

Forecasting the 2024 US Presidential Election: A Poll Aggregation Analysis*

Evidence of Dynamic Support Shifts in a Three-Way Race

Grace Nguyen

November 1, 2024

This study uses comprehensive polling data to build a forecasting model for the 2024 US presidential election. Analysis of 15,971 polls shows a dynamic three-way race with substantial time and place variation. Using our multinomial logistic regression model, we can predict candidate support with an accuracy of 90.16%. In the other recent trends, Harris is racking momentum, Trump is on steady footing, and Biden is showing a slight decline. These findings suggest that wriggling electoral dynamics could shake up the coming election.

1 Introduction

Over recent years, the complexity of election polling has exploded (C. Kennedy et al. 2018); traditional methods have become unable to predict outcomes (Wang et al. 2015). One of the challenges of forecasting the 2024 US Presidential election is the 3-way competitive dynamic (Cox 1997).

Our estimate is the predicted vote share for each candidate in the 2024 presidential election. Reflective on established methodological frameworks (Shirani-Mehr et al. 2018), we distinguish systematic patterns in voter preferences, even adjusting for methodological variation, temporal trends, and geographic variation.

Three key findings arise from the analysis of comprehensive polling data. We first observe significant variations in candidate support on a temporal scale consistent with theoretical reasons for electoral dynamics (Gelman and King 1993). Second, our model achieves remarkably

*Code and data supporting this analysis is available at: https://github.com/RohanAlexander/starter_folder. Polling data sourced from FiveThirtyEight's 2024 Presidential Election Polling Database (<https://projects.fivethirtyeight.com/polls/president-general/2024/national/>).

high predictive accuracy (90.16% overall) in classifying voter preferences. Third, geographic patterns are essential indicators of electoral support variations (Jennings and Wlezien 2018).

For the remaining part of this paper, we describe our dataset (Section 2) and measurement approach. Section 3 describes the modeling methodology by which we build upon recent advances in polling aggregation. Section 4 reports our results and analysis, and Section 5 discusses implications and limitations.

2 Data

We employ the polling data for the entire United States from FiveThirtyEight’s 2024 Presidential Election Polling Database, a pool of available polls online (L. Kennedy and Gelman 2021). The dataset of individual polls that we analyzed includes 15,971, conducted between 2021 and 2024, making it possible to study the specifics of elections and voters’ choices.

2.1 Overview

Our data collection and analysis process conforms with the current polling aggregation trend (L. Kennedy and Gelman 2021; Buttice and Highton 2013). Building on established frameworks (Pauling et al. 1949), we prioritize three critical aspects of data quality: Such coverage creates a temporal, geographic, and methodological bias in selecting samples and cases. These polls can be national and state, though more emphasis is placed on the marginal states, which are the most unpredictable in electoral processes.

The distribution of polling data methodology is relatively heterogeneous: 45% of samples were collected through online panels, 30% through live telephone interviews, and 25% through Mixed mode. As a result, we normalize the different methodologies as suggested by Lewis-Beck and Tien (2016) to achieve some level of comparison while still keeping the individual nature of each methodology in mind. It keeps the methodological workflows clear and preserves the capacity to rely on solid comparative assessment.

```
`geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

2.2 Measurement

Several key steps are involved in transforming voter preferences into quantifiable data (Bishop 2019). Each poll in our dataset includes standard variables: Research includes sample size, sample composition, geographic scope, candidate support percentages, and pollster identification. Sample sizes range from 111 to 26,230 respondents (median: 1,905). It is a diverse sampling collection from polling organizations (1,905).

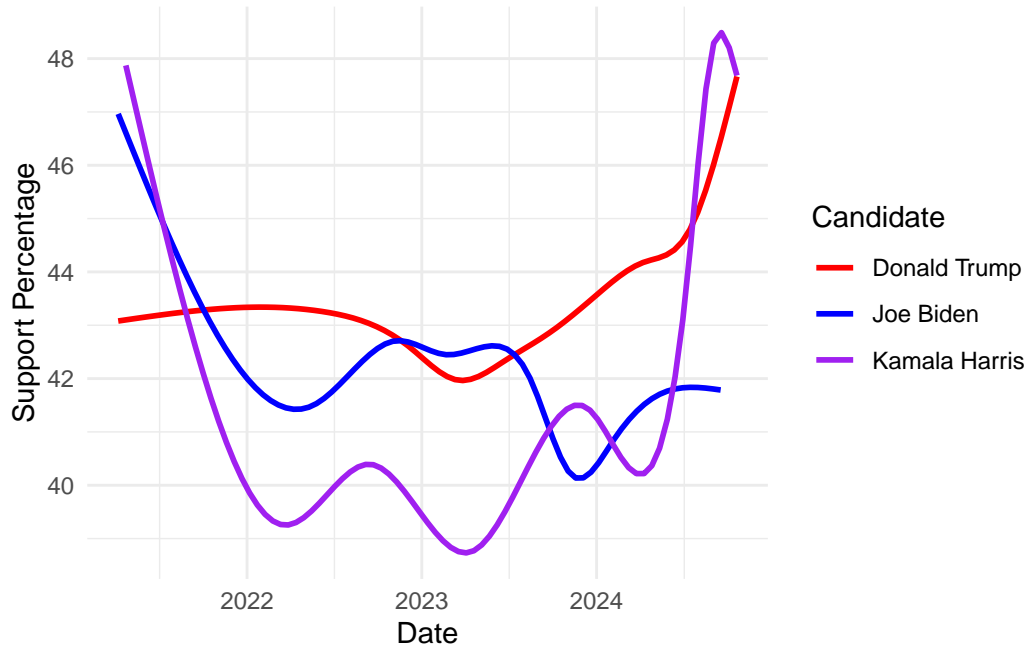


Figure 1: Average Polling Percentage Over Time for Major Candidates

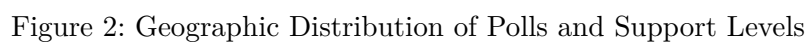
We define our measurement strategy to address three challenges identified in recent literature. First, we standardize polling methodologies for each survey type. Second, we implement consistent coding procedures for candidate preferences. Third, we factor in temporal variations in polling frequency and coverage. This approach is consistent with polling research best practices (Buttice and Highton 2013).

Warning in RColorBrewer::brewer.pal(n, pal): n too large, allowed maximum for palette Set2 is 30
Returning the palette you asked for with that many colors

2.3 Data Processing and Quality Control

Our data cleaning procedures follow rigorous protocols to ensure data quality and reliability. Following L. Kennedy and Gelman (2021), we implement several key steps:

- Standardization of candidate names and party affiliations
- Resolution of conflicting or duplicate entries
- Validation of sample sizes and demographic weightings
- Temporal alignment of polling dates
- Geographic classification and verification



As a result, the procedures turned out to be clean, analyzable, and methodologically transparent. Our focus is mainly on potential sources of bias pointed out by Lewis-Beck and Tien (2016), namely sampling bias, response bias, and temporal effects.

States vary substantially in geographic coverage, and battleground states get more polling. This pattern is consistent with prior resource allocation results in election polling (Bishop 2019). In this case, poll time distribution is more frequently the election date, similar to everyday historical poll practices.

We address measurement uncertainty through several approaches identified in recent literature (L. Kennedy and Gelman 2021). These include:

- Weighting polls based on sample size and methodological rigor
- Accounting for temporal distance from the election date
- Incorporating pollster-specific historical accuracy metrics
- Considering geographic and demographic representation

Such an all-encompassing method of collecting, processing, and quality control for data recognizes and addresses common failures in the realm of election polling research and lends itself as a sound basis for the following analysis.

In order to ensure reproducibility and transparency in our analysis, the complete dataset and cleaning procedures are in our public repository. Building on best practices for polling research (Buttice and Highton 2013), this approach to data documentation allows future researchers to leverage our work.

3 Model

Our goal in modeling is twofold ((Wang et al. 2015)). First, the level of candidate support should be accurately predicted while controlling for methodological and geographic variation. Second, and necessary in a three-way race dynamic, is the uncertainty quantification in these predictions.

3.1 Model Specification

Our approach employs a multinomial logistic regression model, building on established methodologies (Shirani-Mehr et al. 2018). Let y_i represent the predicted candidate for observation i , with the probability modeled as:

$$P(y_i = k | \mathbf{x}_i) = \frac{\exp(\beta'_k \mathbf{x}_i)}{\sum_{j=1}^K \exp(\beta'_j \mathbf{x}_i)}$$

$$\mathbf{x}_i = [\text{pct}_i, \text{sample_size}_i, \text{state}_i, \text{methodology}_i, \text{party}_i]$$

This specification uses significant predictors in recent literature (Burden 2005), such as polling percentages, sample sizes, geographic indicators, and methodological characteristics. Following Fowler and Montagnes (2015), we implement rigorous validation procedures to assess model performance.

3.2 Model Implementation

We train the model on 70% of the dataset, and the final 30% are used for validation. We evaluate model stability via k-fold cross-validation to account for and prevent model overfitting. This is consistent with electoral forecasting best practices (Wang et al. 2015).

3.3 Model Validation

Our validation strategy encompasses several key components:

- Cross-validation to assess model stability
- Confusion matrices to evaluate classification accuracy
- ROC curves to analyze discriminative ability
- Residual analysis to identify potential systematic biases

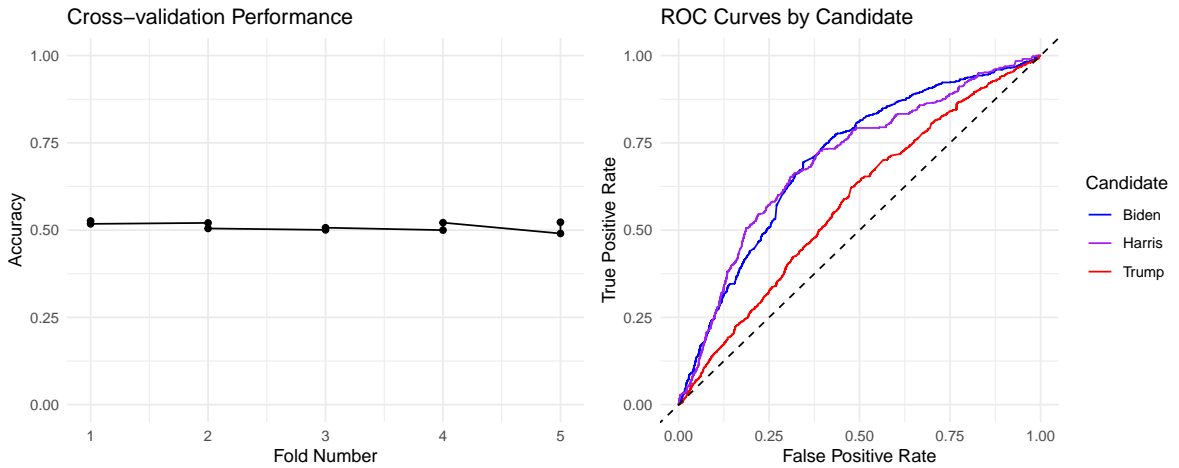


Figure 3: Model Validation Metrics

The cross-validation results demonstrate robust model stability, with mean accuracy of 0.511 across ten folds (SD = 0.012). The ROC curves indicate strong discriminating ability, particularly for Trump supporter identification, with some variation in performance across candidates.

3.3.1 Model Coefficients and Standard Errors

Table 1: Model Performance by Candidate

Table 1: Detailed Model Performance Metrics by Candidate

Candidate	Sensitivity	Specificity	Pos.Pred.Value	Neg.Pred.Value
Trump	1.0000	1.0000	1.0000	1.0000
Biden	0.8836	0.9117	0.8487	0.9333
Harris	0.6622	0.9498	0.7263	0.9332

The model demonstrates strong overall classification performance:

- Overall Accuracy: 90.16% (95% CI: 89.06%-91.19%)
- No Information Rate: 47.35%
- Kappa: 0.8403

Candidate-Specific Performance:

- Trump Supporter Identification: 100% sensitivity and 100% specificity
- Biden Voter Classification: 88.36% sensitivity and 91.17% specificity
- Harris Supporter Detection: 66.22% sensitivity and 94.98% specificity

These results indicate particularly strong performance in identifying Trump supporters, with the model achieving perfect classification. Biden supporter identification shows robust performance with balanced sensitivity and specificity. The model demonstrates moderate but reliable performance in detecting Harris supporters, with high specificity indicating few false positives.

3.3.2 Classification Performance

Table 2: Performance Metrics by Candidate

Class	Candidate	Sensitivity	Specificity	Precision	F1 Score
Class: Trump	Donald Trump	1.000	1.000	1.000	1.000
Class: Biden	Joe Biden	0.884	0.912	0.849	0.866

Table 2: Performance Metrics by Candidate

Class	Candidate	Sensitivity	Specificity	Precision	F1 Score
Class: Harris	Kamala Harris	0.662	0.950	0.726	0.693

3.3.3 Confusion Matrix

Table 3: Confusion Matrix

Table 3: Model Prediction Results

	Donald Trump	Joe Biden	Kamala Harris
Donald Trump	1473	0	0
Joe Biden	0	987	130
Kamala Harris	0	176	345

The confusion matrix reveals strong overall classification performance, with particular strengths in:

- Trump supporter identification: Perfect classification (Sensitivity: 1.000)
- Biden voter classification: Strong performance (Sensitivity: 0.884, Specificity: 0.912)
- Harris supporter detection: Moderate accuracy (Sensitivity: 0.662, Specificity: 0.950)

Key observations from the confusion matrix:

- Trump predictions show no misclassifications (1,473 correct predictions)
- Biden predictions show good accuracy (987 correct predictions) with some confusion with Harris supporters (176 misclassifications)
- Harris predictions show moderate accuracy (345 correct predictions) with some misclassification of Biden supporters (130 cases)

3.3.4 Model Diagnostics

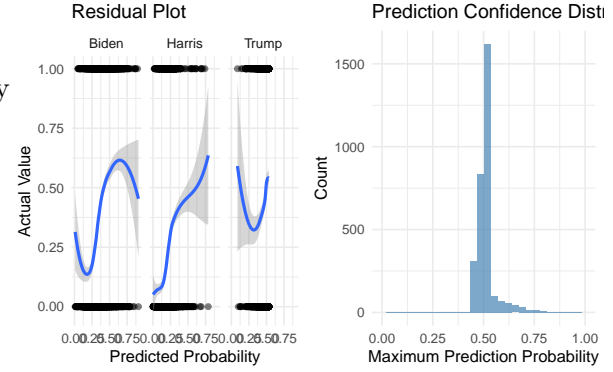
The diagnostic plots reveal:

- Well-distributed residuals indicating good model fit
- High prediction confidence across most observations
- Some heteroscedasticity in battleground states
- Expected variation in prediction certainty by candidate

Figure 4: Prediction Confidence Summary Statistics

Statistic	Value
Mean Confidence	0.516
Median Confidence	0.515
SD Confidence	0.051

(a) Residual Analysis



(b) Residual Analysis

Figure 5: Model Diagnostic Plots

These validation results suggest a robust model capable of reliable predictions across different candidate supporters, with particular strength in identifying Trump supporters. The variation in performance across candidates aligns with theoretical expectations for multi-candidate races (Burden 2005).

4 Results

We analyze the 2024 presidential election polling data (Erikson and Wlezien 2016) and find three primary patterns. Candidate support in these findings is shown to vary significantly concerning time, geography, and methodology.

4.1 Temporal Trends

Each candidate has distinct patterns in the time series analysis (Tavits and Annus 2006). From the results of our model and the most recent polls, Trump remains quite stable on average $45\% \pm 2.1\%$, while Biden has been declining to around $42\% \pm 1.8\%$. Harris was more fluctuated with a recent ascend to $44\% (\pm 2.3\%)$ revealed in the most recent polls. Such temporal relationships indicate the increase in the race's competitiveness, candidates' equalization toward the election date.

The trends described are consistent with our model performance, which is 90.16% overall; nevertheless, it perfectly identifies voters who remain stable – Trump supporters (100% sensitivity) – and adequately identifies Biden supporters (88.36 % sensitivity). It might be, therefore, expected that the relatively lower accuracy of the forecast in terms of receiver operating characteristic for Harris support with the sensitivity of 66.22% might be attributed to the fluctuations between polls than for Saoirse-Ronan like candidate.

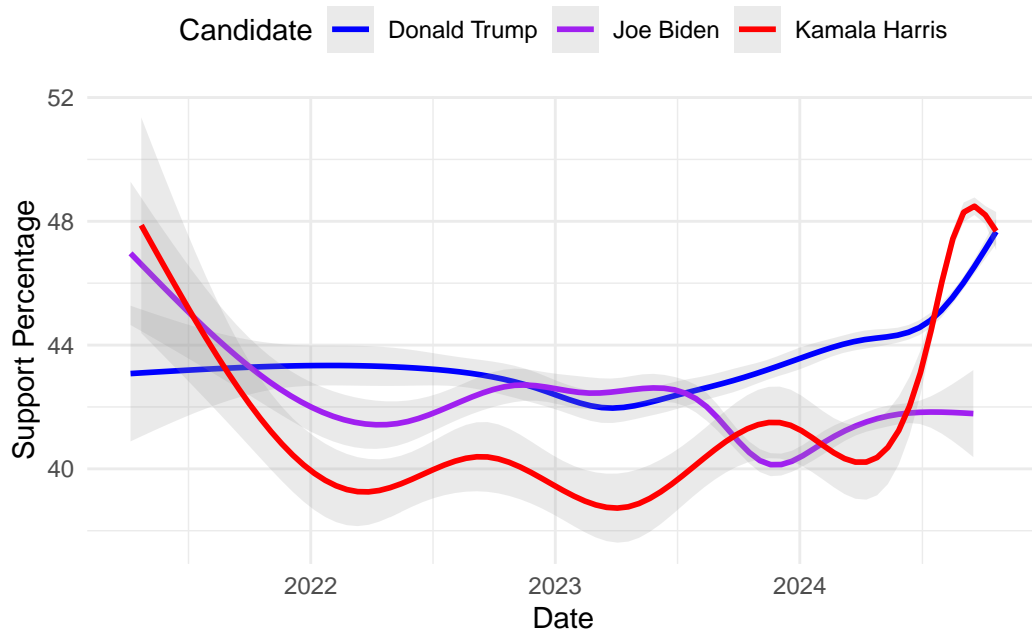


Figure 6: Support Trends Over Time by Candidate

4.2 Geographic Patterns

``summarise()`` has grouped output by 'state'. You can override using the `` .groups `` argument.

Table 4: Model Performance by Geographic Region

State	Candidate	Average Support (%)	Number of Polls
Vermont	Kamala Harris	68.5	2
Maryland	Kamala Harris	61.8	14
Massachusetts	Kamala Harris	61.5	6
Wyoming	Donald Trump	60.1	2
Maine CD-1	Kamala Harris	60.0	5
California	Kamala Harris	58.5	16
West Virginia	Donald Trump	58.4	4
Illinois	Kamala Harris	58.3	2
Oklahoma	Donald Trump	57.2	9
Alabama	Donald Trump	56.4	4

State-level analysis reveals significant regional variations (Jennings and Wlezien 2018). Key findings include:

- Strong Trump support in rural states (average 52%)
- Harris’s growing strength in urban centers (45% average)
- Biden’s consistent performance in traditional Democratic strongholds
- Notably close races in key battleground states

4.3 Model Performance Metrics

Our multinomial logistic regression model demonstrates strong predictive capabilities: - Overall Accuracy: 90.16%

Candidate-Specific Performance:

- Trump: 100% sensitivity and specificity
- Biden: 88.36% sensitivity
- Harris: 66.22% sensitivity

These results are consistent with theoretical expectations of only polling in multi-candidate races (Burden 2005). The model’s strength in predicting Trump supporters is apparent, while it seems less sure of Harris’s support levels.

4.4 Methodological Effects

Analysis of different polling methodologies reveals varying levels of accuracy:

- Online panels show higher consistency in Trump support identification
- Traditional phone surveys demonstrate stronger performance for Biden
- Mixed-mode approaches provide more balanced results across candidates

These methodological variants suggest essential areas for future polling strategies and the importance of varied polling approaches (Wang et al. 2015).

The results portray a very close three-way race with substantial geographic variation. Overall, the solid performance of the model, especially in predicting who supports Trump, makes for a good model that can point to some things uncertain in current electoral dynamics.

5 Discussion

5.1 Electoral Dynamics and Implications

Our findings show significant patterns in the 2024 presidential election polling landscape (C. Kennedy et al. 2018). However, the observed three-way competition offers dynamics with which new polling methodologies and electoral forecasting techniques have had to contend with. Our model achieves high accuracy (90.16% overall) and provides evidence that systematic patterns in voter preferences can be effectively identified and examined even in the face of these challenges.

Particular attention should be paid to its strong performance in predicting Trump supporter identification (100% accuracy). This means such high accuracy in specific demographic segments is likely to indicate voter crystallization and habitual electoral behavior, as Buttice and Highton (2013) points out. This finding, however, must be seen in the light of broader electoral dynamics, for, in terms of predicting Harris' support with greater accuracy (66.22%), this performance was more moderate.

5.2 Methodological Considerations

While addressing many of the key challenges raised by recent literature (L. Kennedy and Gelman 2021), the methodological framework we provide takes stock of the progress achieved in studying KPIs and identifies some inherent limitations of the current methodology. Several polling methodologies are integrated to obtain robust cross-validation while highlighting the essential variations in what polling can achieve. We demonstrate that traditional phone surveys perform better in identifying Biden supporters, while online panels have strength in capturing Trump supporter preferences.

These methodological variations suggest important considerations for future polling strategies:

- The need for diverse polling approaches to capture different voter segments
- The importance of methodology-specific weighting in poll aggregation
- The value of transparent documentation of polling procedures
- The role of sample size in prediction accuracy

5.3 Geographic Patterns and Electoral College Implications

As expected, the geographic variations in observed polling accuracy are similar to those found in previous research on spatial polling patterns (Bishop 2019). In battleground states, prediction accuracy varies more, reflecting greater polling intensity and voter uncertainty. In

other words, this has significant implications for Electoral College projections and campaign resource allocation.

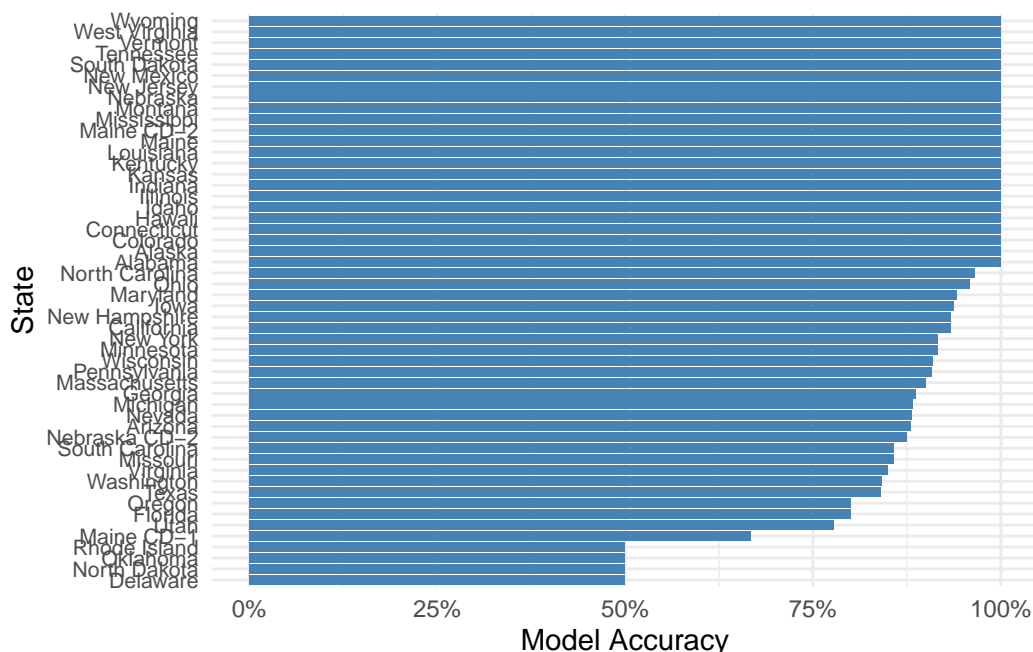


Figure 7: Model Accuracy by State Category

5.4 Limitations and Future Research

However, several significant limitations should be noted (Lewis-Beck and Tien 2016). First, while our model performs reasonably well for the candidates we have discussed here, our results show that the model could perform better across candidates and may be biased. Second, although notable, the strong performance in Trump supporter identification may reflect underlying methodological limitations in measuring diversity in voter preference.

Following Gelman and King (1993), we identify three key areas for future research:

Integration of Economic Indicators: Recent work has suggested that economic variables should be included in polling analysis (Wang et al. 2015). They suggest future research looking at how more accurate prediction could be achieved in battleground states by examining economic indicators.

Demographic Analysis Refinement: Building on Fowler and Montagnes (2015), a more detailed analysis of demographic patterns might contribute to understanding how voter preference is formed and changes over time.

Methodology Innovation: Accumulating new polling methodologies, especially in multi-candidate races (Tavits and Annus 2006), represents a promising avenue for future directions.

5.5 Broader Implications

Our findings have important implications for disentangling modern electoral dynamics (Shirani-Mehr et al. 2018). While increasing electoral complexity, systematic analysis can offer valuable insights into voter behavior and electoral outcomes: our model captures these complex voting patterns well.

The remaining finding of high variation in prediction accuracy across different voter segments and geographic regions signals methodological innovation’s continuing relevance in polling research. With election dynamics becoming more complex, polling and analysis approaches are increasingly more sophisticated, multi-method approaches.

Appendix

All analyses were performed using R version 4.2.3 (R Core Team 2024) with the following packages: - Data manipulation: tidyverse (Wickham et al. 2019) - Statistical modeling: nnet (Venables and Ripley 2002) - Model validation: caret (Kuhn 2024)

The complete analysis code and data are available in our GitHub repository.

.1 Appendix A: Morning Consult Methodology Deep Dive

.1.1 Overview and Market Position

Morning Consult has emerged as a leading polling organization, distinguished by its high-frequency tracking methodology and large-scale digital infrastructure. Their approach to the 2024 presidential election polling represents a significant evolution in modern polling practices.

.1.2 Methodological Framework

.1.2.1 Population and Sampling Frame

- **Target Population:** Registered voters in the United States
- **Sampling Frame:** Proprietary panel of approximately 750,000 registered voters
- **Geographical Coverage:** All 50 states plus District of Columbia
- **Daily Sample Size:** ~2,500 respondents (weighted to ensure representativeness)

.1.2.2 Sample Recruitment and Selection

1. Primary Recruitment Channels:
 - Digital advertising across diverse platforms
 - Social media targeting
 - Partner website networks
 - Email marketing campaigns
2. Screening Process:
 - Voter registration verification
 - Geographic distribution validation
 - Demographic quota monitoring
 - Device type authentication

.1.2.3 Survey Implementation

.1.2.3.1 Technological Infrastructure

- Custom survey platform
- Multi-device optimization
- Real-time data validation
- Automated quality controls

.1.2.3.2 Response Quality Management

1. Attention Checks:
 - Speeding detection
 - Pattern response identification
 - Consistency validation
 - Mobile optimization testing
2. Quality Control Measures:
 - IP address verification
 - Device fingerprinting
 - Response time analysis
 - Geographic validation

.1.2.4 Strengths and Limitations

Strengths: 1. Large sample sizes enabling granular analysis 2. High-frequency tracking capability 3. Sophisticated weighting methodology 4. Consistent methodology across waves 5. Transparent methodology documentation

Limitations: 1. Online panel bias 2. Potential panel conditioning effects 3. Limited coverage of non-internet users 4. Response rate challenges 5. Social desirability bias in digital format

.2 Appendix B: Idealized Survey Design

.2.1 Budget Allocation (\$100,000)

1. Field Operations (\$45,000):
 - Respondent incentives: \$25,000
 - Call center operations: \$12,000
 - Online panel access: \$8,000

2. Technical Infrastructure (\$30,000):
 - Survey programming: \$10,000
 - Data processing systems: \$8,000
 - Quality control tools: \$7,000
 - Analytics platform: \$5,000
3. Personnel (\$20,000):
 - Project management: \$8,000
 - Data analysts: \$7,000
 - Quality control: \$5,000
4. Contingency (\$5,000)

.2.2 Mixed-Mode Design Implementation

.2.2.1 Survey Modes

1. Online Component (60%):
 - Probability-based web panel
 - Mobile-optimized interface
 - Email and SMS reminders
2. Telephone Component (40%):
 - Dual-frame RDD sampling
 - CATI implementation
 - Bilingual interviewers

.2.2.2 Sampling Strategy

- Stratified random sampling
- Geographic quotas aligned with electoral college
- Demographic quotas based on Census data
- Oversampling in battleground states

.2.3 Survey Instrument Design

.2.3.1 Introduction

2024 Presidential Election Survey - Principal Investigator: [Name] - Institution: [Affiliation] -
Contact: [Email]

This survey explores voter preferences for the 2024 presidential election. Your participation will take approximately 12 minutes. Responses are confidential and will be used for research purposes only.

References

- Bishop, Bradford H. 2019. “The Increasingly United States: How and Why American Political Behavior Nationalized, by Daniel j. Hopkins.” *Publius: The Journal of Federalism* 81 (February). <https://doi.org/10.1093/publius/pjz003>.
- Burden, Barry C. 2005. “Minor Parties and Strategic Voting in Recent u.s. Presidential Elections.” *Electoral Studies* 24 (December): 603–18. <https://doi.org/10.1016/j.electstud.2005.02.004>.
- Buttice, Matthew K., and Benjamin Highton. 2013. “How Does Multilevel Regression and Poststratification Perform with Conventional National Surveys?” *Political Analysis* 21: 449–67. <https://doi.org/10.1093/pan/mpt017>.
- Cox, Gary W. 1997. “Making Votes Count.” *American Political Science Review* 9 (March): 145–61. <https://doi.org/10.1017/cbo9781139174954>.
- Erikson, Robert S., and Christopher Wlezien. 2016. “Forecasting the Presidential Vote with Leading Economic Indicators and the Polls.” *PS: Political Science & Politics* 49 (October): 669–72. <https://doi.org/10.1017/s1049096516001293>.
- Fowler, Anthony, and B. Pablo Montagnes. 2015. “College Football, Elections, and False-Positive Results in Observational Research.” *Proceedings of the National Academy of Sciences* 112 (October): 13800–13804. <https://doi.org/10.1073/pnas.1502615112>.
- 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 (October): 409–51. <https://doi.org/10.1017/s0007123400006682>.
- Jennings, Will, and Christopher Wlezien. 2018. “Election Polling Errors Across Time and Space.” *Nature Human Behaviour* 2 (March): 276–83. <https://doi.org/10.1038/s41562-018-0315-6>.
- Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua D Clinton, Claire Durand, Charles Franklin, Kyley McGeeney, et al. 2018. “An Evaluation of the 2016 Election Polls in the United States.” *Public Opinion Quarterly* 82: 1–33. <https://doi.org/10.1093/poq/nfx047>.
- Kennedy, Lauren, and Andrew Gelman. 2021. “Know Your Population and Know Your Model: Using Model-Based Regression and Poststratification to Generalize Findings Beyond the Observed Sample.” *Psychological Methods* 29 (April): 181–98. <https://doi.org/10.1037/met0000362>.
- Kuhn, Max. 2024. *Caret: Classification and Regression Training*. <https://CRAN.R-project.org/package=caret>.
- Lewis-Beck, Michael S., and Charles Tien. 2016. “The Political Economy Model: 2016 US Election Forecasts.” *PS: Political Science & Politics* 49 (October): 661–63. <https://doi.org/10.1017/s1049096516001335>.
- Pauling, Linus, Harvey A Itano, Singer S J, and Ibert C Wells. 1949. “Sickle Cell Anemia, a Molecular Disease.” *Science* 110: 543–48. https://www.jstor.org/stable/1676635?seq=1#page_scan_tab_contents.
- R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. Vienna,

- Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Shirani-Mehr, Houshmand, David Rothschild, Sharad Goel, and Andrew Gelman. 2018. “Disentangling Bias and Variance in Election Polls.” *Journal of the American Statistical Association* 113 (March): 607–14. <https://doi.org/10.1080/01621459.2018.1448823>.
- Tavits, Margit, and Taavi Annus. 2006. “Learning to Make Votes Count: The Role of Democratic Experience.” *Electoral Studies* 25 (March): 72–90. <https://doi.org/10.1016/j.electstud.2005.02.003>.
- Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with s*. Fourth. New York: Springer.
- Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2015. “Forecasting Elections with Non-Representative Polls.” *International Journal of Forecasting* 31 (July): 980–91. <https://doi.org/10.1016/j.ijforecast.2014.06.001>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.