**POLITICS**

*Special Issue on Forecasting the 2024 US Elections*

# PREDICTING POPULAR VOTE SHARES AT US PRESIDENTIAL ELECTIONS: A MODEL-BASED STRATEGY RELYING ON ANES DATA

**Stefano Camatarri**, *Autonomous University of Barcelona, Spain*

Election forecasting in modern democracies faces significant challenges, including increasing survey nonresponse and selection bias. Added to this are the limitations of current predictive approaches. While structural models focus solely on macro-level variables—such as economic conditions and leader popularity—thereby overlooking the importance of individual-level factors, survey-based aggregation methods often rely on intuitive procedures that lack theoretical foundations.
To address these gaps, this contribution proposes a combined logistic regression approach (both standard and Bayesian) that leverages voter-level data and incorporates a theory-based specification. By testing these models on recent waves of the American National Election Studies (ANES) Time Series, this study demonstrates that the proposed approach yields notably accurate predictions of Republican popular support in each election.

**Keywords:** election forecasting, voting intentions, US Presidential elections, regression analysis, ANES



## Introduction

It is well known that in modern democracies, including the U.S., election forecasting faces significant challenges. The increasing trend of survey refusals (Plewes and Turangeau 2013), particularly among segments of the population more likely to support specific parties or candidates (e.g., Kennedy et al. 2018), exacerbates these difficulties. Additionally, factors like the spiral of silence and cross-cutting pressures can prevent some respondents from accurately disclosing their true political preferences. This often results in biased self-reported data (Blair et al. 2020) and skewed election forecasts.
However, these issues are not the only hurdles faced by electoral researchers and analysts. There are also significant concerns about the effectiveness of the methods used to capture trends and infer likely aggregate outcomes of electoral processes. Most studies rely primarily on inductive aggregation procedures—whether weighted or unweighted—of individual self-reported voting intentions, or on theory-based models that examine structural relationships between macro-level variables, such as economic conditions and electoral outcomes (see Dassonneville and Lewis-Beck 2015, for an overview). While both approaches have contributed significantly to the field of election forecasting, they often overlook the substantive mechanisms underlying the outcomes they aim to predict - namely, voter-level decision-making dynamics.
Building on this premise, this article argues that citizens' reasoning is a crucial lens for understanding decision-making processes and predicting their aggregate effects. It demonstrates that voter-based regression models can enhance election predictions while complementing existing aggregation and structural strategies. Notably, this analysis draws on high-quality data from the American National Election Study (ANES), an approach that is rarely employed in forecasting exercises. Using pre-electoral datasets from the last three

This is a "preproof" accepted article for *PS: Political Science & Politics*. This version may be subject to change during the production process.

DOI: 10.1017/S1049096524000933

presidential elections (2012 2016, and 2020), I show that predictions of support for Republican candidates - often underrepresented in U.S. election polls - are equally or even more accurate than those generated by other methods. Working with high quality individual-level data to tackle prediction issues has not been used often for forecasting purposes and represents a valuable step forward in addressing prediction challenges

**Background: Leveraging ANES Data to Predict Popular Vote Shares at US Presidential Elections**

Election forecasting in established democracies, including the U.S., is a well-established practice with a long history. As mentioned above, Dassonneville and Lewis-Beck (2015) offer a comprehensive overview of the evolution of various prediction approaches. They distinguish between a theory-based approach, known as structural modeling, which predicts election outcomes using multivariate equations that account for a range of macro-level factors (e.g., Lewis-Beck and Stegmaier 2013), and a more inductive approach, known as aggregation, where analysts estimate vote shares by applying combination rules to multiple election polls (e.g., Traugott 2014). As a hybrid of these two methods, 'synthesizers' propose model-based predictions that also incorporate polls as additional data (e.g., Mongrain et al. 2021).

Importantly, all these methods share a common characteristic: their units of analysis, and therefore the level at which they perform inferences, tend to be at the state or national level. In the case of structuralist approaches, this strategy is deliberately employed to avoid relying on surveys, which can introduce estimation errors into the predictions. However, macro-level approaches are consistently affected by the issue of ecological fallacy, as they attempt to predict outcomes of a process that inherently involves a strong micro-level component (i.e., voters' decision-making) using only macro-level factors such as the state of the economy (Kramer, 1983). This limitation highlights the analytical advantages of survey approaches, which have made significant progress in recent years by addressing issues such as nonresponse bias and systematic measurement error through various means, including controlling for reported past voting behavior and sociodemographics. On the other hand, mounting polarization, combined with spiral of silence mechanisms - where citizens may be reluctant to express their pre-electoral preferences due to fear of social isolation for holding minority opinions in their environment -makes this challenge particularly difficult to address. Camatarri et al. (2023) provide clear evidence of the spiral of silence effect among Republican voters in 2020, based on an analysis of survey data and county-level election results from the prior Presidential Election. This aligns with findings from extant research (e.g., Urquizo Sancho 2006; Dinas et al. 2024), indicating that right-wing respondents are less likely to disclose their political preferences in environments perceived as hostile to their views. In contrast, Democrats do not exhibit the same pattern. This difference is likely attributable to their greater psychological openness—a trait that enables them to express their views more freely and engage in political discussions even in contexts where their opinions may be in the minority (Gelder 2010; Mutz 2002). However, it is important to acknowledge that not all Republican supporters share the same approach to expressing their preferences. While some may fear social isolation and underreport their support, more nonconformist Republicans may feel less constrained by holding minority opinions within their social circles (e.g., Kushin et al. 2019). This diversity among Republican supporters likely contributes to the fact that, despite noticeable polling misses, the discrepancies between election forecasts and actual outcomes are generally not massive.

On top of this, it is also important to note that a potential cluster of Republican supporters - such as those with lower educational levels, anti-elite views, relatively weak partisan affiliations, and political disaffection - are generally less likely to participate in surveys. This more general survey nonresponse bias, in addition to the sensitivity of the vote intention

question especially for Trump supporters, can increase prediction errors in estimating Trump's aggregate support (Kennedy et al. 2018).
To address these challenges and improve the quality of survey estimates, a comprehensive approach is essential. First, using individual-level data, rather than aggregate-level data, enables more accurate inferences about the micro-level factors that influence election outcomes. Second, applying post-estimation weighting helps to correct for selection bias and ensures more representative estimates across different population segments.

**Data and estimation procedure**

Against the background described above, I propose a straightforward approach to predicting Republican candidate popular support. This method addresses the need to leverage individual-level data to avoid ecological fallacies and adjusts for selection bias through weighting. The process begins with estimating an individual-level vote function grounded in electoral behaviour theory. The predicted values from this model are then used to infer individual support for Trump (predicted votes), which are subsequently aggregated to estimate the candidate's overall vote share.[i]
ANES data is particularly well-suited for this approach, as longitudinal analyses (1952-2020) have consistently demonstrated that its vote intention measures reliably reflect aggregate popular vote outcomes (Ko et al. 2024). For this analysis, I utilize the pre-electoral waves of the 2012 2016, and 2020 ANES Time Series Study, which included 5,914, 4,270, and 8,280 overall respondents, respectively. The interviews were randomly assigned to different modes, including face-to-face and web for 2012 and 2016, as well as telephone and video interviews for 2020. Importantly, in addition to capturing vote intentions, the surveys collect extensive information on respondents' socio-demographic and attitudinal backgrounds. This comprehensive data allows for the construction of a robust vote function that accounts for most factors influencing individual voting behavior, both traditional and novel.
Starting from a retrospective framework, the model incorporates variables such as (dis)approval of presidential economic performance over the past year and respondents' feelings about their financial situation compared to the previous year. It also includes general (dis)trust toward the federal government. From a positional perspective, the model considers respondents' self-placement on a 7-point ideological scale ranging from liberal to conservative. In conclusion, the estimations also account for several socio-demographic controls, including gender, age, highest education level, and ethnic background (simplified dichotomously for comparison: White vs. all other categories such as Black, Hispanic, Asian/Hawaiian, Native American/Alaska Native, or others including multiple non-Hispanic backgrounds).[ii] It is crucial to note that the model incorporates state-level fixed effects and normalizes all variables between 0 and 1 to ensure comparability of coefficients' sizes.
The employed models are logistic regressions predicting voting intentions for the Republican candidate versus all other presidential candidates in 2012 2016, and 2020. One key advantage of logistic regression in this context is its simplicity and interpretability, along with its suitability for predicting binary outcomes, such as the one of interest in this analysis. Odds ratios can be directly interpreted as changes in probability associated with a one-unit change in the predictor variables, making the model's outputs relatively intuitive.
To further validate the results and address the issue of statistical uncertainty, this study also employs Bayesian logistic regression models using Markov Chain Monte Carlo (MCMC) methods.[iii] By complementing mainstream logistic regression models with Bayesian estimations, the study enhances the robustness of the findings regarding the factors influencing the binary outcome of interest and the resulting predictions.[iv]
The predicted probabilities from both methods categorize respondents as either Republican supporters or non-Republican supporters, alternating the standard 0,5 probability threshold and the weighted mean probability as cut-off points. The predicted votes for the Republican

candidate versus other candidates are then aggregated (and weighted) to derive a federal-level estimate of vote shares in each scenario, reflecting the overall popular vote as closely as possible to the reference population. For both steps, the full sample pre-election weight (V200010a) is applied.

**Results**

As previously discussed, the initial step of the analysis involves establishing a baseline individual-level model to generate aggregate predictions of electoral support. The results of the mainstream logistic estimations are detailed in Table A1 in the Appendix. As shown, most predictors have a significant effect on voting intentions for the Republican candidate, aligning with expectations based on existing theories. For example, holding more conservative positions on the ideological scale consistently correlates with a significant increase in the probability of intending to vote for the Republican candidate. Conversely, economic disapproval is identified as one of the strongest factors predicting Republican support, particularly when the Presidential candidate is not an incumbent. In the case of Trump's incumbency, this coefficient reflects a significantly negative effect. A similar trend, though less pronounced, is observed regarding perceptions of a worsening personal economic situation. Regarding socio-demographic factors, it is noteworthy that higher education levels are negatively associated with support for the Republican candidate only during Trump's presidential runs (2016 and 2020). Conversely, White ethnic background consistently emerges as a strong predictor of Republican support across all data points.[v] Importantly, the predictors exhibit highly comparable effects in the Bayesian analysis as well, as indicated by their posterior means for the outcome variable (see Table A2).

The next step is converting the predicted probabilities from each estimation into aggregate estimates of support for the Presidential candidate, specifically focusing on popular vote shares.[vi] To achieve this, we use both the default 0.5 threshold and the average predicted probability (weighted according to the full sample pre-electoral weight) as cut-offs for predicting whether respondents would vote for the Republican candidate.[vii] The results, alongside estimates derived from simple aggregation of voting intentions (both weighted and unweighted) on the full sample, are presented in Table 1. This table also includes the actual percentage of popular support for each Republican candidate based on election results, facilitating comparison and assessment of the different strategies.

First, it should be noted that the estimate derived from straightforward aggregation of actual voting intentions in the data significantly underestimates support for the Republican presidential candidate throughout the entire period. This underscores the pressing challenge of mitigating selection bias in electoral surveys, even with high-quality data such as ANES. Weighting the estimates using the full sample weight slightly improves accuracy; however, the confidence intervals remain significantly below the actual electoral support received by each Republican candidate in the corresponding elections.

Turning to logistic-based approaches (both standard and Bayesian), a clear improvement in the estimates of each candidate is evident, particularly when using the average predicted probability as the cut-off. This method successfully identifies confidence intervals compatible with the actual results for the Republican candidates. Notably, in 10 out of 12 overall models, the forecasting error (i.e., the absolute difference between the predicted estimate and the actual result) is well below 3%. This performance is significantly better than the mean absolute error of major commercial polls and the entire 1952-2020 ANES series, which averages around 3% for predicting the incumbent candidate's vote share (see Ko et al. 2024). Moreover, this result aligns with alternative methods tested for the same election (see Erikson and Wlezien 2021), further supporting the efficacy of this approach.

**Table 1. Republican Presidential Candidate Support: Estimates vs. Actual Results (%)**

| Year | Republican Candidate | Simple Aggregation (unweighted) | Simple Aggregation (weighted) | Logit Model (0.5 Threshold) | Logit Model (Average Predicted Probability) | Bayesian Logit (0.5 Threshold) | Bayesian Logit (Average Predicted Probability) | Actual Popular Support |
|---|---|---|---|---|---|---|---|---|
| 2012 | Romney | 38,1 (36,7-39,5) | 43,7 (41,7-45,6) | 47,5 (44,6-50,5) | 49,1 (46,1-52,09) | 47,9 (45,3-50,6%) | 49,1 (46,5-51,7) | 47.2 |
| 2016 | Trump | 41,3 (39,6-43) | 40,7 (38,7-42,6) | 44,4 (42,0-46,7) | 46,28 (43,9-48,6) | 44,4 (42,2-46,6) | 46,5 (44,3-48,7) | 46.1 |
| 2020 | Trump | 42,8 (41,6-43,9) | 43,7 (42,2-45,3) | 43,5 (41,8-45,3%) | 45,6 (43,8-47,3) | 43,2 (41,5-44,8) | 45,4 (43,7-47) | 46.8 |

*Note: Parentheses indicate 95% confidence intervals*

## Conclusions

This study addresses the unique challenges of forecasting electoral support in the current U.S. political climate, marked by increasing nonresponse and selection bias, which compromise the quality of estimates (e.g., Enns 2017). To tackle these challenges, I showed that a model-based estimation approach relying on individual-level data and theoretically relevant variables can yield results closely aligned with actual election outcomes. The rationale for using a model-based approach lies in its analytical appropriateness, as it targets the decisive levels where electoral decisions are influenced, i.e., individual voters. This adds an important layer of validation to studies that rely solely on aggregation or macro-level modeling. Overall, standard logistic approaches provided highly accurate predictions across all three elections, with minimal forecasting errors (as low as 0.18% in 2016). In contrast, unweighted and weighted simple aggregation methods consistently produced higher forecasting errors, underestimating support for Republican candidates. The Bayesian logistic models followed similar patterns to the standard logistic models, although they occasionally exhibited slightly higher forecasting errors.

Overall, these results highlight that regression-based voter-level forecasting models have the potential to offer strong predictive performance in estimating popular vote shares. In doing so, they position the ANES 2020 Time Series alongside other forecasting approaches that maintain an absolute error of less than 3 percentage points (e.g., Graefe 2021).

Despite the encouraging results, several important limitations warrant consideration. First, the analysis primarily focuses on predicting popular support. While this emphasis is vital for understanding general opinion trends and their macro-consequences, it overlooks critical components of the electoral system, particularly the role of the Electoral College in U.S. Presidential elections. This omission is significant and suggests that future efforts should expand the methodology to encompass both federal and state levels, enabling a more nuanced prediction not only of election outcomes but also of actual winners.

Moreover, future research should explore alternative model specifications and evaluate the potential benefits of incorporating election-specific predictors, including for the forthcoming 2024 elections, beyond the standard model employed here. To maximize comparability, our estimates did not account for factors such as the impact of COVID-19 and immigration attitudes. The former proved central to voters' decisions in 2020 (see Luartz et al. 2024), while border security and immigration were pivotal themes in Trump's successful 2016 campaign. Reports indicate that these issues persist -at least in terms of candidate rhetoric -as we approach the 2024 election (Tourangbam 2024). Additionally, the significance of women's rights and civil rights has become increasingly pronounced, particularly following the overturning of Roe v. Wade during the Trump administration. In contrast, the 2012

election was characterized by a greater emphasis on economic and healthcare issues, reflecting the electorate's concerns during the recovery from the Great Recession (see Kiousis et al. 2015).
Looking ahead to the 2024 election, a combination of domestic and foreign issues is likely to play a crucial role in shaping voter decisions. In addition to the enduring importance of economic concerns, topics such as women's rights, immigration security, the Ukraine war, and climate change are emerging as significant focal points in the current presidential debates, indicating potential shifts in voter priorities. A systematic approach to incorporating these salient issues for each election cycle will be essential in informing future models, striking a balance between maintaining comparability and adequately capturing the complexities of voters' decision-making processes.

**ACKNOWLEDGEMENTS**
I would like to sincerely thank the Editors of the Special Issue on Forecasting the 2024 Presidential Elections, Philippe Mongrain and Mary Stegmaier, for the opportunity to contribute with this manuscript, as well as for their valuable and constructive feedback on the draft version. I also extend my gratitude to the two anonymous reviewers for their thoughtful comments and engagement, which helped me strengthen this contribution both conceptually and analytically.

**DATA AVAILABILITY STATEMENT**
Research documentation and data that support the findings of this study have not yet been verified by PS's replication team. Data will be openly available at the Harvard Dataverse upon publication of the final article.

**CONFLICTS OF INTEREST**
The authors declare no ethical issues or conflicts of interest in this research.

**References**
Acquah, Henry De-Graft. 2013. "Bayesian Logistic Regression Modelling via Markov Chain Monte Carlo Algorithm." *Journal of Social and Development Sciences* 4(4): 193-197.
Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. "Parallel worlds: fixed effects, differences-in-differences, and panel data." In *Mostly Harmless Econometrics*, eds. Joshua D. Angrist and Jörn-Steffen Pischke, 221-248. Princeton, NJ: Princeton University Press.
Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." *American Political Science Review* 114 (4): 1297-1315.
Camatarri, Stefano. 2022. "What If It Kicked in More Strongly? A Counterfactual Analysis of Protest Voting's Electoral Consequences in Greece, Italy, and Spain." *Representation* 58 (1): 103-118.
Camatarri, Stefano, Lewis A. Luartz, and Marta Gallina. 2023. "Always Silent? Exploring Contextual Conditions for Nonresponses to Vote Intention Questions at the 2020 US Presidential Election." *International Journal of Public Opinion Research* 35 (3): edad025.
Coppock, Alexander. 2017. "Did Shy Trump Supporters Bias the 2016 Polls? Evidence from a Nationally-Representative List Experiment." *Statistics, Politics, and Policy* 8 (1): 29-40.
Cramer, Jan Salomon. 2003. *Logit Models from Economics and Other Fields.* Cambridge: Cambridge University Press.

Dinas, Elias, Sergi Martínez, and Vicente Valentim. 2024. "Social Norm Change, Political Symbols, and Expression of Stigmatized Preferences." *The Journal of Politics* 86 (2): 488-506.

Lewis-Beck, Michael S., and Ruth Dassonneville. 2015. "Comparative Election Forecasting: Further Insights from Synthetic Models." *Electoral Studies* 39: 275-283.

Enns, Peter K., Julius Lagodny, and Jonathon P. Schuldt. 2017. "Understanding the 2016 US Presidential Polls: The Importance of Hidden Trump Supporters." *Statistics, Politics, and Policy* 8 (1): 41-63.

Erikson, Robert S., and Christopher Wlezien. 2021. "Forecasting the 2020 Presidential Election: Leading Economic Indicators, Polls, and the Vote." *PS: Political Science & Politics* 54 (1): 55-58.

Graefe, Andreas. 2021. "Of Issues and Leaders: Forecasting the 2020 US Presidential Election." *PS: Political Science & Politics* 54 (1): 70-72.

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): 1-33.

Kiousis, Spiro, Ji Young Kim, Matt Ragas, Gillian Wheat, Sarab Kochhar, Emma Svensson, and Maradith Miles. 2015. "Exploring New Frontiers of Agenda Building During the 2012 US Presidential Election Pre-Convention Period: Examining Linkages Across Three Levels." *Journalism Studies* 16 (3): 363-382.

Ko, Hyein, Natalie Jackson, Tracy Osborn, and Michael S. Lewis-Beck. 2024. "Forecasting Presidential Elections: Accuracy of ANES Voter Intentions." *International Journal of Forecasting*. [DOI if available]

Kramer, Gerald H. 1983. "The Ecological Fallacy Revisited: Aggregate-Versus Individual-Level Findings on Economics and Elections, and Sociotropic Voting." *American Political Science Review* 77 (1): 92-111.

Kushin, Matthew J., Masahiro Yamamoto, and Francis Dalisay. 2019. "Societal Majority, Facebook, and the Spiral of Silence in the 2016 US Presidential Election." *Social Media + Society* 5 (2): 2056305119855139.

Lewis-Beck, Michael S., and Mary Stegmaier. 2013. "The VP-Function Revisited: A Survey of the Literature on Vote and Popularity Functions After Over 40 Years." *Public Choice* 157: 367-385.

Menard, Scott. 2002. *Applied Logistic Regression Analysis*. Vol. 106. Thousand Oaks, CA: Sage.

Mongrain, Philippe, Richard Nadeau, and Bruno Jérôme. 2021. "Playing the Synthesizer with Canadian Data: Adding Polls to a Structural Forecasting Model." *International Journal of Forecasting* 37 (1): 289-301.

Mutz, Diana C. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46 (4): 838-855.

Plewes, Thomas J., and Roger Tourangeau, eds. 2013. "Nonresponse in Social Science Surveys: A Research Agenda." *Washington, DC: National Academies Press*.

Tourangbam, Monish. 2024. "Issues and Trends in US Presidential Election 2024." *Institute for Security and Development Policy* [issue brief], March 14, 2024. https://isdp.eu/wp-content/uploads/2024/03/Brief-Monish-Mar-14-2024-final.pdf.

Traugott, Michael W. 2014. "Public Opinion Polls and Election Forecasting." *PS: Political Science and Politics* 47 (2): 342-344.

Urquizu-Sancho, Ignacio. 2006. "The Non-Declared Vote in the Surveys: The Spanish Case in the 1980s." *Electoral Studies* 25 (1): 103-128.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

## Appendix

**Table A1. Factors explaining vote intention for the Republican candidate at US Presidential Elections, logistic regressions***

| | Model 1 (2012) | Model 2 (2016) | Model 3 (2020) |
|---|---|---|---|
| Liberal-conservative | 401.2*** | 374.0*** | 6,247*** |
| | (206.0) | (145.9) | (1,775) |
| Economic disapproval | 84.56*** | 14.28*** | 0.0493*** |
| | (22.25) | (2.285) | (0.00855) |
| Personal economic situation (worse) | 1.946*** | 1.870** | 0.825 |
| | (0.385) | (0.585) | (0.179) |
| Government distrust | 1.149 | 1.036 | 0.442*** |
| | (0.549) | (0.383) | (0.102) |
| Age | 1.550 | 1.307 | 0.828 |
| | (0.606) | (0.405) | (0.147) |
| Education (highest) | 1.590 | 0.139*** | 0.473*** |
| | (0.573) | (0.0710) | (0.0862) |
| White Non-Hispanic | 3.863*** | 3.620*** | 3.215*** |
| | (0.863) | (0.683) | (0.398) |
| Female | 0.881 | 0.712** | 1.061 |
| | (0.161) | (0.101) | (0.0974) |
| Constant | 0.000152*** | 0.0136*** | 0.0932*** |
| | (0.000123) | (0.0125) | (0.0516) |
| Pseudo-R2 | 0,68 | 0,58 | 0.58 |
| Observations | 1,961 | 2,394 | 5,605 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**State-level fixed effects are not shown for simplicity and readability purposes. Additional versions of the model tables are available upon request.*

**Table A2. Posterior Distribution Summaries of parameters from MCMC Bayesian Logistic regression***

| Electoral year | Variables | Posterior Means | Standard Deviation | Quantiles of Posterior Distributions | |
|---|---|---|---|---|---|
| 2012 | | | | 2,5% | 97,5% |
| | Liberal-conservative | 6.245 | 0.157 | 5.919 | 6.515 |
| | Economic disapproval | 4.407 | 0.148 | 4.158 | 4.789 |
| | Personal economic situation (worse) | 0.631 | 0.168 | 0.338 | 1.088 |
| | Government distrust | 0.103 | 0.249 | -0.346 | 0.613 |
| | Age | 0.227 | 0.204 | -0.171 | 0.648 |
| | Education (highest) | 0.552 | 0.135 | 0.285 | 0.770 |
| | White Non-Hispanic | 1.383 | 0.098 | 1.199 | 1.595 |
| | Female | -0.175 | 0.195 | -0.461 | 0.233 |
| 2016 | | | | | |
| | Liberal-conservative | 6.103 | 0.157 | 5.740 | 6.370 |
| | Economic disapproval | 2.711 | 0.114 | 2.485 | 2.928 |
| | Personal economic situation (worse) | 0.593 | 0.142 | .350 | .865 |
| | Government distrust | .162 | 0.234 | -.314 | .521 |
| | Age | .060 | 0.174 | -.329 | .397 |
| | Education (highest) | -1.793 | 0.099 | -2.046 | -1.588 |
| | White Non-Hispanic | 1.401 | 0.151 | 1.133 | 1.681 |
| | Female | -0.364 | 0.124 | -.658 | -.126 |
| 2020 | | | | | |
| | Liberal-conservative | 8.678 | .065 | 8.557 | 8.809 |
| | Economic disapproval | -3.142 | .083 | -3.292 | -2.969 |
| | Personal economic situation (worse) | -.219 | .087 | -.377 | -.059 |
| | Government distrust | -.644 | .048 | -.733 | -.541 |
| | Age | -.278 | .081 | -.412 | -.096 |
| | Education (highest) | -.798 | .092 | -.956 | -.611 |
| | White Non-Hispanic | 1.185 | .054 | 1.075 | 1.283 |
| | Female | .048 | .050 | -.042 | .150 |

**State-level fixed effects are not shown for simplicity and readability purposes. Additional versions of the model tables are available upon request.*

## Endnotes

[i] For a more detailed overview of this approach, please refer to recent works that estimate aggregate electoral scenarios as a by-product of individual-level vote functions (e.g., Camatarri 2022; Luartz et al. 2024).
[ii] Please refer to the ANES website for a more detailed description of variables and survey methodology: https://electionstudies.org/data-center/anes-time-series-cumulative-data-file/
[iii] MCMC is a powerful computational technique used to approximate the posterior distributions of the Bayesian model parameters. It does this by generating a sequence of samples that converge to the true distribution, allowing us to estimate the parameters accurately even in complex models.
[iv] For a comprehensive overview of the advantages of combining traditional logistic regression with Bayesian estimates, including the stabilization of estimates and the reduction of standard errors, please refer to Acquah (2013).
[v] For the record, state-level fixed effects, which are not shown in the table, were found to be largely insignificant, confirming the apparent primacy of individual-level factors for predictions of voting behavior and therefore election results.
[vi] In logistic regression, predicted probabilities are calculated by converting the log-odds—a linear combination of predictors and their coefficients—into probabilities using the logistic function. In Bayesian logistic regression, this process uses the posterior means of the coefficients, incorporating both the predictors' effects and the uncertainty in the estimates.
[vii] In contrast to the conventional 0.5 threshold in logistic regression (Menard 2002), using the mean probability cut-off can yield more balanced and accurate classification outcomes, particularly for minority classes and underrepresented groups in the data, such as Republican supporters in this case (Cramer 2003).