelman, Andrew, and Gary King. 1993. "Why Are American Presidential Election Campaign Polls so Variable When Votes Are so Predictable?" *British Journal of Political Science* 23 (4): 409–51.

Horst, Allison, Alison Hill, Kristen B Gorman, and Raymond Barbour. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. https://allisonhorst.github.io/palmerpenguins/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, and Apache Arrow Developers. 2023. *Arrow: Interoperable Data Frames for r and Python with Apache Arrow*. https://arrow.apache.org/.

Scherpenzeel, Annette. 2011. "Data Collection in a Probability-Based Internet Panel: How the LISS Panel Was Built and How It Can Be Used." *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 109 (1): 56–61.

School, Marquette University Law. 2024. "About the Poll." https://law.marquette.edu/poll/about-the-poll/.

Walker, David A. 2023. "Forecasting the 2020 and 2024 US Presidential Elections." *Journal of Forecasting* 42 (6): 1519–35.

Warren, John Robert, and Andrew Halpern-Manners. 2012. "Panel Conditioning in Longitudinal Social Science Surveys." *Social Forces* 90 (4): 1369–98. https://doi.org/10.1093/sf/sos073.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley et al. 2019. *The Tidyverse*. https://CRAN.R-project.org/package=tidyverse.

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

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in r." *Implementing Reproducible Computational Research*.

Zakaria, Fareed. 1999. *From Wealth to Power: The Unusual Origins of America's World Role*. Vol. 84. Princeton University Press.