University of California, Berkeley

Data 102: Data, Inference, and Decisions

May 12, 2025

Exploring Endorsements in the Primary Election

Group: 24

Contributors: Fiona Chang, Liam Dupeyron, Emily Gabion, Joseph Zhai

Project Repository: GitHub

Word Count: 2984

Contents

1 Introduction

Endorsements from political parties, advocacy groups, and public figures have long played a role in shaping electoral outcomes. These endorsements can act as informational shortcuts, helping voters—especially those less engaged with down-ballot races—navigate a crowded and complex field of candidates. By signaling credibility, alignment with key values, or simply increasing name recognition, endorsements have the potential to elevate candidates who might otherwise be overlooked.

This influence may be particularly critical in primary elections, where turnout tends to be lower and the pool of candidates is often less familiar to the public. In such contexts, endorsements may not only reflect a candidate's viability but also amplify it by directing media attention, swaying undecided voters, and attracting institutional support.

Our project aims to examine the relationship between endorsements and primary election outcomes in the 2022 U.S. midterm elections. Specifically, we ask:

- What effect do endorsements from political action committees (PACs), political figures, and advocacy groups have on primary election outcomes?
- Does a candidate's party affiliation amplify or diminish the effect of endorsements?

To answer these questions, we combine exploratory data analysis with causal inference and Bayesian hierarchical modeling, allowing us to estimate both average effects and endorsement-specific heterogeneity across parties.

2 Background and Related Work

Scholars have long studied the influence of political endorsements on electoral behavior. Endorsements serve as cues for voters who may lack detailed policy knowledge or familiarity with all candidates. Studies have shown that endorsements from trusted elites or well-known organizations can significantly increase voter support, especially in primaries where party labels alone are less informative (Boudreau & MacKenzie, 2014). Other work has demonstrated that in election cycles, endorsements can be particularly persuasive when they come from media sources, particularly those whose typical partisan lean differs from the endorsement itself (Knight & Chiang, 2011).

Our project builds on this literature by comparing the causal impact of endorsements across parties and using Bayesian techniques to model endorser-specific variation.

3 Data Overview

We use publicly available data compiled by FiveThirtyEight (FiveThirtyEight, 2022), which includes detailed information on Democratic and Republican primary candidates across the United States in 2022. This data set aggregates political, demographic, and electoral information for each candidate, allowing us to examine how endorsements correlate with primary performance.

We work with two separate data sets—one for each major political party. Each row corresponds to an individual candidate in a specific race and includes variables such as:

- Candidate characteristics (e.g., gender, race, veteran status, incumbency),
- Endorsements (binary indicators from political figures, PACs, and interest groups),
- Primary outcomes (e.g., vote share, win/loss),
- Metadata (e.g., office sought, state, election type).

These data contain a mix of Boolean, numeric, string, and date fields. Our analysis focuses on endorsement variables and vote outcomes, while adjusting for potential confounders such as incumbency and state-level effects.

Before analysis, we removed unopposed candidates (to avoid 100% vote shares), and filtering incomplete or inconsistent entries. We also engineered a key binary treatment variable (`endorsed_any`) to capture whether a candidate received at least one major endorsement, and created aggregated measures such as a "Net Endorsement Score" to summarize endorsement patterns.

This dataset provides a strong foundation for investigating both the general and nuanced effects of endorsements in U.S. primary elections.

4 Exploratory Data Analysis

Our exploratory data analysis (EDA) examines the relationships between both quantitative and qualitative variables to inform our modeling approach. We began removing entries with extreme values (e.g., 100% primary results) to ensure meaningful comparisons.

We focused on the impact of political endorsements on primary election outcomes. For each party, we identified relevant endorsement variables—such as support from EMILY's List, Biden, and Sanders for Democrats, and Trump, the NRA, and Susan B. Anthony List for Republicans.

The bar plot below (Figure 1) shows the proportion of Democratic primary candidates who won broken down by whether they received a specific type of endorsement (e.g., from Biden, Emily's List, Justice Democrats). The "Yes" and "No" groups allow us to visually compare win rates across endorsement categories.

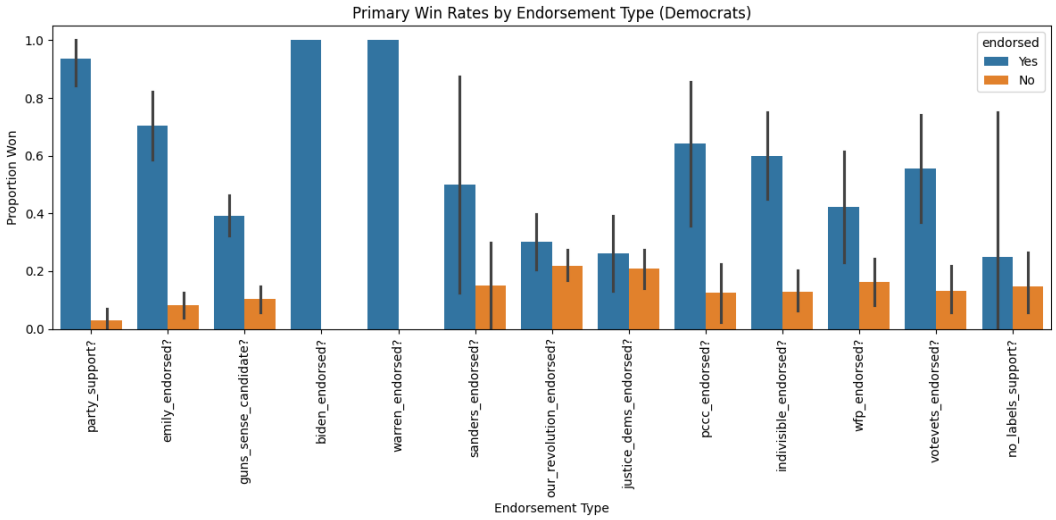


Figure 1: Primary Win by Endorsement Type (Democrats)

Candidates endorsed by Biden, Warren, and with party support had near-perfect win rates. In contrast, candidates endorsed by Our Revolution or Justice Democrats had significantly lower win rates, similar to or only slightly above those not endorsed at all.

On the other hand, we look at the endorsements for Republicans in Figure 2.

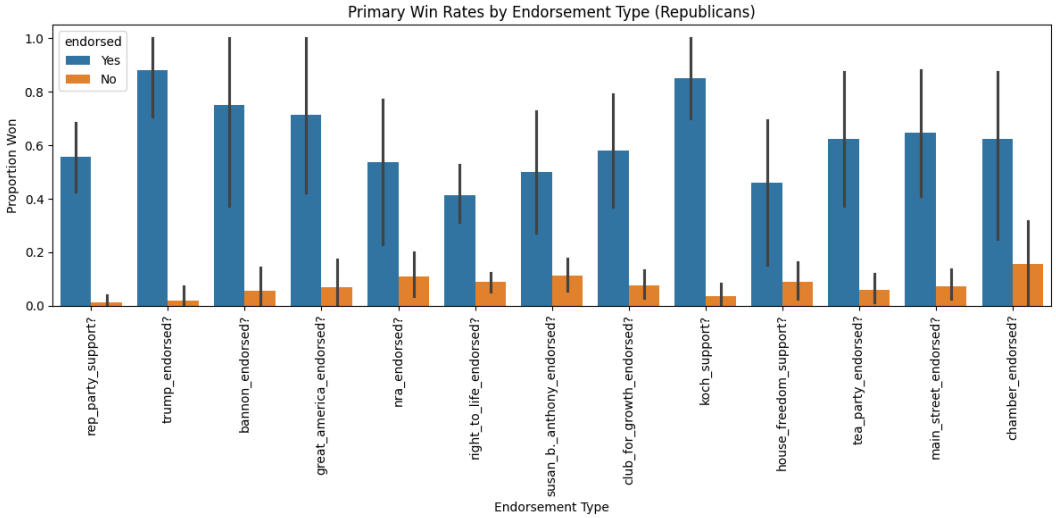


Figure 2: Primary Win by Endorsement Type (Republicans)

The bar plot (Figure 2) displays the proportion of Republican primary candidates who won, broken down by endorsement status. This visualization addresses the first research question by revealing which republican endorsements are most strongly associated with electoral success: Candidates endorsed by Trump, Koch, and Club of Growth showed the biggest win rates.

It suggests that endorsements from mainstream or establishment figures may be more influential in primaries than endorsements from extreme organizations. It also motivates a potential answer to our second research question: the impact of endorsements appears to depend on party dynamics. For example, centrist or party-aligned endorsements seem to carry more weight.

We also look at the number of total endorsements given in Figure 3.

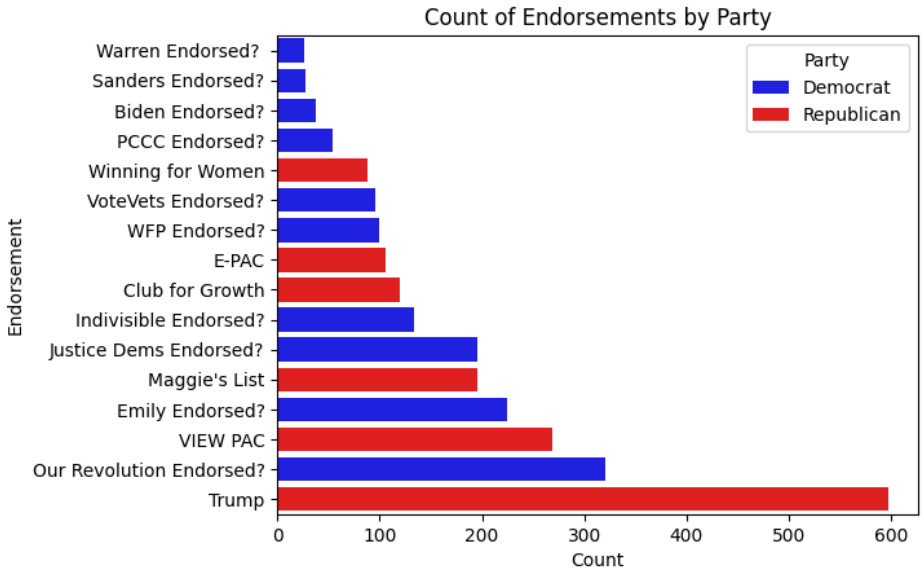
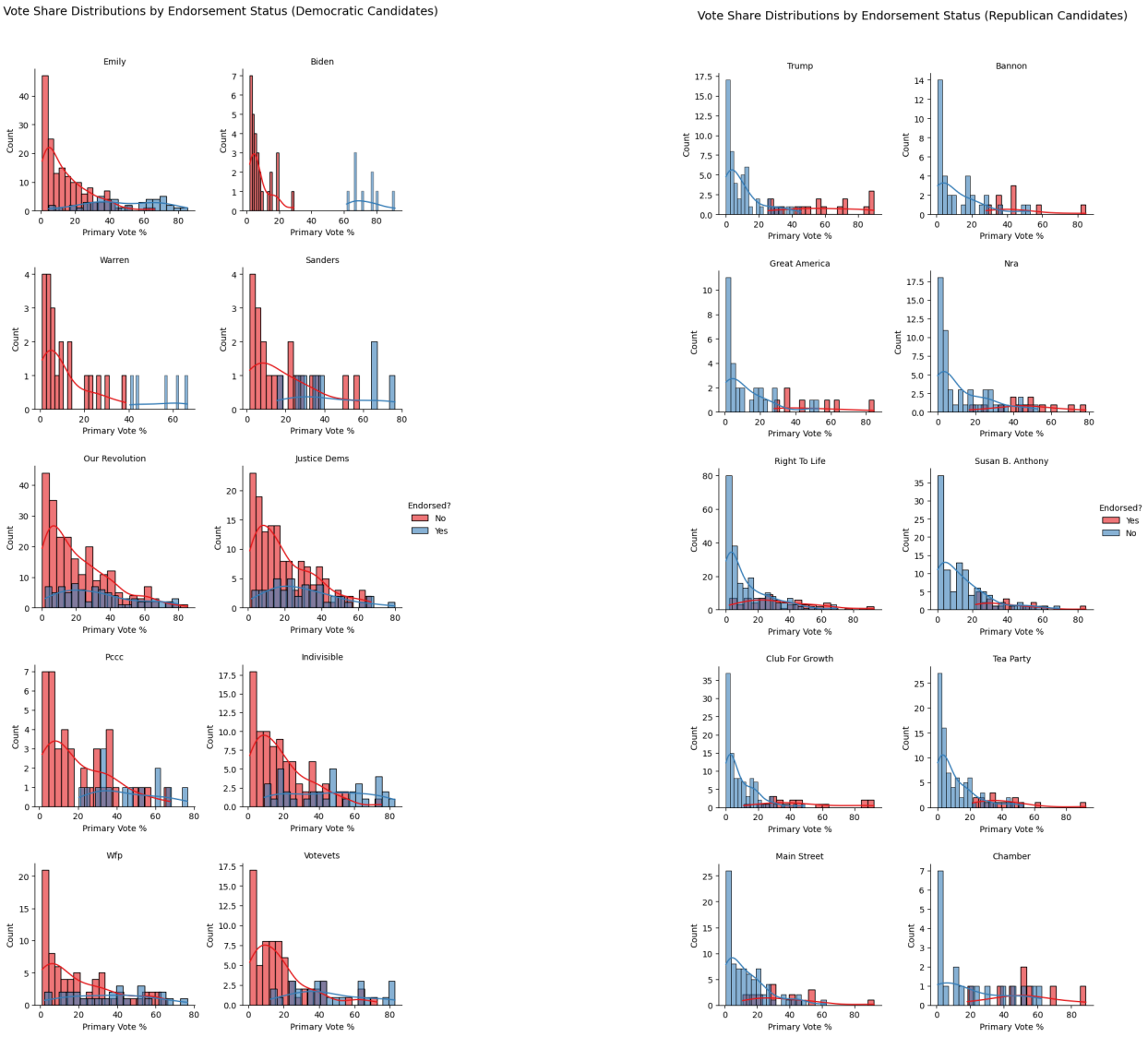


Figure 3: Endorsement Count by Party

We observed that the largest endorser, regardless of party, is Donald Trump. Republicans have more endorsements overall but this is largely attributed to Donald Trump whereas the largest endorser for Democrats is Our Revolution, a political action committee.

By identifying which endorsements are most common and for which party, we gain a deeper understanding of different parties rely on endorsements as a tatic for political signaling. It also allows us to investigate questions such as : Do candidates with Trump's endorsement win more often than those endorsed by PACs? Or is the effect size of an endorsement stronger in one party vs another?

We explored this further by looking at voting distributions by endorsement status (Figure 4).



(a) Democratic Party

(b) Republican Party

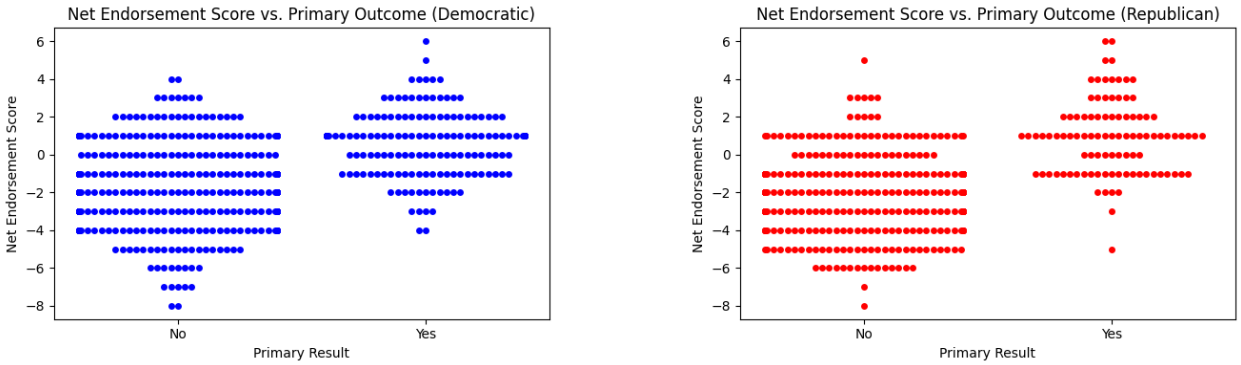
Figure 4: Vote Share Distributions

We visualized the distribution of primary vote percentages for both Democratic and Republican candidates grouped by endorsement status for major endorsers. Each histogram illustrates the distribution of primary vote percentages for whether the candidates are endorsed by a particular figure/organization or not. This helps us see whether candidates endorsed by a figure/organization will tend to receive more/less/similar votes. For example, Biden, Sanders, and Warren endorsed candidates appear to be right skewed compared to candidates not being endorsed by them. Other endorsers like Our Revolution show more overlap between the two groups, suggesting that some endorsements may correlate with better electoral performance—driving our investigation using casual inference.

Lastly, we also created a "Net Endorsement Score" which is calculated by

Net Endorsement Score = # of Endorsements − # of non/anti-Endorsements

We visualize the Net Endorsement Score versus the election outcome in Figure 5.



(a) Democratic Party

(b) Republican Party

Figure 5: Net Endorsement Scores

Endorsements appear to matter slightly more for Democratic candidates: very negative net endorsement scores are rare among Democratic winners but common among Republican winners. This suggests that Democratic primaries are somewhat more influenced by endorsement networks than Republican primaries.

Overall, this analysis helped us identify which variables (e.g., endorsement presence, party affiliation, and candidate outcome) might serve as useful predictors or confounders in subsequent inferential and predictive modeling.

5 Research Question 1: Causal Inference

Research Question: Does receiving at least one major endorsement causally increase a candidate's primary-election vote share?

5.1 Methods

We estimate the effect of major endorsements on primary-election performance, analyzing Democratic and Republican candidates separately. The treatment variable is `endorsed_any`, indicating whether a candidate received at least one major endorsement. The outcome is `primary%`, representing their share of the primary vote.

We control for pre-treatment confounders including:

- Incumbency/experience (`elected_official?`),
- District partisan lean (Democrats only),

- Candidate demographics: race, veteran status, LGBTQ status (Democrats only),
- Office type (House, Senate, Governor),
- Race type (regular vs. special; Republicans only),
- Primary calendar timing (month),
- State fixed effects.

We estimate the Average Treatment Effect on the Treated (ATT) using ordinary least squares (OLS) with HC3 robust standard errors:

$$\texttt{primary\_\%} = \beta_0 + \beta_1 \cdot \texttt{endorsed\_any} + \beta_2 \cdot \texttt{confounders} + \texttt{state effects} + \varepsilon$$

5.2 Results

The regression results indicated that candidates who received at least one major endorsement gained significantly higher vote shares compared to their non-endorsed counterparts. This effect was observed in both parties, though somewhat stronger among Democrats.

5.2.1 Democratic Candidates

We fit an OLS regression with HC3 robust standard errors to estimate the causal effect of endorsements on Democratic candidates' primary vote share. The model included 607 observations and 40 covariates, including state fixed effects.

The estimated coefficient for `endorsed_any` was statistically significant and substantial:

- $\hat{\beta}_{\texttt{endorsed\_any}} = 15.37$ (SE = 1.82, $p < 0.001$)
- $R^2 = 0.341$, Adjusted $R^2 = 0.295$

This indicates that, all else equal, receiving at least one major endorsement is associated with a 15.4 percentage point increase in primary vote share among Democratic candidates. Other significant predictors included incumbency (`experienced`, $\hat{\beta} = 11.29$) and district partisan lean.

5.2.2 Republican Candidates

We estimated the causal effect of receiving a major endorsement on Republican primary performance using OLS with HC3 robust standard errors. The model included 675 candidates and 48 predictors, including race type and state fixed effects.

The estimated coefficient for `endorsed_any` was highly significant:

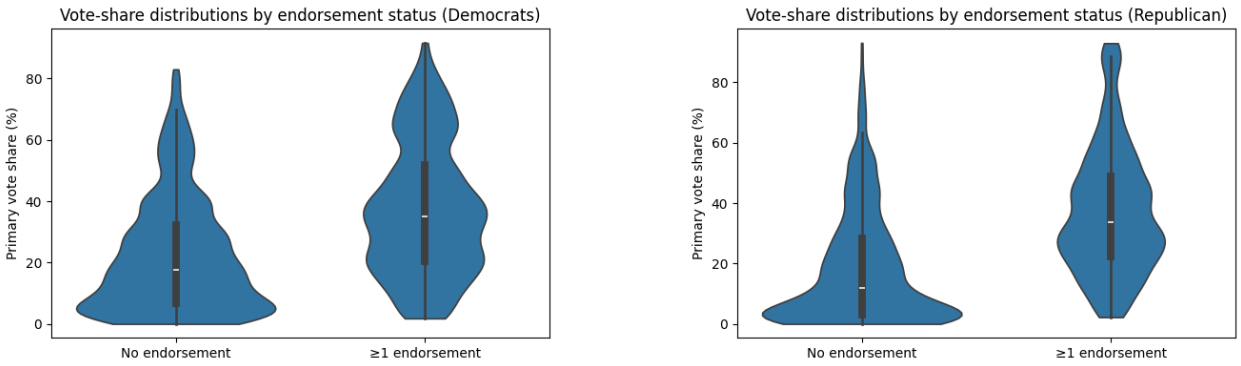
- $\hat{\beta}_{\texttt{endorsed\_any}} = 20.84$ (SE = 2.16, $p < 0.001$)
- $R^2 = 0.300$, Adjusted $R^2 = 0.247$

Our analysis confirms a strong and statistically significant relationship between endorsements and Republican primary vote share. On average, endorsed candidates received 20.8 more percentage points—an effect even larger than that observed among Democrats.

This result reinforces the broader inference that endorsements are not only influential signals but potentially decisive assets in primary campaigns. In Republican races, this may reflect the centralizing influence of key endorsers—particularly high-profile figures—whose support may consolidate early momentum or sway partisan voter bases.

5.3 Discussion

To visually support these findings, we plotted the distribution of primary vote shares for endorsed versus non-endorsed candidates using violin plots (Figure 6). The distributions show clear rightward shifts for endorsed candidates, suggesting both higher medians and fewer extremely low vote shares.

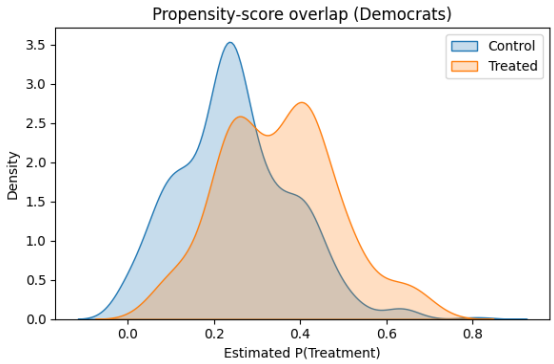


(a) Democratic Party

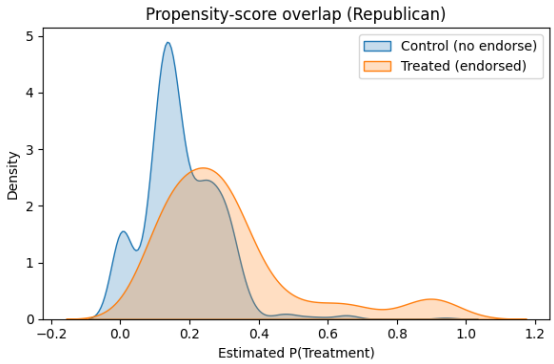
(b) Republican Party

Figure 6: Vote-share Distributions

Additionally, we plotted the estimated propensity score distributions before matching to assess overlap between treated and control groups (Figure 7). These plots confirmed sufficient covariate overlap, supporting the validity of ATT estimation.



(a) Democratic Party



(b) Republican Party

Figure 7: Propensity-score overlaps

We also observed some structural variation: special election status correlated with significantly lower vote shares, possibly due to lower turnout or different campaign dynamics.

5.3.1 Limitations

We caution that causal interpretation depends on assumptions such as unconfoundedness and SUTVA. While we adjusted for a wide array of observed confounders, it's possible that unobserved variables (e.g., candidate charisma, insider support) still bias the estimate. Nevertheless, strong overlap in propensity scores and robust model diagnostics support the claim that endorsements play a causal role in shaping electoral outcomes in Republican primaries.

5.3.2 Implications and Future Work

These findings have important implications for political strategists and organizations: securing endorsements early in the campaign may significantly enhance a candidate's competitiveness, particularly in crowded primaries. This suggests that endorsements not only reflect candidate quality but may actively shape electoral outcomes by signaling viability to voters and donors.

For future work, deeper analysis of endorsement types (e.g., elite vs. grassroots organizations) could provide more granular insights into which sources carry the most weight. In addition, exploring potential heterogeneity in endorsement effects across regions, election years, or candidate demographics would enrich the causal narrative. Finally, applying methods like matching, inverse propensity weighting, or instrumental variables could help triangulate the causal estimate under alternative assumptions.

6 Research Question 2: Bayesian Hierarchical Inference

Research Question: How does the estimated effect of high-profile endorsements on primary election outcomes vary across political parties?

6.1 Methods

In this question, we assess how the estimated effect of high-profile endorsements on primary election outcomes varies across political parties by using a Bayesian Hierarchical Model. This approach allows us to estimate party-specific effects for each endorser while sharing statistical strength across groups, particularly useful given that endorsements are usually party specific and are distributed unevenly at times.

The variables we observe for this question are the candidate's primary vote share (as the outcome variable), binary indicators of endorsements from high-profile individuals or organizations (e.g., Trump, EMILY's List) and the candidate's party affiliation, allowing us to compare effects between Democrats and Republicans. In our Bayesian Hierarchical Model, for each party, we model a candidate's vote share y_i as a function of:

- An intercept α (baseline support with no endorsements),
- A weighted sum of endorsements $x_{ij} \cdot \beta_j$,
- Gaussian noise σ.

We place hierarchical priors on the endorsement coefficients:

$$\beta_j = \mu_\beta + \sigma_\beta \cdot z_j, \quad \text{where } z_j \sim \mathcal{N}(0, 1)$$

$$\mu_\beta \sim \mathcal{N}(0, 1), \quad \sigma_\beta \sim \text{Exponential}(1.0)$$

$$\alpha \sim \mathcal{N}(0.5, 0.25), \quad \sigma \sim \text{Exponential}(1.0)$$

Estimation and Posterior Inference

We implement our model in `PyMC` and perform **Markov Chain Monte Carlo (MCMC)** sampling to estimate the posterior distributions of endorsement effects. We ran 2,000 tuning steps followed by 2,000 sampling steps across two chains, with a target acceptance rate of 0.995 to ensure stable and efficient exploration of the posterior space.

We estimate the model separately for Democratic and Republican candidates to enable comparison across party lines. Posterior means and 95% credible intervals are computed for each endorsement effect, and a sensitivity analysis is conducted using a more diffuse prior on the variance of endorsement effects.

This structure allows each endorsement type to have its own estimated effect within each party while still borrowing strength across endorsement types through shared hyperparameters.

6.2 Results

We estimate the effect of high-profile endorsements on primary vote share separately for Republican and Democratic candidates. The results are presented using posterior means and 95% credible intervals, forest plots, and posterior density distributions.

6.2.1 Republican Candidates

All selected Republican endorsements are associated with **positive effects** on vote share. Figure 5 shows the following:

- The **Trump endorsement** has the largest effect (approximately 0.094), with a credible interval that does not overlap zero, suggesting a strong, statistically credible boost.
- Endorsements from the **Main Street** and **Club for Growth** follow closely, also with positive means and relatively narrow intervals.
- Even endorsements with wider intervals, such as **Right to Life** and **Susan B. Anthony**, lean positively.

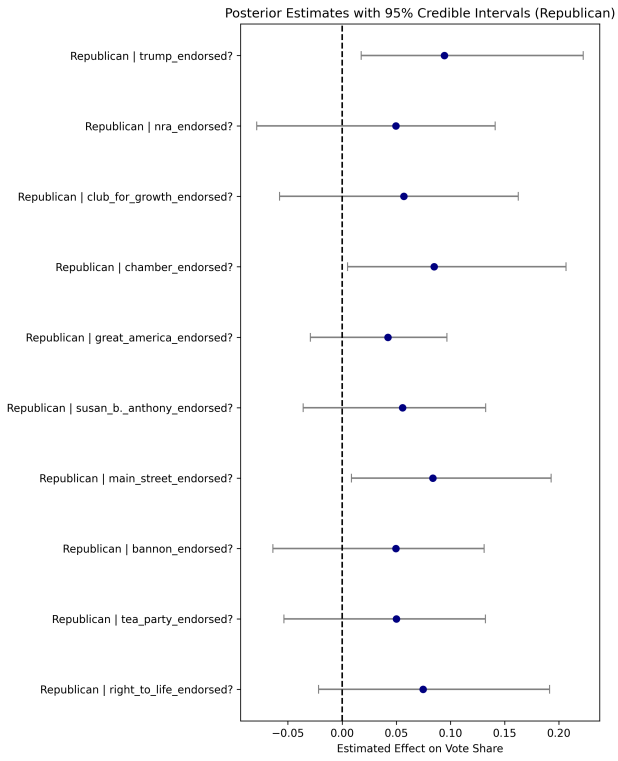


Figure 8: Estimated Effects of Endorsements on Republican Primary Vote Share

6.2.2 Democratic Candidates

The effects of endorsements for Democrats are more varied

- **Biden**, **Indivisible**, **Justice Democrats**, **PCCC**, and **Our Revolution** show **credible positive effects**, with posterior means above zero and 95% credible intervals that do not include zero.
- Endorsements from **Emily's List**, **Warren**, **Sanders**, and **VoteVets** have positive means, but their intervals overlap with zero, indicating more uncertainty.

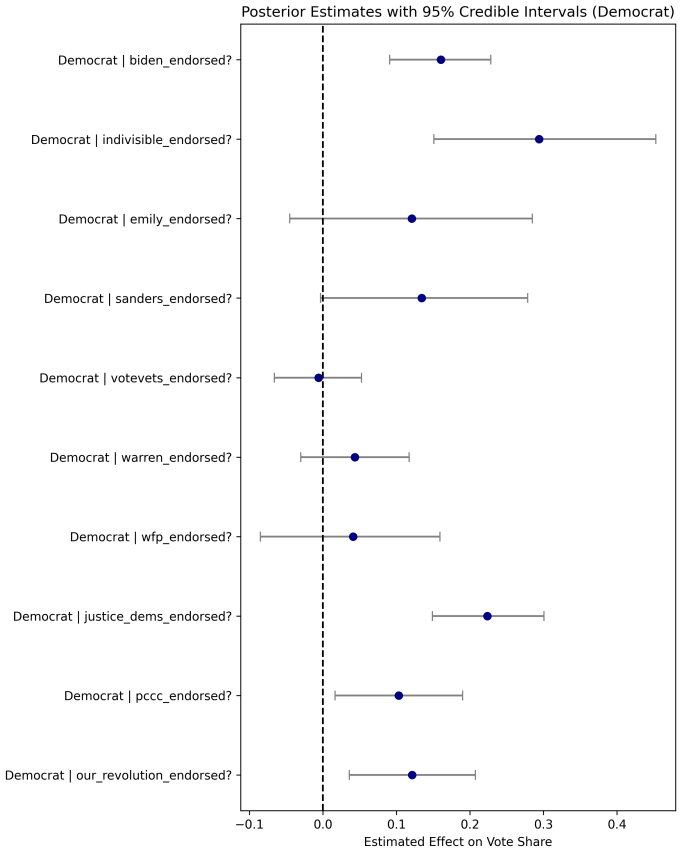
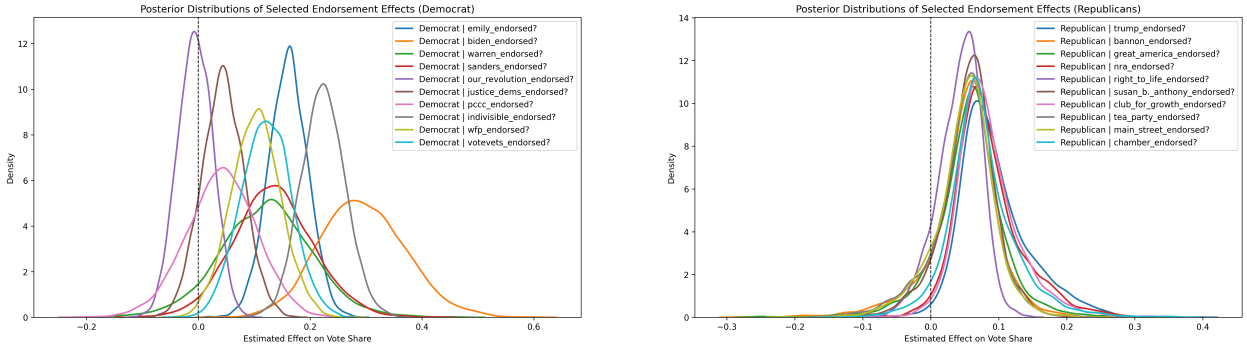


Figure 9: Estimated Effects of Endorsements on Democratic Primary Vote Share

6.3 Discussion

The posterior density plot below (Figure 10) highlights greater variation and uncertainty among Democratic endorsements. Several distributions are wider or overlapping, indicating weaker or inconsistent effects. While endorsements generally benefit candidates from both parties, Republican endorsements appear more uniformly positive, with tighter credible intervals and less variability.



(a) Democratic Party

(b) Republican Party

Figure 10: Posterior Distributions of Endorsement Effects

One possible explanation for this difference is the more centralized influence structure within the Republican Party. High-profile figures such as former President Donald Trump play a dominant role in shaping voter preferences, which may amplify the impact of his endorsement. In contrast, the Democratic Party is more ideologically and organizationally diverse, with multiple progressive, moderate, and institutional factions. This diversity may dilute the overall effect of individual endorsements, especially when candidates are supported by ideologically similar but separate groups.

We conducted a sensitivity analysis by increasing the hyperprior scale σ_β from 1.0 to 3.0 to allow greater variation in endorsement effects. The posterior means changed by less than 0.003 for all endorsements, and the 95% credible intervals remained broadly consistent. This suggests that our findings are robust to prior assumptions and that endorsement effects are primarily driven by the observed data.

6.3.1 Limitations

Our model does not include key contextual factors such as the demographic composition or voter history of the counties in which the primaries took place. As such, this analysis isolates endorsement effects and does not account for potentially confounding variables that could influence electoral outcomes. Future work could integrate geographic and socioeconomic data to provide a more comprehensive view of what drives candidate success.

6.3.2 Implications and Future Work

These findings suggest that endorsements from widely recognized or institutionally credible figures may be particularly effective in shaping primary outcomes, especially in races with low visibility or limited candidate differentiation. Future analyses could explore interactions with campaign spending and media exposure or examine longitudinal patterns across multiple election cycles.

7 Conclusions

This project explored the relationship between political endorsements and primary election outcomes in the 2022 U.S. elections. Through a combination of exploratory visualizations, causal inference models, and Bayesian hierarchical modeling, we assessed both the average and endorsement-specific impacts on candidate performance across parties.

Our causal analysis found that receiving a major endorsement significantly increases a candidate's primary vote share—by approximately 15 percentage points for Democrats and 21 points for Republicans. These findings were robust across model specifications and supported by diagnostic checks, such as propensity score overlap and outcome distribution comparisons.

Our Bayesian analysis further revealed that endorsement effects vary widely by endorser and by party. For Republicans, endorsements—particularly from Donald Trump and major conservative PACs—exhibited consistently strong and positive effects. For Democrats, the effects were more heterogeneous: some endorsements (e.g., Biden, Justice Democrats) had statistically credible benefits, while others were more uncertain.

Taken together, our results suggest that endorsements are not only reflective of candidate viability but also act as powerful tools that shape electoral outcomes. The influence of endorsements appears stronger when they come from figures with broad recognition and centralized influence within their party. These insights have practical implications for campaign strategy and political communication, especially in low-information primaries.

7.1 Recommendations

Campaigns may benefit from strategically prioritizing endorsements from well-recognized figures early in primary races, especially when running in low-information districts. However, reliance on endorsements also raises concerns about fairness and voter autonomy. Endorsements from centralized party elites could reduce ideological diversity and marginalize grassroots candidates, especially in polarized environments.

Future work could also build on this study by examining the effects of multiple endorsements in combination, integrating voter demographics and district characteristics, or analyzing whether endorsements influence general elections differently from primaries. Expanding this framework across election cycles may also reveal how the power of endorsements shifts over time and in response to political trends.

7.2 Broader Impacts

Beyond strategic implications, our findings raise important ethical considerations around political equity and democratic representation. Endorsements, particularly from high-profile figures or centralized party elites, can disproportionately shape media narratives and voter perception—potentially crowding out grassroots candidates or those from underrepresented communities. This dynamic may exacerbate inequities in political visibility and limit ideological diversity within primaries. Furthermore, heavy reliance on endorsements risks creating feedback loops where electability is perceived as a function of elite support rather than voter

engagement or policy alignment. Future work should investigate how endorsement systems intersect with campaign finance, media access, and systemic barriers to entry in politics to better understand their broader societal impact.

References

Boudreau, C., & MacKenzie, S. A. (2014). The persuasive effects of political endorsements: Reevaluating the informational value of campaign speeches. *American Politics Research*, *42*(3), 509–538. https://doi.org/10.1177/1532673X13507802

FiveThirtyEight. (2022). 2022 primary election data [Accessed: 2025-05-12].

Knight, B. G., & Chiang, C.-F. (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, *78*(3), 795–820. https://doi.org/10.1093/restud/rdq030