Can Exit Polls Predict Elections? Evaluating the Accuracy of Multilevel Regression and Poststratification Models in Predicting State-Level Vote Share

Joelle Gross
University of Pennsylvania
Master's of Science in Engineering in Data Science *

August 7, 2025

Abstract

With historical exit poll data, U.S. Census microdata, and official election returns, this paper aims to evaluate the effectiveness of multilevel post-stratification (MRP) models in predicting vote share in presidential elections. By applying MRP across five election cycles (2008–2024), this analysis assesses the model's accuracy in both states with and without exit polls. Results show that MRP consistently produces reliable estimates, with no statistically significant accuracy advantage in states where exit polls were conducted. Additional regression analysis reveals strong associations between predicted Democratic vote share and state-level demographic factors such as educational attainment, gender, and race. These findings highlight both the strengths and limitations of MRP in electoral forecasting and underscore its value as a complementary tool for exit polling and election night race calls.

1 Introduction

Early on election night, understanding voter behavior and public sentiment is both critical and difficult, especially with few to no vote returns available. One key tool at our disposal is the exit poll. Conducted by the National Election Pool (a media consortium including NBC, ABC, CNN, and CBS News) and Edison Research, the exit poll surveys voters as they leave polling places on Election Day and, for early voters, via phone interviews. These surveys provide demographic and attitudinal data that often help shape the early narrative of how America voted on election night.

Exit polls are the only survey instrument we have to reliably ask verified voters about their choices before votes are counted. However, the polls come with limitations. They are only conducted in a subset of states during each election year due to their expense, restricting their ability to support nationwide analysis. Moreover, exit polls are subject to selection bias: only voters who opt into the

*In addition to being a MSE in Data Science student, I also work full-time for Penn as the Assistant Director for the Data Science & Analytics minor, and I serve as an Exit Poll Analyst for NBC News Decision Desk.

survey are included, and these individuals tend to skew Democratic. (Notably, demographic data on refusals is also estimated and included in the exit poll, but of course, without vote choice.)

This project seeks to extend the geographic and demographic reach of exit polls and evaluate their broader utility in informing the national election story. Using multilevel regression and poststratification (MRP), a widely used method for generalizing survey data on public opinion, I estimate presidential vote share in all 50 states. By comparing these predictions to actual election outcomes across multiple cycles, this analysis assesses both the performance of MRP and the added value of exit poll data in modeling election outcomes.

2 Background and Prior Research

Surveys are only as reliable as they are representative of the population they are aiming to estimate. A common issue with survey research is the challenge in polling a representative sample from every demographic group. Multilevel regression with poststratification (MRP) has become a cornerstone of modern public opinion and survey research. Originally developed to generate small-area estimates from national surveys, MRP is now widely used when direct data for certain subgroups are limited or unavailable.

One of the first influential applications of MRP in political science was by Park et al. who demonstrated how MRP could be used to estimate state presidential vote from national pre-election surveys by combining multilevel modeling with census-based post-stratification. Since then, MRP has been applied to a wide array of political and social outcomes such as vote choice, policy preferences, and partisanship (Lewis & Jacobsmeier, 2017; Park et al., 2004).

Gelman et al.'s extended MRP's application further by evaluating how it performs across different states and survey conditions, including sparse sample sizes and non-representative surveys. Their work emphasized that MRP's value lies in its ability to "borrow strength" across groups via partial pooling, and to incorporate contextual variables like prior vote share or economic indicators at the geographic level (Gelman et al., 2013).

MRP has also been used specifically in the context of exit polls, particularly in relation to a known "competitor" of the National Election Pool, the AP VoteCast survey (Associated Press, 2024). AP VoteCast, developed by The Associated Press and NORC at the University of Chicago, departs from traditional in-person exit polling by relying on a combination of pre-election and Election Day surveys conducted online and by phone. At the AP, researchers apply MRP to the survey data drawing on demographic estimates from the census's American Community Survey, Current Population Survey, as well as a proprietary voterfile.

3 Data and Methods

3.1 Data Sources

This project relies on historical exit poll data, U.S. Census microdata, and official election returns to generate and evaluate state-level vote share. The analysis focuses on the last five presidential elections: 2024, 2020, 2016, 2012, and 2008. I do not include elections prior to 2008 due to limitations in the American Community Survey (ACS), which first provided 5-year estimates in 2009. The project is limited to presidential elections for several key reasons. Presidential primaries draw a distinct subset of the population that differs significantly from the general electorate, undermining the assumptions we make to conduct census-based modeling. Midterm elections pose additional challenges, as states and localities vote on different candidates, making it difficult to meaningfully extrapolate exit poll data across states. In contrast, general presidential elections offer a consistent baseline: all states vote on the same two candidates at the same time, and turnout can be reasonably assumed to approximate a representative sample of the population surveyed in the American Community Survey. Notably, the fact that the electorate is a representative sample from the census is the fundamental assumption made to conduct this analysis.

The foundation of the analysis is a comprehensive archive of national and state-level exit polls, which I compiled from both the Roper Center for Public Opinion Research and NBC's internal database of exit poll archive files. These files were systematically cleaned and standardized to align variable names and demographic groupings across election years (National Election Poll, 2025).

To represent the demographic makeup of each state in a given election year, the model uses data from the American Community Survey (ACS) 5-Year Public Use Microdata Sample (PUMS), available through the U.S. Census Bureau. These microdata provide a detailed sample of individuals and households, including variables such as age, sex, race, education, and Hispanic origin. The data were recoded to match the demographic categories used in the exit poll data, and poststratification frames were built for each state in each elction year (U.S. Census Bureau, 2024).

Historical election returns and prior vote share data were sourced from the MIT Election Data and Science Lab, which offers cleaned and standardized state-level election results for presidential elections (MIT Election Data and Science Lab, 2024). These returns were used both to validate model predictions and to include as a variable in the model, as historical vote is often the greatest predictor of current vote share in any given election. Together, these data sources were integrated into a hierarchical Bayesian model, estimated using the `rstanarm` package in R to produce state-level Democratic vote share predictions (Kastellec et al., 2019; Kennedy & Gabry, 2020).

4 Modeling Approach

Multilevel regression with poststratification (MRP) operates in two main stages. In the first stage, a multilevel logistic regression model is fit to individual-level survey data. For each individual i, we model the probability of voting for the Democratic candidate based on their demographic characteristics as:

$$\text{logit}(\Pr(y_i = 1)) = \beta_0 + \beta_1 \cdot \text{age}_i + \beta_2 \cdot \text{sex}_i + \beta_3 \cdot \text{race}_i + \beta_4 \cdot \text{educ}_i + \beta_5 \cdot \text{prevvote}_{s(i)} + u_{s(i)}$$

where y_i is a binary indicator for a Democratic vote, $\text{prevvote}_{s_{(i)}}$ is the prior Democratic vote share in state s, and $u_{s(i)} \sim \mathcal{N}(0, \sigma_s^2)$ is a state-level random effect capturing unobserved heterogeneity across states.

The full model is estimated within a Bayesian framework, which incorporates prior beliefs about the parameters and updates them based on the observed data, primarily from the exit poll and Census microdata. Estimation is performed using the rstanarm package in R (Goodrich et al., 2024).

The power of MRP lies in its use of weakly informative priors: they provide enough constraint to regularize estimates in small or sparse subgroups, while remaining flexible enough to let the data speak when sufficient information is available.

In the second stage, the model's predicted probability of Democratic vote is computed for each post-stratification cell j (defined by combinations of demographic and geographic features). An example of a cell would be a micro-group such as, "Black women aged 30-44 with a college degree in Massachusetts" yielding:

$$\hat{p}_j = \Pr(y = 1 \mid \text{cell } j)$$

These probabilities are then aggregated using known population frequencies N_j from the ACS PUMS dataset for each state:

$$\hat{\theta}_s = \frac{\sum_{j \in s} N_j \hat{p}_j}{\sum_{j \in s} N_j}$$

where $\hat{\theta}_s$ represents the estimated Democratic vote share in state s. This poststratification step ensures that final state-level predictions reflect the true demographic composition of the electorate in each state (Gao, 2022). Additionally, prior vote share is included as a co-variate. The model is fit separately for each election year using historical exit poll data, and predictions are generated for all 50 states, regardless of whether an exit poll was conducted in that state. (University of Virginia Library, StatLab, 2025)

4.1 Evaluation Metrics

Since ground truth data for actual vote share is available at the state level for each election year, I compare each state's predicted Democratic vote share to its true value. As a first pass, I assess whether the model correctly projects the winner in each state. While this binary classification is crucial for understanding electoral outcomes, it may overepresent near-miss cases. For example, a predicted vote share of 49.8% when the true value is 50.01% results in an incorrect winner classification, despite being off by only a fraction of a percentage point.

I collect the mean absolute error (MAE) for each year, which is the average absolute difference between the predicted and actual Democratic vote share across all states. Like MAE, RMSE captures the average prediction error, but penalizes larger errors more heavily by squaring the differences. Finally, R^2 measures the proportion of variation in actual vote share that is accounted for by the model's predictions. An R^2 value close to 1 indicates that the model explains most of the variance in the observed data, suggesting strong predictive accuracy.

Together, these metrics offer a comprehensive view of the model's performance. Beyond raw error estimates, I also explore potential biases in the models. Are there particular years where the model performs better or worse? For instance, in 2024, we only conducted exit polls in a limited set of states. Does the presence of an exit poll significantly affect prediction accuracy? Are the estimates systematically biased toward certain demographic groups that may be more likely to participate in exit polling?

One important methodological consideration is that the number of states with exit polls has declined over time, which could either hinder or enhance model performance. In 2008, exit polls were conducted in all 51 jurisdictions (50 states plus DC). That number dropped to 31 in 2012, 28 in 2016, 25 in 2020, and just 10 in 2024.

5 Results

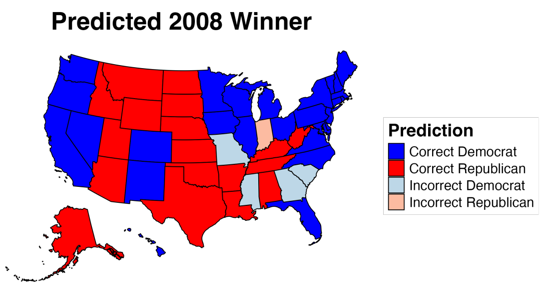
5.1 Model Performance Summary

Table 1 show the overall performance of the MRP model for each of the five presidential cycles. The results show a clear improvement in model performance between 2008 and 2016, with MAE decreasing from 0.050 to 0.033 and R^2 increasing from 0.598 to 0.844. Although performance dipped slightly in 2020, with a higher MAE (0.046) and a lower R^2 (0.744), the model rebounded in 2024, achieving its best R^2 value (0.856) and maintaining a low error rate. Overall, the model demonstrates strong performance overall, with some notable front runners (like 2024) and dips in performance (2008). The consistently low MAE and RMSE values, in 2012, 2016 and 2024, indicate that the model is capable of producing reliable estimates at the state level.

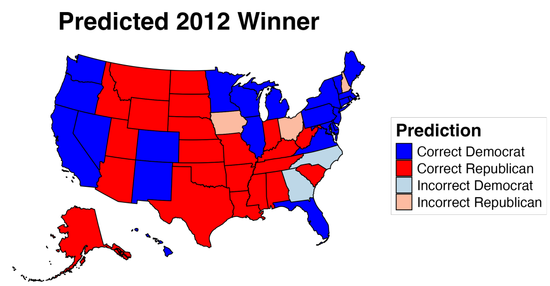
Table 1: MRP Evaluation Metrics by Year

Year	MAE	RMSE	R^2
2008	0.050	0.063	0.598
2012	0.042	0.049	0.840
2016	0.033	0.042	0.844
2020	0.046	0.057	0.744
2024	0.046	0.052	0.856

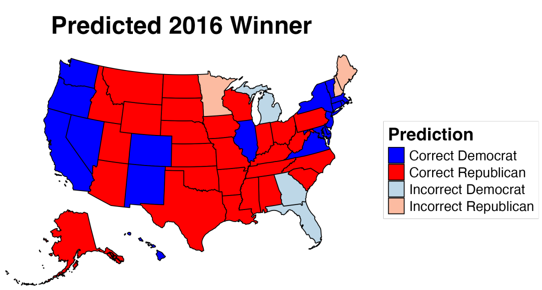
Figure 1 displays the model's predicted winners by state across five presidential election years. States are red or blue to indicate whether the model predicted a Republican or Democratic winner, respectively, with lighter red and blue used to highlight incorrect predictions. Overall, the model performs well across cycles, with few errors in 2012, 2016 and 2024. In 2016 and 2008, the model over-predicted Democratic support in several Republican-leaning states, while 2020 exhibits a more balanced error pattern. Notably, in 2020 and 2008, the model includes several false Democratic predictions in the South. At first glance, the model appears to preference southern states as Democratic and northern states as Republican. Considering the demographics of these regions, we might assume that the model is finding a strong correlation between states with a higher Black population and democratic support. For example, Minnesota is a state that has voted for the Democratic candidate for the last 13 consecutive elections and has a lower Black population than some of the southern states, yet the MRP model predicts it incorrectly as red in 2016. (Stephen Swanson, 2024)



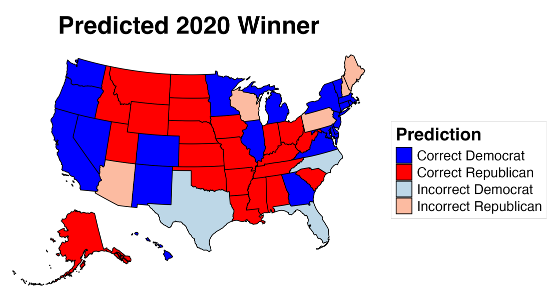
(a) 2008



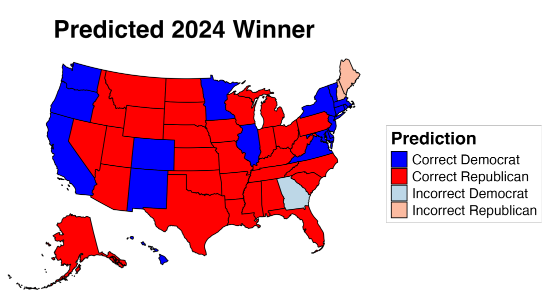
(b) 2012



(c) 2016



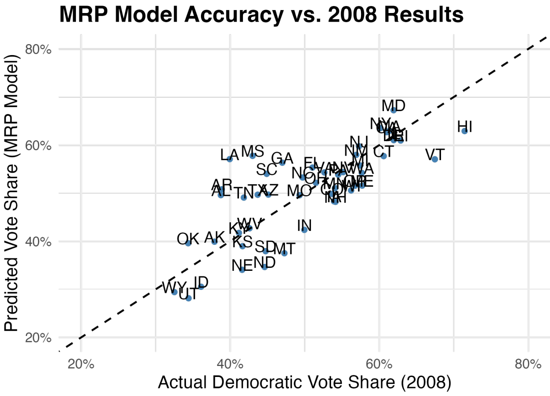
(d) 2020



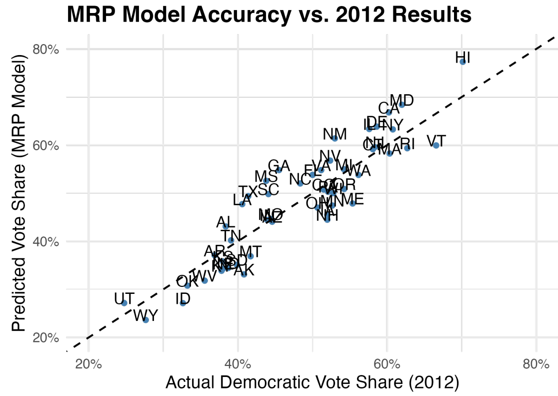
(e) 2024

Figure 1: Predicted winners by state using MRP models across five presidential elections. Lighter shades indicate incorrect predictions. Light red is a false Republican state, light blue is a false Democrat state.

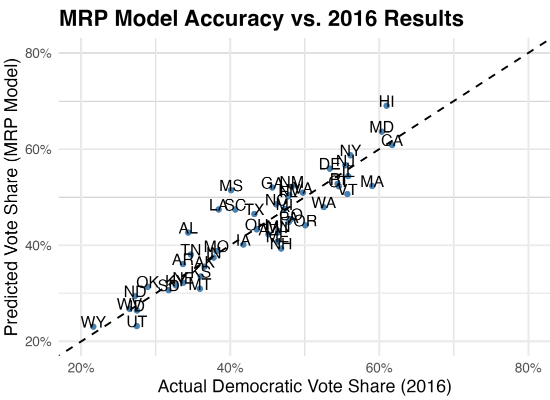
In Figure 2, we gain a more granular view of the predicted vote shares. Notably, certain states, such as Maryland, Hawaii, Mississippi, and Georgia, are consistently overestimated in terms of Democratic vote share.



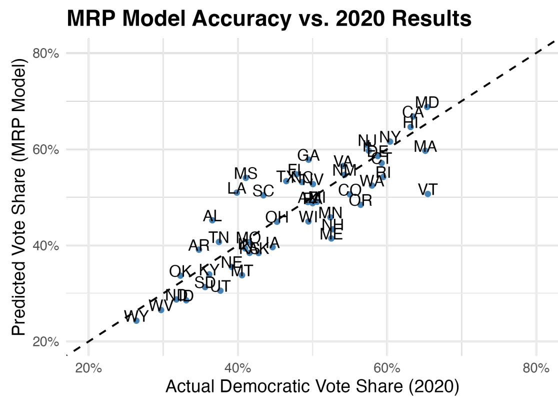
(a) 2008



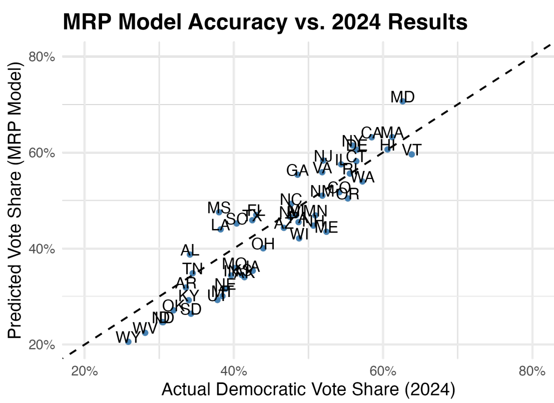
(b) 2012



(c) 2016



(d) 2020



(e) 2024

Figure 2: Democratic vote share estimate by state compared to actual vote share estimate in that year with a 45 degree line indicating perfect prediction.

Conversely, states such as Vermont, New Hampshire, and Minnesota are frequently underestimated in terms of Democratic vote share, leading the model to incorrectly predict Republican victories sometimes. These plots help identify where the MRP model tends to deviate from actual outcomes, distinguishing between years and states with high variance in prediction accuracy (2008, 2020) and those where the estimates are consistently close (2016, 2024).

5.2 Regression Analysis

In this section, I move beyond evaluating overall model performance and begin exploring the underlying dynamics that may be driving variation in performance. I examine several key questions: Do statistically significant differences in error persist across election years? Are there systematic differences in performance between states with and without exit polls? And how do demographic characteristics of states influence the model's vote share predictions?

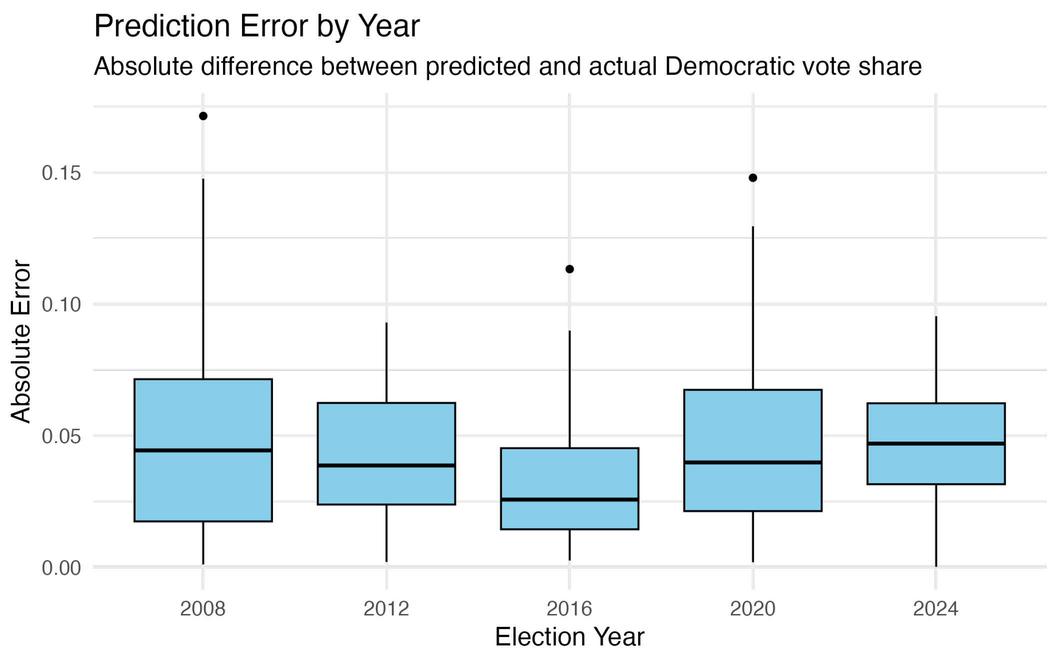


Figure 3: Distribution of absolute prediction error by election year. The MRP model performs consistently across cycles, with 2024 exhibiting the lowest median error and least spread.

Figure 3 displays the distribution of absolute prediction errors for the MRP model across five presidential elections. The median error remains relatively stable at around 4.5 percentage points. A linear regression model was used to assess whether absolute error significantly varied by year. Using 2020 as the reference year, only the 2016 election shows a statistically significant difference, with errors approximately 1.3 percentage points lower on average ($p = 0.036$). No other year exhibited a statistically significant difference from the baseline (2020). These results suggest that while model accuracy is largely consistent across years, predictions in 2016 were notably more precise.

An assumption one might make is that states with exit polls would far out-perform states without exit polls. To test this, I ran a regression with exit poll as a binary variable. To see where exit polls were conducted in each election year, see table 3 in the Appendix.

Table 2: P-values for "Exit Poll State" Effect on Error by Year

Year	p-value
2008	NA
2012	0.2484
2016	0.2758
2020	0.4396
2024	0.0895

The results, shown in Table 2, do not support a strong or consistent effect of exit poll presence on model accuracy. In most years, the p-values are well above conventional significance thresholds, suggesting that the presence of an exit poll in a state did not meaningfully reduce prediction error. Notably, in 2008, the regression term was dropped because exit polls were conducted in all 50 states.

This finding reinforces the robustness and reliability of the MRP estimates. Since exit polls do not appear to systematically reduce error in the states where they are conducted, the model demonstrates its ability to generate accurate and meaningful estimates even in states without survey data. This underscores the value of MRP as a tool as there could be fewer states where we conduct exit polls in the future.

5.3 Demographic Analysis

To better understand how demographic characteristics relate to predicted Democratic vote share, I regressed state-level predictions on three variables: percent with a college degree, percent female, and percent Black. For each, I ran separate linear regressions across all five election years, visualized using scatterplots with state shaded by region (North, West, South, East), since earlier figures highlighed that errors were clustered in some parts of the country more than others.

As shown in Figure 4, there is a consistently positive relationship between the share of adults with a college degree and predicted Democratic vote share across all years, with no notable regional pattern. The slope increases over time, particularly in 2020 and 2024. States that have higher educational attainment exhibits markedly higher Democratic support. The p-values across all years are statistically significant (all $p < 0.05$), suggesting that those with college degrees may be more likely to be surveyed by the exit poll and vote Democrat. While these individuals are historically more likely to vote Democrat, the model may be overestimating vote share for states that have a higher percent of the population with a college degree. Figure 5 shows the relationship between percent female and predicted Democratic vote share, which is consistently positive. States where democratic vote share dips are notably in the West. Again, while women are more likely to vote Democrat, they also may be more likely to take the exit poll, explaining this trend further.

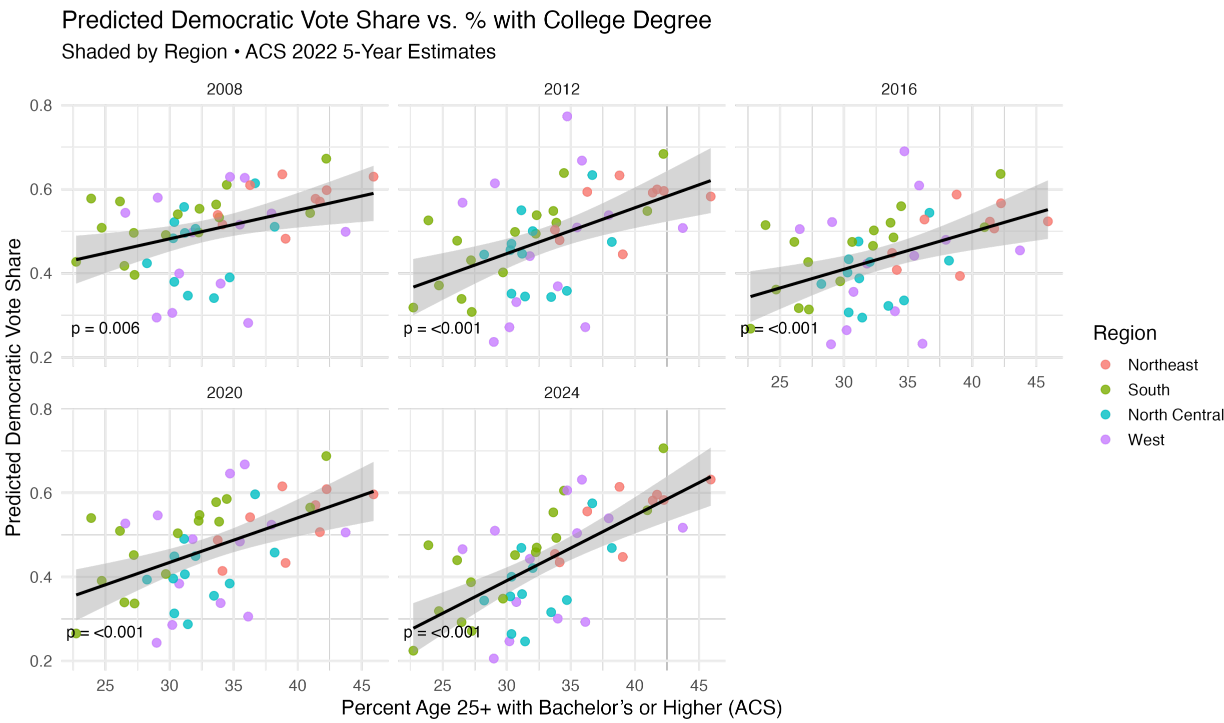


Figure 4: Predicted Democratic Vote Share vs. Percent with College Degree (Age 25+).

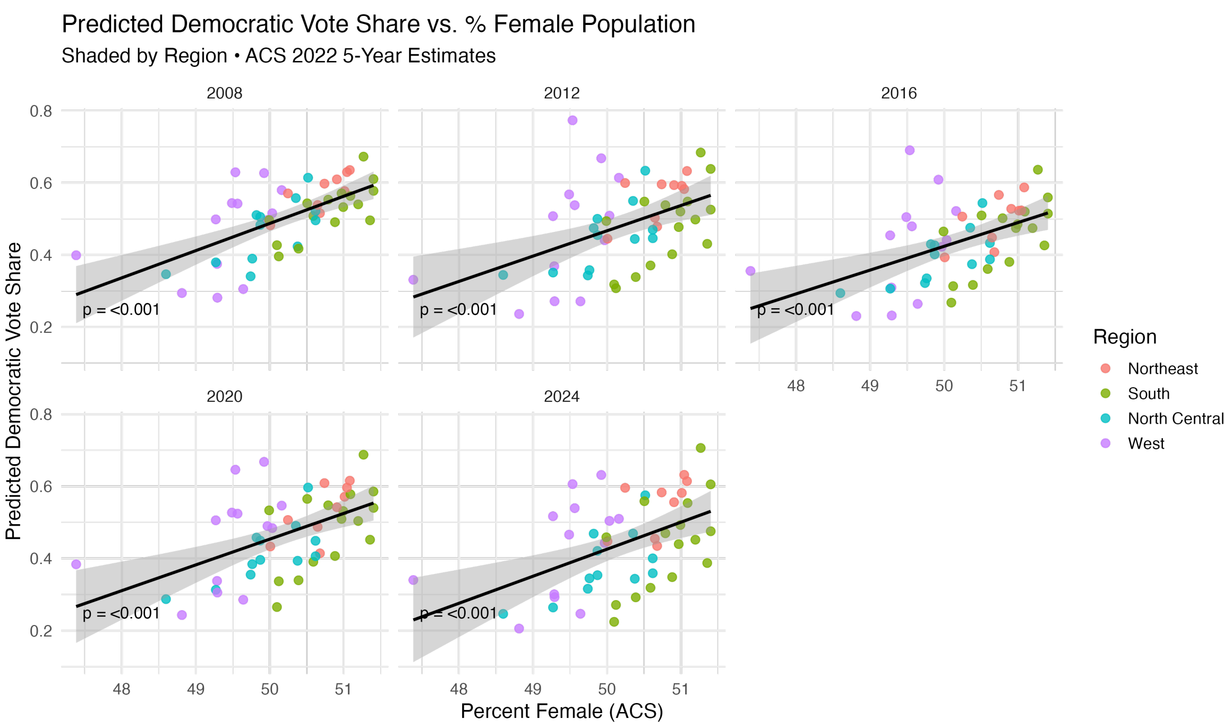


Figure 5: Predicted Democratic Vote Share vs. Percent Female

As shown in Figure 6, the percentage of Black residents in a state is also positively associated with predicted Democratic vote share. This relationship is somewhat noisier across years compared to sex and educational attainment, but a strong regional pattern emerges. Southern states, which have higher percentages of Black residents, are often predicted to support Democratic candidates more often than they have. This may help explain some of the misclassifications observed in Figure 1, where the model frequently overestimated Democratic performance in the South. The model's emphasis on racial composition, particularly the share of Black voters, could be contributing to this persistent bias.

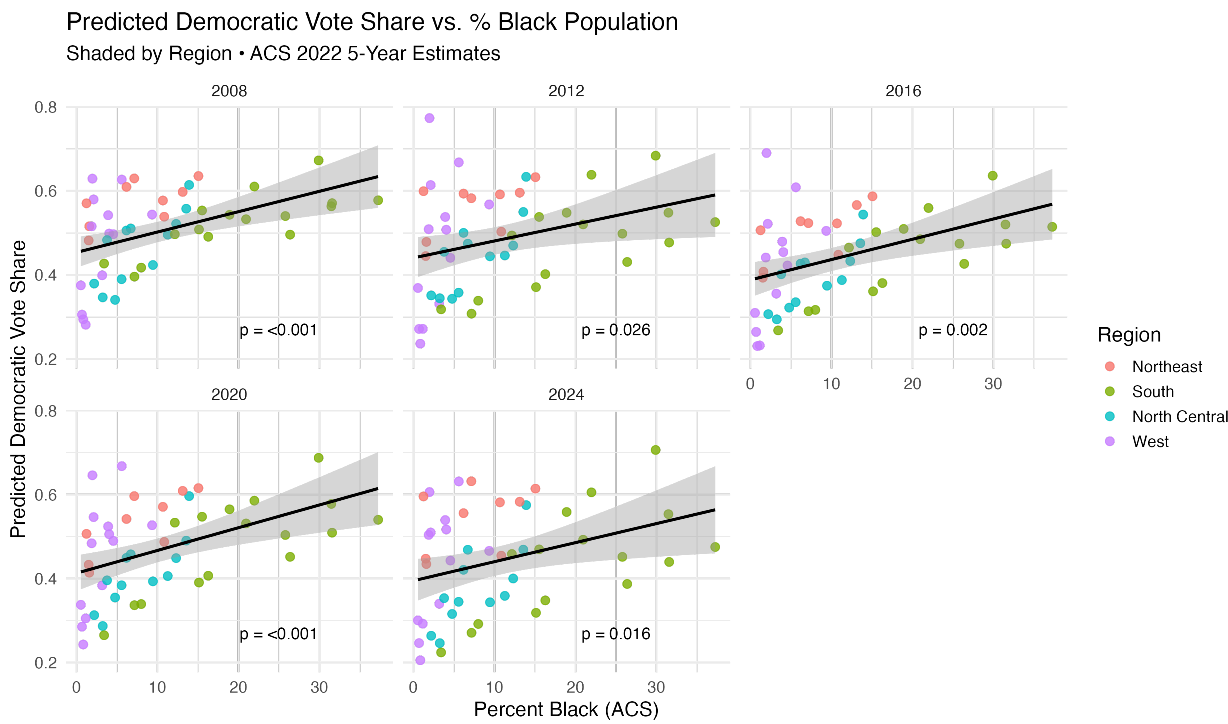


Figure 6: Predicted Democratic Vote Share vs. Percent Black

6 Conclusion

This project set out to evaluate the performance and utility of multilevel regression and post-stratification models in predicting state-level presidential vote share with exit poll data. Across five presidential elections, the results demonstrate that MRP is robust and flexible, capable of producing moderately accurate estimates even in states where no direct survey data were available.

The model performed consistently well across cycles, with particularly strong results in 2016 and 2024. Notably, the analysis found no statistically significant difference in model accuracy between states with and without an exit poll, suggesting that well-specified demographic and historical covariates can be sufficient to generate useful predictions in the absence of survey data.

Demographic regression analyses further revealed statistically significant relationships between predicted Democratic vote share and state-level characteristics such as educational attainment, gender composition, and racial makeup. While these patterns reflect well-documented trends in American electoral behavior, they also point to areas where the model may overemphasize certain features, particularly Black population share, contributing to consistent prediction of Democratic support in the South.

Together, these findings highlight both the strengths and limitations of MRP in election forecasting. As the future of the exit poll remains uncertain, MRP provides a compelling framework for extending exit poll analysis. For the NBC News Decision Desk in particular, this analysis suggests that the exit poll remains a vital resource – not only as an editorial tool in the early hours of election night– but as a foundation for improving race calls and offering a timely snapshot of the American electorate.

7 Bonus: Exit Poll Dashboard

A major focus of my summer work was compiling our exit poll archives into a clean, readable, and centralized format. After working on the exit poll team both on election night and in the weeks leading up to it, I saw firsthand the value of historical exit poll data. Not only would a repository of exit poll data significantly ease the burden on the exit poll team, especially during the high-pressure hours of election night, but it would also address a broader gap: there is currently no comprehensive public record of historical exit poll data accessible to journalists, researchers, or the general public.

Despite its limitations (as with any survey), I believe the exit poll is both a critical tool for real-time election analysis and an important public good. The archive represents a small but meaningful piece of American history, a snapshot of how people across the country felt on some of the most consequential days in our democracy.

To address this, I developed a public-facing archive of exit poll results, accessible through an interactive dashboard. My long-term vision includes features like dynamic cross-tabulation by demographic groups and the ability to track trends over time. If the Decision Desk and Exit Poll team sees value in this tool for the 2025 general election and the 2026 midterms, I would be eager to continue developing it into a resource that serves both our internal needs and the broader public interest.

View the dashboard here. (a note: I am using Render's free plan to host the dashboard, and therefore it might be a little bit slow to run). Additionally, here is a list of all of the elections for which I have collected exit poll data so far.

8 Replication Code

Replication code and data for the analysis in this paper can be found on github. The replication code and data for the exit poll dashboard can be found on this github.

9 Acknowledgments

I would like to thank my incredible advisors and mentors – Stephanie Perry (Manager of Exit Polling at NBC News), Dr. Marc Meredith (Professor of Political Science at the University of

Pennsylvania), and Dr. Sharath Chandra Guntuku (Associate Professor of Computer and Information Science at the University of Pennsylvania)– for their support throughout the summer in helping me turn this project into something I'm really proud of.

Stephanie provided invaluable guidance on the inner workings of the exit poll and offered a thoughtful journalistic perspective on the practical use and design of the dashboard. Dr. Meredith's expertise in statistics and elections helped me bridge my interests in MRP and election forecasting to shape the core of this analysis. Over the past year, Dr. Guntuku has answered many emails and Zoom calls, guiding me through the evolution of this project and helping me achieve a goal I have had since day one of starting the DATs program: designing and executing an independent research project as part of my master's degree.

I am immensely grateful for the generosity of time, thoughtful feedback, and replies to many (many) questions. Your mentorship has meant the world to me.

10 Appendix

Table 3: States with Exit Polls by Year

Year	Where exit polls where conducted
2008	AK, AL, AR, AZ, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, SD, TN, TX, UT, VA, VT, WA, WI, WV, WY
2012	AL, AZ, CA, CO, CT, FL, IA, IL, IN, KS, MA, MD, ME, MI, MN, MO, MS, MT, NC, NH, NJ, NM, NV, NY, OH, OR, PA, VA, VT, WA, WI
2016	AZ, CA, CO, FL, GA, IA, IL, IN, KY, ME, MI, MN, MO, NC, NH, NJ, NM, NV, NY, OH, OR, PA, SC, TX, UT, VA, WA, WI
2020	AL, AZ, CA, CO, DC, FL, GA, IA, KY, ME, MI, MN, MT, NC, NH, NV, NY, OH, OR, PA, SC, TX, VA, WA, WI
2024	AZ, FL, GA, MI, NC, NV, OH, PA, TX, WI

Every year, a national exit poll is conducted sampling voters from all states in addition to state-level exit polls. This table shows where state-level exit polls occurred in the respective year.

References

Associated Press. (2024). Ap votecast — a modern approach to election research [Accessed August 2, 2025].

Gao, Y. (2022). *Advances in multilevel regression and poststratification for non-representative data modeling* [PhD thesis]. University of Toronto [Graduate Department of Statistical Sciences].

Gelman, A., Lax, J. R., & Phillips, J. H. (2013). How should we estimate public opinion in the states? *American Journal of Political Science*, *57*(3), 538–550.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2024). Rstanarm: Bayesian applied regression modeling via stan [R package version 2.32.1]. https://mc-stan.org/rstanarm/

Kastellec, J. P., Lax, J. R., & Phillips, J. H. (2019). Estimating state public opinion with multilevel regression and poststratification using r [Accessed August 2, 2025].

Kennedy, L., & Gabry, J. (2020). Mrp with rstanarm [Accessed August 2, 2025].

Lewis, D. C., & Jacobsmeier, M. L. (2017). Evaluating policy representation with dynamic mrp estimates: Direct democracy and same-sex relationship policies in the united states. *State Politics & Policy Quarterly*, *17*(4), 441–464. https://doi.org/10.1177/1532440017739423

MIT Election Data and Science Lab. (2024). U.s. election returns data [Accessed: 2025-08-01]. https://electionlab.mit.edu/data

National Election Poll. (2025). Exit polls — elections and presidents [Accessed July, 2025].

Park, D. K., Gelman, A., Bafumi, J., & Katz, J. N. (2004). Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis*, *12*(4), 375–385.

Stephen Swanson, E. M. (2024, November). *Minnesota extends blue streak for 13th consecutive democratic presidential election*. https://www.cbsnews.com/minnesota/news/minnesota-democratic-presidential-election-history/

University of Virginia Library, StatLab. (2025). Getting started with multilevel regression and poststratification [Accessed August 2 ,2025]. *StatLab Articles, UVA Library*. https://library.virginia.edu/data/articles/getting-started-multilevel-regression-and-poststratification

U.S. Census Bureau. (2024). American community survey (acs) 5-year pums api [Accessed: 2025-08-01]. https://www.census.gov/data/developers/data-sets/census-microdata-api/acs-5y-pums.html