Demographic Groups with Stronger Biden Support in the Election Race'*

Yang Zhou

April 1, 2024

This paper explores the relationship between voting preferences and the predictors of education, income, and personal gun ownership in the United States through analyzing 2022 US election database. By examining data from Cooperative Election Study, we identify clear patterns indicating that higher income are more likely to vote for Biden. Besides, the higher personal gun ownership will also significantly increase the likelihood of voting for Biden. By contrast, the people with higher education level show less probability to vote for Biden. Our findings underscore the critical impact of socio-economic status and personal security concerns on electoral outcomes, shedding light on the underlying dynamics that shape voter behavior in contemporary American politics. Ultimately, this research contributes to our understanding of the complex factors that drive electoral decisions, offering valuable insights for policymakers, political strategists, and citizens aiming to foster more informed and equitable democratic processes.

Table of contents

*Code and data are available at: <https://github.com/yangzhoucoco/Political-support-in-the-United-State.git >

1 Introduction

2020 United States Presidential Election stood as an important political event, reflecting societal divisions and highlighting the influence of various demographic factors on voting behavior. This election, which culminated in the victory of Joe Biden over incumbent Donald Trump, has spurred a renewed interest in understanding the dynamics of voter preferences and the underlying factors that drive electoral decisions. Our study aims to dissect the relationship between voters' educational background, income levels, personal gun ownership, and their voting preferences, employing logistic regression analysis to reveal the potential associations between different factors.

The Cooperative Election Study (CES)conduct a throughout survey encompassing a wide range of voter demographics and attitudes.CES collaborate with 60 researches teams to assembles sample comprising 60,000 cases. The cases are based on the survey of adult pariticipants in 2022 fall. YouGov was commissioned to help the teams to conduct national sample surbey. They conducted the survey before and after elections. The first round survey were completed from September 29 to November 8, 2022 and the second period is gathered from November 10 to December 15, 2022. By offering an throughout views to different constituencies, voter demographics and voter behaviors, CES reveals the complexity of electoral behavior and voter demographics. By understanding voter behaviors„ we gain valuable insights that pave the way for refining and advancing democratic practices.

In addition, by leveraging data from the 2022 CES, this paper delves into how variables such as education, income, and gun ownership serve as predictors for electoral choices, specifically the likelihood of voting for Biden or Trump. This analysis not only reveal the individual impact of these factors but also contributes to a broader understanding of the socio-economic and cultural impact of American electoral behavior.

Based on our model, I found that individuals with higher incomes demonstrate a greater propensity to vote for Biden. Additionally, elevated levels of personal gun ownership significantly correlate with an increased likelihood of supporting Biden. Conversely, individuals possessing higher levels of education exhibit a reduced probability of voting for Biden.

To be more specific, this paper is structured as follow: in {#sec-data}, I will introduce the data variable and how the data is used for the overall analysis.Visualization would also be included. In {#sec-model}, I will set up the model to predict the relationship between the predictors and electoral outcomes. Model justification will provide a rationale for selecting logitics model. {#sec-results} is focused on explain the result of the model. Lastly, {#sec-Discussion} provides a discussion on what we found in the model. I will also talk about the weakness of this paper along with the further study on this topic.

I will uses the programming language `R`(R Core Team 2023) to analyze the data. `dplyr`(Wickham et al. 2023), `tidyverse`(Wickham et al. 2019), `ggplot2`(Wickham 2016), `knitr` (Xie 2014), `here` (Müller 2020) , and `kableExtra` (Zhu 2024),`modelsummary`(Arel-Bundock 2022), and `rstanarm`(Brilleman et al. 2018) will help me in the model and visualization part.

2 Data

2.1 Source and Methodology

2.2 Sampling and Sampling Matching

2.3 Weightening

2.4 Vote Vaildation

2.5 Variables

Table 1: Brief summary of variables types

votereg	presvote20post	educ	faminc_new	gunown
1	1	6	11	3
1	1	3	8	3
1	1	5	6	3
1	1	6	11	3
1	6	6	9	3
1	2	5	7	3
1	1	2	1	3
1	1	6	11	3
1	1	5	3	3
1	1	6	4	3

Figure 1: Brief summary of variables types

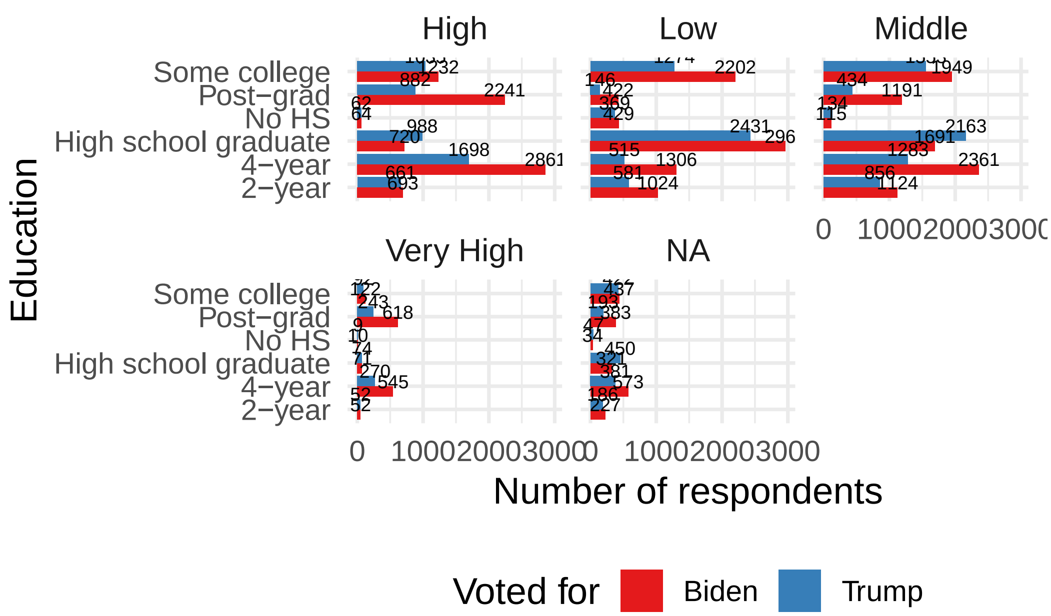

Figure 2: The distribution of presidential preferences, by gender, and race

2.6 Measurements

3 Model

3.1 Model set-up

Define y_i is the political preference of the respondent and equal to 1 if Biden and 0 if Trump. Then education_i is education level of the respondent and income_i is the income of the respondent.

The parameters can be derived utilizing the `stan_glm()` function. Practically, `rstanarm` transforms categorical variables into a set of binary indicators, leading to the estimation of numerous coefficients. To optimize computational efficiency, we will select a random subset of 500 data points for model fitting instead of using the entire dataset.

$$
\begin{aligned}
y_i|\pi_i &\sim \text{Bern}(\pi_i) &(1)\\
\text{logit}(\pi_i) &= \beta_0 + \beta_1 \times \text{education}_i + \beta_2 \times \text{income}_i &(2)\\
\beta_0 &\sim \text{Normal}(0, 2.5) &(3)\\
\beta_1 &\sim \text{Normal}(0, 2.5) &(4)\\
\beta_2 &\sim \text{Normal}(0.2.5) &(5)
\end{aligned}
$$

Table 2: Explanatory models of political preferences based on gender and race (n = 500)

	Support Biden
(Intercept)	−0.929
	(0.359)
educ4-year	−0.454
	(0.372)
educHigh school graduate	0.181
	(0.374)
educNo HS	0.778
	(0.828)
educPost-grad	−1.202
	(0.469)
educSome college	0.139
	(0.374)
income_groupLow	−0.118
	(0.309)
income_groupMiddle	0.522
	(0.274)
income_groupVery High	0.806
	(0.511)
gunownPersonally own a gun	1.136
	(0.255)
gunownSomeone in the household owns a gun	0.199
	(0.346)
Num.Obs.	394
R2	0.133
Log.Lik.	−236.683
ELPD	−248.3
ELPD s.e.	8.4
LOOIC	496.7
LOOIC s.e.	16.8
WAIC	496.6
RMSE	0.45

3.2 Model justification

4 Results

5 Discussion

5.1 Findings

5.2 High income

5.3 Education

5.4 Gun ownership

5.5 Application of Logistic Regression

5.6 Linear Regression Model

5.7 Model Building

5.8 Weaknesses and Next Steps

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.

Brilleman, SL, MJ Crowther, M Moreno-Betancur, J Buros Novik, and R Wolfe. 2018. "Joint Longitudinal and Time-to-Event Models via Stan." https://github.com/stan-dev/stancon_talks/.

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/.

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.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. http://haozhu233.github.io/kableExtra/.