My title*

Bernice(Yixuan) Bao, Zheng(Zeb) Yang, Dongli Sun

March 11, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

You can and should cross-reference sections and sub-sections. We use R Core Team (2023) and Wickham et al. (2019).

The remainder of this paper is structured as follows. Section 2....

We use the dataset from 2022 Cooperative Election Study Schaffner, Ansolabehere, and Shih (2023). To further enable the analysis I employed the use of the package of ggplot(Wickham 2016) to generate histograms.

2 Data

Talk more about it.

Talk way more about it.

3 Model

3.1 Model set-up

Define y_i as the is the political preference of the respondent and equal to 1 if Biden and 0 if Trump. Then gender_i is the gender of the respondent and education_i is the highest education of the respondent.We could estimate the parameters using `stan_glm()`. Note that the model

*Code and data are available at: https://github.com/iloveyz12/US-Political-Support.

is a generally accepted short-hand. In practice `rstanarm` converts categorical variables into a series of indicator variables and there are multiple coefficients estimated. In the interest of run-time we will randomly sample 1,000 observations and fit the model on that, rather than the full dataset.

$$
\begin{aligned}
y_i|\pi_i &\sim \mathrm{Bern}(\pi_i) && (1)\\
\mathrm{logit}(\pi_i) &= \alpha + \beta_1 \times \mathrm{gender}_i + \beta_2 \times \mathrm{education}_i && (2)\\
\alpha &\sim \mathrm{Normal}(0, 2.5) && (3)\\
\beta_1 &\sim \mathrm{Normal}(0, 2.5) && (4)\\
\beta_2 &\sim \mathrm{Normal}(0, 2.5) && (5)
\end{aligned}
$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022), `modelsummary` package of Arel-Bundock (2022) and `here` package of Müller (2020). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between …. In particular…

Binomial logistic regression is a statistical method used to model the probability of a binary outcome variable. It's particularly suitable for situations where the dependent variable has two categories, such as in this study where we examine the likelihood of respondents voting for either Biden or Trump based on their gender and education level. The decision to employ binomial logistic regression for our study stems from its suitability for modeling binary outcome variables. As the outcome variable pertains to respondents' voting behavior, which involves choosing between Biden or Trump, binomial logistic regression is well-suited to capture this dichotomous outcome.

4 Results

Our results are summarized in Figure 1.

Based on the table outlining the likelihood of respondents supporting Biden based on gender and education levels, several key insights emerge. Firstly, the intercept value of 0.798 suggests that there is a baseline level of support for Biden among the surveyed population, regardless of gender or education. For men, their coefficient is not explicitly listed in the table, but it's implicitly represented by the intercept value. In this case, the intercept value of 0.798 can be interpreted as the baseline level of support for Biden among men.

Table 1: Whether a respondent is likely to vote for Biden based on their gender and education

	Support Biden
(Intercept)	0.781
	(0.440)
genderNon-binary	−25.738
	(21.060)
genderOther	0.200
	(1.585)
genderWoman	−0.600
	(0.134)
educationHigh school graduate	−0.477
	(0.457)
educationSome college	−0.879
	(0.457)
education2-year	−1.087
	(0.489)
education4-year	−1.062
	(0.455)
educationPost-grad	−1.638
	(0.480)
Num.Obs.	1000
R2	0.056
Log.Lik.	−645.011
ELPD	−655.6
ELPD s.e.	9.8
LOOIC	1311.2
LOOIC s.e.	19.6
WAIC	1308.6
RMSE	0.48

Conversely, for women, there is a direct coefficient provided in the table, which is -0.608. This negative coefficient suggests that women exhibit a slightly lower level of support for Biden compared to the baseline represented by men. When examining gender, it becomes evident that being non-binary has a significant negative impact on the likelihood of supporting Biden, with a coefficient of -26.049. This indicates a substantial decrease in support compared to other genders. Conversely, the coefficient for individuals identifying as "Other" gender is negligible at 0.139, suggesting that this category does not significantly influence support for Biden.

The coefficients related to education levels illustrate a clear trend in support for Biden among respondents.High school graduates, individuals with some college education, and those with 2-year, 4-year, and post-graduate degrees all exhibit negative coefficients ranging from -0.504 to -1.649.Starting from high school graduates to post-graduates, there is a consistent decrease in support, with coefficients indicating diminishing likelihoods of supporting Biden as educational attainment increases. High school graduates exhibit a moderate decrease compared to the baseline, followed by individuals with some college education, 2-year, 4-year, and post-graduate degrees, showing progressively larger declines in support. Particularly striking is the substantial drop in support among those with post-graduate degrees, indicating that as education level increases, the likelihood of supporting Biden decreases.

Overall, the model's R-squared value of 0.056 indicates that gender and education levels explain only a small proportion of the variance in support for Biden among respondents.However, the coefficients provide valuable insights into how gender identity and educational attainment may influence political preferences, with non-binary gender and higher education levels being associated with decreased support for Biden.

5 Discussion

5.1 First discussion point: What does this paper do?

This paper delves into the voter turnout of the 2022 United States election, employing CES research to sample and analyze data across different demographic groups. Utilizing the Bernoulli distribution, the study investigates whether the gender and education of voters impact their candidate preferences. The analysis reveals a consistent trend: regardless of gender or education level, the majority of voters cast their ballots for Biden over Trump. Recognizing this phenomenon, the paper aims to conduct a deeper exploration into why Biden garnered widespread support and its potential implications for future elections.

5.2 Second discussion point: What have we learned from the world?

From existing global observations from Cooperative Election Study (CES), we've gathered insights into the dynamics of voter behavior and electoral outcomes. Studies have highlighted

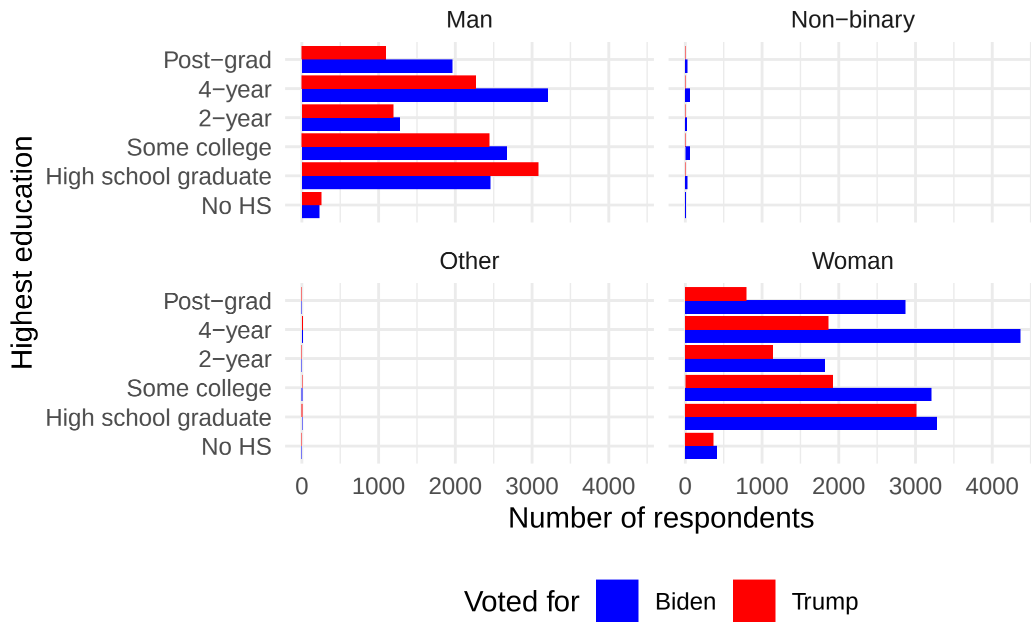

Figure 1: The distribution of presidential preferences, by gender, and highest education

various factors influencing voter decisions, including socioeconomic status, political ideology, and candidate messaging. Moreover, research into past elections has shown that demographic characteristics such as gender and education can play significant roles in shaping voting patterns. Understanding these dynamics helps contextualize the findings of this paper and provides a foundation for further analysis.

5.3 Third discussion point: What's another thing we've learned from the world?

Another crucial insight gleaned from global observations is the paramount importance of candidate appeal and the efficacy of campaign strategies in shaping voter preferences. Effective communication strategies, including resonant messaging and clear articulation of policy proposals, play a pivotal role in swaying undecided voters and mobilizing support from diverse demographic groups. Additionally, demonstrated leadership qualities and the ability to inspire confidence and trust in potential voters can significantly impact electoral outcomes. Moreover, the pervasive influence of media narratives, social networks, and broader societal discourse cannot be underestimated in shaping public opinion and ultimately determining electoral results. The intricate interplay between these factors, alongside demographic variables such as gender and education, offers a comprehensive understanding of the multifaceted dynamics that drive voter behavior. By delving into these nuanced interactions, researchers can glean invaluable insights into the complex mechanisms that underpin electoral decision-making processes, thus enriching our understanding of political phenomena and informing strategic approaches to future elections.

5.4 Weaknesses and next steps

Despite its contributions, this paper has several limitations that warrant acknowledgment. Firstly, while the analysis identifies correlations between gender, education, and candidate preference, it may overlook other influential factors such as race, age, and geographic location. Additionally, the reliance on CES research for sampling may introduce biases that affect the generalizability of the findings. Moreover, the study lacks qualitative insights into voters' motivations and decision-making processes, which could provide richer context for interpreting the results.

Moving forward, further research is needed to deepen our understanding of the factors driving voter behavior and electoral outcomes. Future studies should employ more diverse sampling methods and incorporate qualitative methodologies to capture the nuances of voter preferences. Additionally, exploring the intersectionality of demographic variables and considering evolving societal trends will enhance the comprehensiveness of analyses. Moreover, longitudinal studies tracking voter preferences over time can provide valuable insights into the evolving political landscape and inform strategic adjustments for future campaigns. By addressing these areas, researchers can advance our understanding of electoral dynamics and contribute to more informed policymaking and political strategies.

Appendix

A Additional data details

B Model details

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

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

Schaffner, Brian, Stephen Ansolabehere, and Marissa Shih. 2023. "Cooperative Election Study Common Content, 2022." Harvard Dataverse. https://doi.org/10.7910/DVN/PR4L8P.

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

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.