

# Higher Information and Higher Trust in Government ...\*

My subtitle if needed

Talia Fabregas

December 1, 2024

First sentence. Second sentence. Third sentence. Fourth sentence. Key Trump 2024 voting blocs, including young men without a college education were less likely to vote in the 2022 midterms when he was not on the ballot compared to 2020

## 1 Introduction

Overview paragraph

Estimand paragraph This study aims to examine the effects of 2020 presidential vote preference, age, highest level of education completed, trust in the federal government, knowledge of which party holds a majority in Congress, and interest in politics on the likelihood that an individual who voted for Donald Trump or Joe Biden in the 2020 U.S. presidential election would also vote in the 2022 U.S. midterm election. The estimand, which can never be known with complete certainty, is the true effect of presidential vote preference, age, education, trust, civic knowledge, and interest in politics on the likelihood that an individual who voted in a presidential election will also vote in the subsequent midterm election (Alexander 2023) .

Results paragraph Based on the results of applying logistic regression models to predict voter turnout and vote preference, I estimate that ...

Why it matters paragraph In the 2024 election, Democratic Senate candidates (both incumbent and non-incumbent) were elected in states that President-Elect Trump carried: Michigan, Wisconsin, Nevada, and Arizona. In Nevada and Wisconsin, Senators Jackie Rosen and Tammy Baldwin

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2 contains an overview of the survey data set from Schaffner, Ansolabehere, and Shih (2023) that

---

\*Code and data are available at: <https://github.com/taliafabs/US-Midterms-2022.git>.

was used, visualizations of different variables, and summary statistics. Section 3 contains the logistic regression models used to predict the probability that an individual who voted for either Trump or Biden in the 2020 presidential election would also vote in the 2022 midterm election and the probability that an individual would support Trump over Biden in 2020. **?@sec-results** contains tables and data visualizations that present results about what characteristics made individuals who voted in the 2020 presidential election more likely to also vote in the 2022 midterm election and what characteristics make individuals more likely to support Trump. **?@sec-discussion** contains detailed discussiona about the results presented in **?@sec-results** ...

## 2 Data

### 2.1 Overview

The survey data set from Schaffner, Ansolabehere, and Shih (2023) contains ...

We use the statistical programming language R (R Core Team 2023).... Our data (Schaffner, Ansolabehere, and Shih 2023).... Following Alexander (2023), we consider...

Overview text

### 2.2 Measurement

The primary measurement task is to capture voter characteristics and translate it into conclusions and predictions about voting behavior in U.S. presidential, senate, and congressional elections. During every midterm and presidential election cycle, teams of pollsters work together to conduct surveys that measure public opinion, voter demographics, and candidate preferences. Each survey response is an individual voter's preferences, which pollsters collect to build the Comprehensive Election Study (CES) dataset for each election cycle.

### 2.3 Outcome variables

Add graphs, tables and text. Use sub-sub-headings for each outcome variable or update the subheading to be singular.

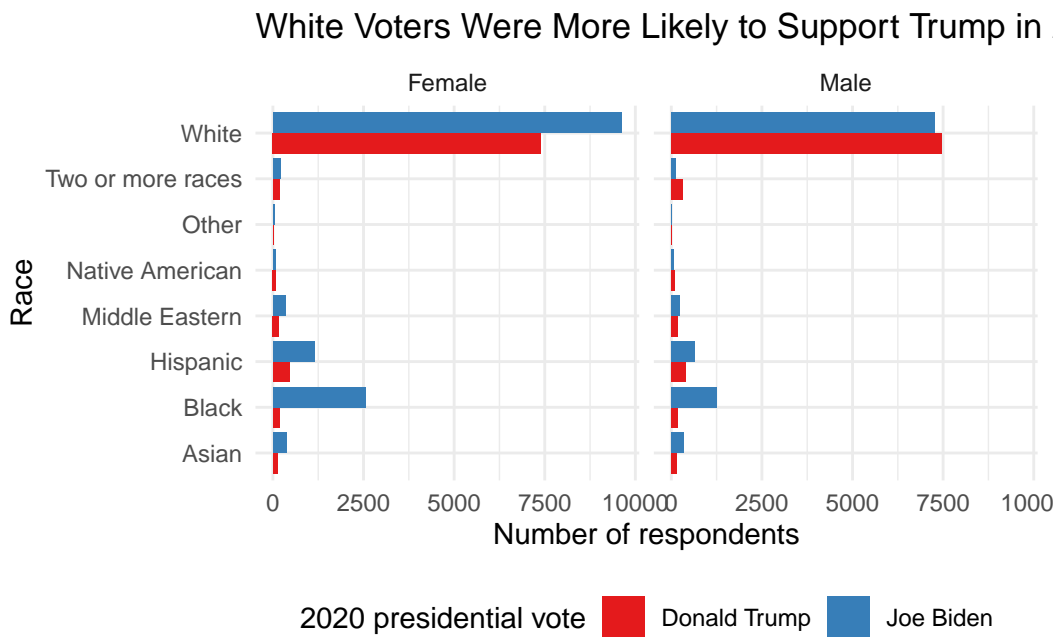
Some of our data is of penguins (Figure 1), from ...

Table 1: Biden 2020 Voters Voted in 2022 midterms slightly more than Trump 2020 voters.  
73.26% of respondents who voted for Trump in 2020 voted in 2022, compared to  
74.67% of respondents who voted for Biden.

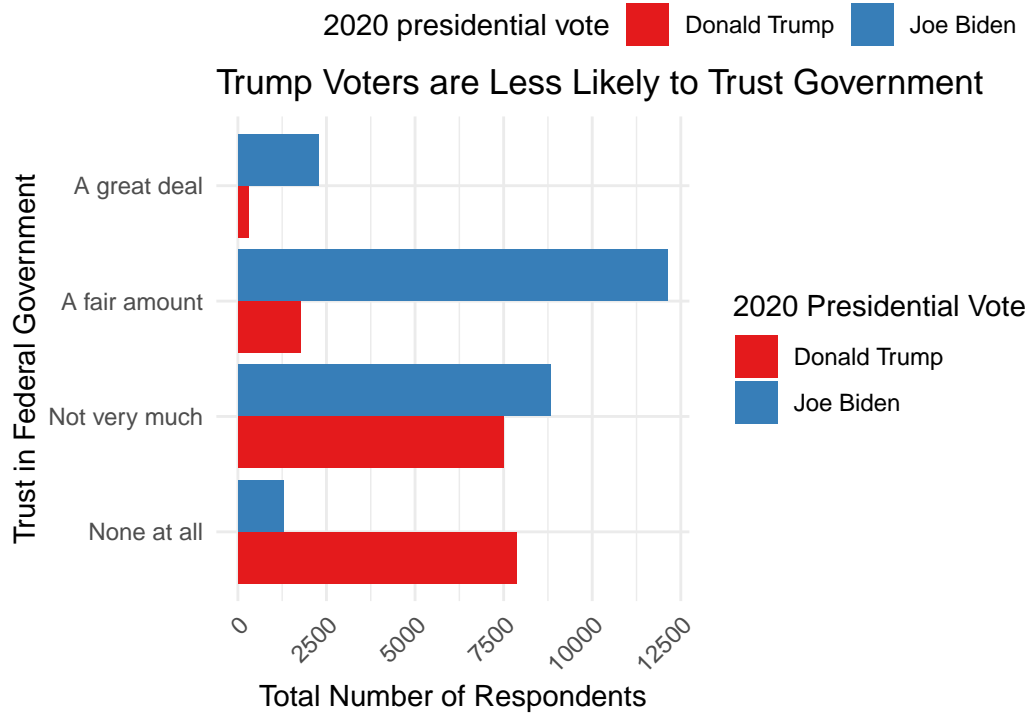
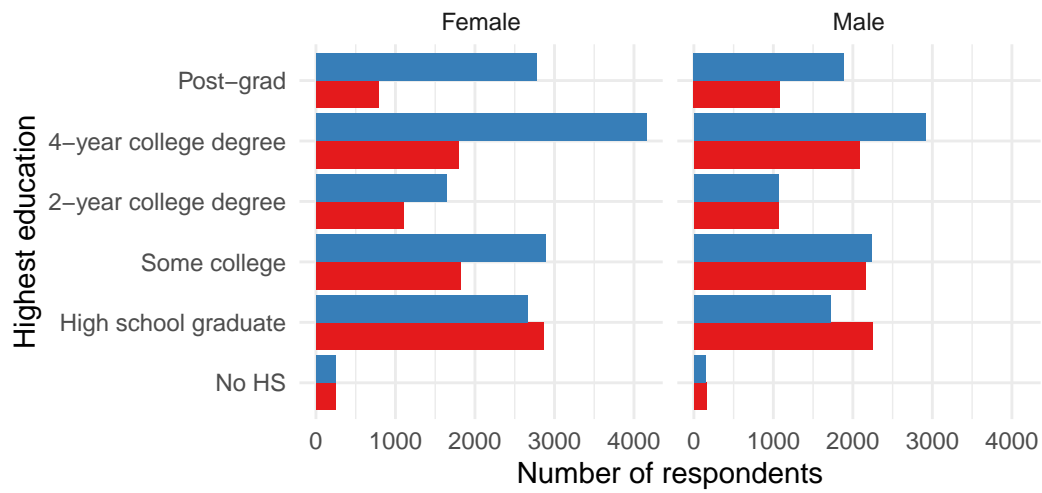
## 2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.



## College-Educated Voters are Less Likely to Support



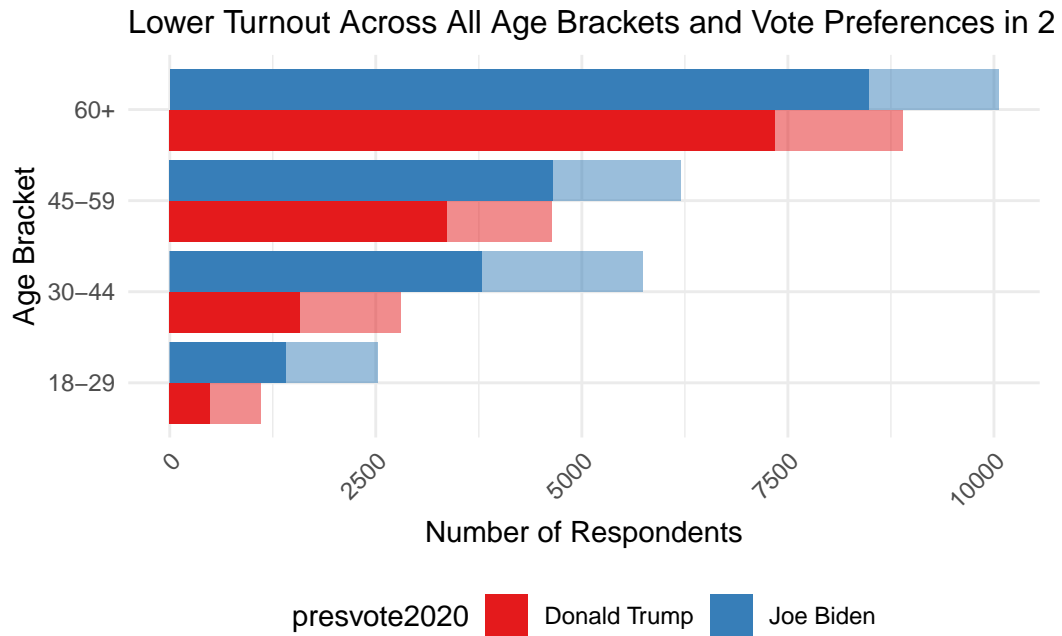
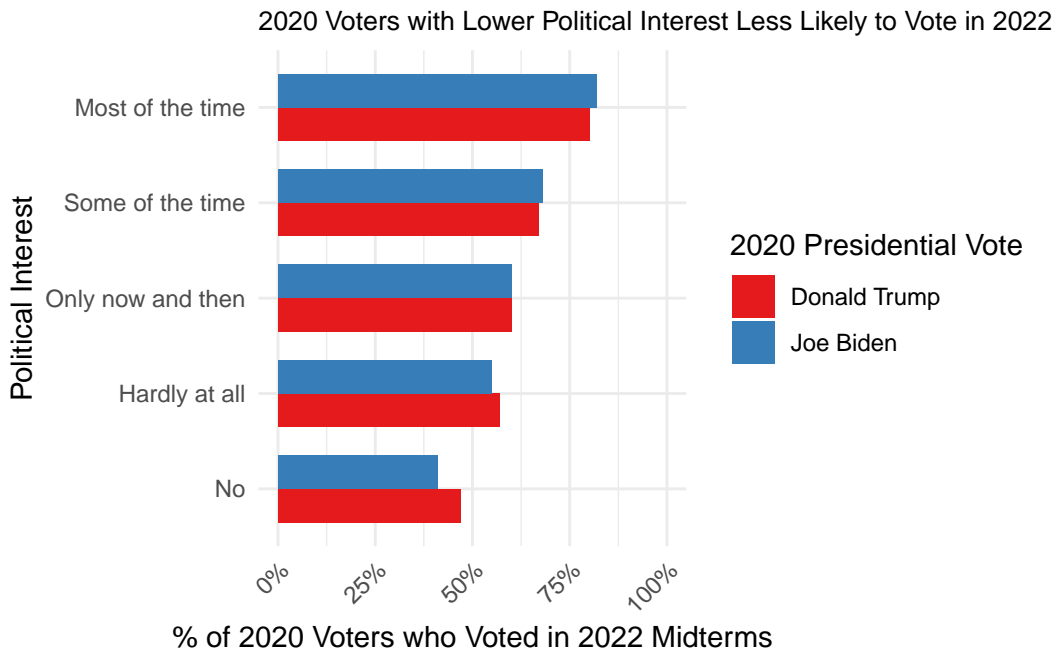


Figure 1: More than half of 18-29 year old respondents who voted for Trump in 2020 presidential election did not vote in the 2022 midterm election.



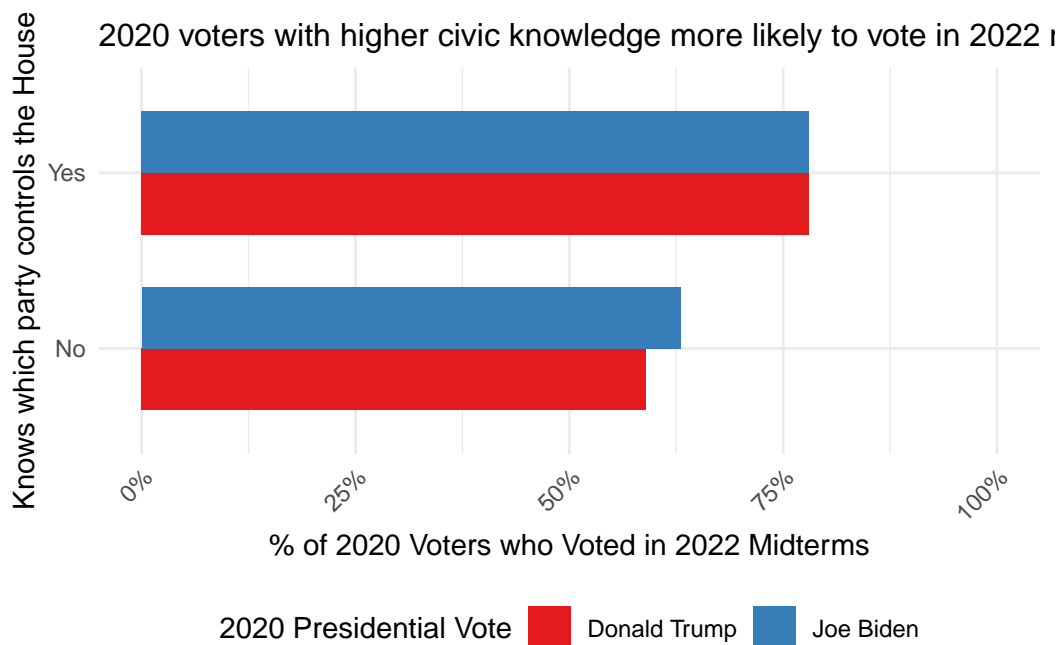


Figure 2: There is a nearly 15-point gap in 2022 turnout amongst respondents who voted for Trump or Biden in 2020 and know which party has a majority in the U.S. House of Representatives versus those who do not.

## 3 Model

The goal of my modeling strategy is to use logistic regression models to investigate ... I use two logistic regression models: one to model the probability that an individual who voted in the 2020 presidential election would also vote in the 2022 midterm election, and one to model the probability that an individual voter supports Trump.

The goal of my modeling strategy is twofold. Firstly, I want to The goal of our modelling strategy is twofold. Firstly,...

The model to predict *voted\_in\_2022* uses the following predictors:

The model to predict *voted\_for\_trump* uses the following predictors:

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix [C](#).

### 3.1 Model set-up

#### 3.1.1 Turnout model

Define  $y_i$  as the number of seconds that the plane remained aloft. Then  $\beta_i$  is the wing width and  $\gamma_i$  is the wing length, both measured in millimeters.

Models the probability that someone who voted for either Trump or Biden in 2020 also voted in the 2022 mid-term elections.

Define  $\pi_i$  as the probability that survey respondent  $i$  voted in the 2022 midterm election

$$\begin{aligned} y_i | \pi_i &\sim \text{Bern}(\pi_i) \\ \text{logit}(\pi_i) &= \beta_0 + \beta_1 \times \text{presvote2020}_i + \beta_2 \times \text{educ}_i + \beta_3 \times \text{trust}_i + \beta_4 \times \text{know\_us\_house}_i \\ &\quad + \beta_5 \times \text{know\_us\_senate}_i + \beta_6 \times \text{political\_interest}_i \\ \beta_0 &\sim \text{Normal}(0, 2.5) \\ \beta_1 &\sim \text{Normal}(0, 2.5) \\ \beta_2 &\sim \text{Normal}(0, 2.5) \\ \beta_3 &\sim \text{Normal}(0, 2.5) \\ \beta_4 &\sim \text{Normal}(0, 2.5) \\ \beta_5 &\sim \text{Normal}(0, 2.5) \\ \beta_6 &\sim \text{Normal}(0, 2.5) \end{aligned}$$

Models the probability that a respondent voted for Trump in 2020:

$$\begin{aligned}
y_i | \pi_i &\sim \text{Bern}(\pi_i) \\
\text{logit}(\pi_i) &= \beta_0 + \beta_1 \times \text{age\_bracket}_i + \beta_2 \times \text{gender}_i + \beta_3 \times \text{race}_i + \beta_4 \times \text{educ}_i \\
&\quad + \beta_5 \times \text{trust}_i + \beta_6 \times \text{know\_us\_house} - i \\
\beta_0 &\sim \text{Normal}(0, 2.5) \\
\beta_1 &\sim \text{Normal}(0, 2.5) \\
\beta_2 &\sim \text{Normal}(0, 2.5) \\
\beta_3 &\sim \text{Normal}(0, 2.5) \\
\beta_4 &\sim \text{Normal}(0, 2.5) \\
\beta_5 &\sim \text{Normal}(0, 2.5) \\
\beta_6 &\sim \text{Normal}(0, 2.5)
\end{aligned}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

### 3.1.2 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance  $\theta$ . Confounders have not been eliminated -

I found that trust in the federal government is related to both support for Trump and likelihood that someone who voted in 2020 also voted in 2022. education same thing.

## 3.2 Model Weaknesses and Limitations

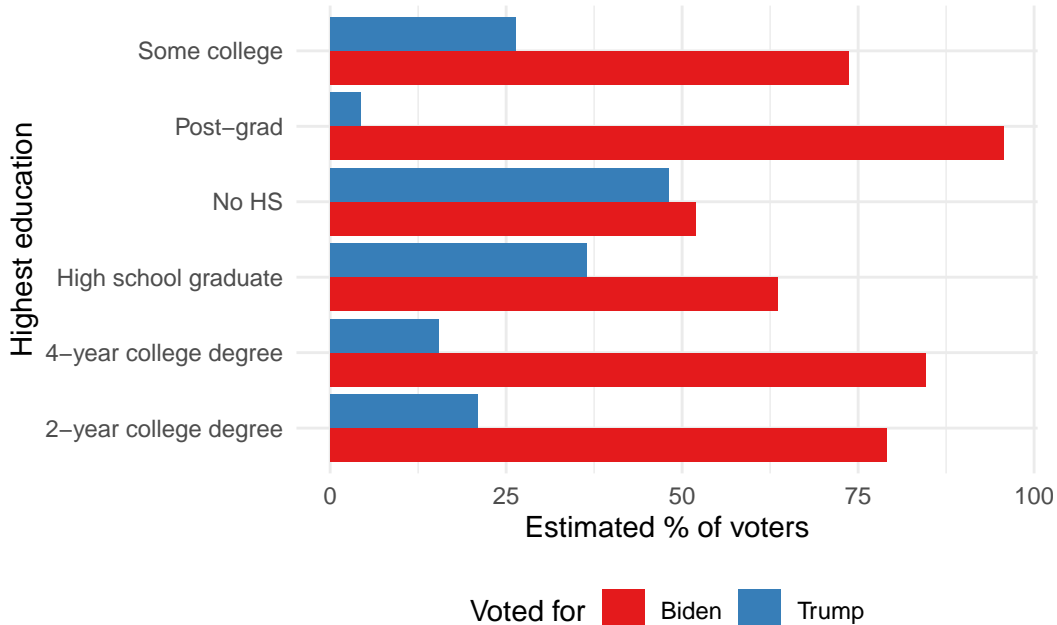
Trained using a random subset of the data. Refer to codebook about the risks of that. I randomly sampled 5,000 observations from the 2022 CES survey data set. This was a tradeoff; training a model on more observations would have had a significantly longer runtime. However, randomly subsetting the 2022 CES survey data set is risky because ...

## 4 Results

The results are summarized in Appendix [C.1](#).



## 4.1 Trump support amongst young men



## 4.2 Low Propensity Trump Voters are Less Likely to Vote in Midterms

This result foreshadowed the fact that young men, especially those without a college degree who are infrequent voters, would become a key part of Trump's winning coalition in 2024.

```
# young men who voted for trump, no college education, low trust, do not know us house
young_men_lp <- turnout_young_men %>%
  filter(presvote2020 == "Donald Trump" &
    (educ == "High school graduate" | educ == "No HS" | educ == "Some college") &
    (trustfed == "None at all" | trustfed == "Not very much") &
    know_us_house == "No"
  )

young_men_lp_trump <- new_preference_data |>
  filter(gender == "Male" & age_bracket == "18-29") %>%
  filter( (educ == "High school graduate" | educ == "No HS" | educ == "Some college") &
    (trustfed == "None at all" | trustfed == "Not very much") &
    know_us_house == "No"
  )
```

#### **4.3 Voters with Little Interest in Politics are Less Likely to Vote in Midterms**

#### **4.4 College Education is Associated with Higher Support for the Democratic Party and Higher Likelihood to Vote in Midterms**

#### **4.5 Young men are more likely to support Trump and less likely to vote when he is not on the ballot**

### **5 Discussion**

#### **5.1 First discussion point**

If my paper were 10 pages, then should be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

#### **5.2 Second discussion point**

Please don't use these as sub-heading labels - change them to be what your point actually is.

#### **5.3 Polling Misses and Trump's Unique Appeal to Low Propensity Voters**

Young men who are not college-educated, do not know which party controls the U.S. House of Representatives,

#### **5.4 2024 Polling Misses and Non-Response Bias**

#### **5.5 Weaknesses and next steps**

Weaknesses and next steps should also be included.

## Appendix

### A Surveys, Sampling, and Observational Data

#### A.1 Deep Dive into CES Methodology

##### A.1.1 Sampling Method

##### A.1.2 Weighting

##### A.1.3 Verification

##### A.1.4 Weaknesses and Limitations

#### A.2 Second Point: Idealized Survey Methodology to Study 2024 Down-Ballot Voter Attrition

#### A.3 Social Desirability Bias and Identifying Non-Voters

### B Additional data details

#### B.1 Data cleaning

### C Model details

#### C.1 Model Results

The model summary is shown in Table [2](#).

?@tbl-modelsummary makes

#### C.2 Posterior predictive check

In ?@fig-ppcheckandposteriorvsprior-1 we implement a posterior predictive check. This shows the comparison of the outcome variable *vote\_2022*, with simulations from the posterior distribution (Alexander 2023).

In ?@fig-ppcheckandposteriorvsprior-2 we compare the posterior with the prior. This shows...

Table 2: Explaining whether someone who voted in the 2020 presidential election also voted in the 2022 midterm election, based on who they voted for, age, education, trust in government, civic knowledge, and political interest

	Voter Turnout Model
(Intercept)	−0.79 (0.21)
presvote2020Joe Biden	0.24 (0.08)
age_bracket30-44	0.55 (0.10)
age_bracket45-59	1.16 (0.11)
age_bracket60+	1.79 (0.11)
educ4-year college degree	0.36 (0.11)
educHigh school graduate	−0.27 (0.12)
educNo HS	−1.01 (0.17)
educPost-grad	0.42 (0.13)
educSome college	0.10 (0.12)
trustfedA great deal	−0.86 (0.13)
trustfedNone at all	0.35 (0.11)
trustfedNot very much	0.39 (0.08)
truststateA great deal	−0.07 (0.11)
truststateNone at all	0.15 (0.10)
truststateNot very much	−0.07 (0.08)
know_us_houseYes	0.36 (0.07)
know_us_senateYes	0.02 (0.06)
political_interestMost of the time	0.28 (0.15)
political_interestNo	−1.07 (1.12)
political_interestOnly now and then	−0.32 (0.17)
political_interestSome of the time	0.00 (0.16)
Num.Obs.	7500
R <sup>2</sup>	0.134
Log.Lik.	−3270.706
WAIC	6606.0
RMSE	0.42

Examining how the model fits, and is affected  
by, the data

### C.3 Credibility Intervals

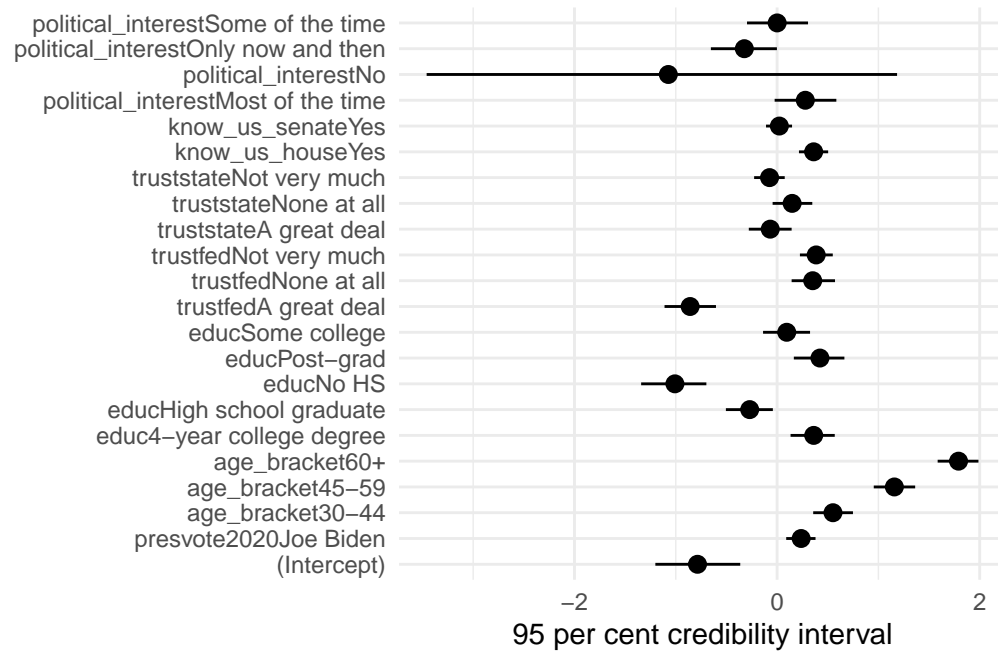


Figure 3: Credible intervals for predictors of voter turnout in the 2022 U.S. midterm elections

### C.4 Diagnostics

Figure 4a is a trace plot. It shows... This suggests...

Figure 4b is a Rhat plot. It shows... This suggests...

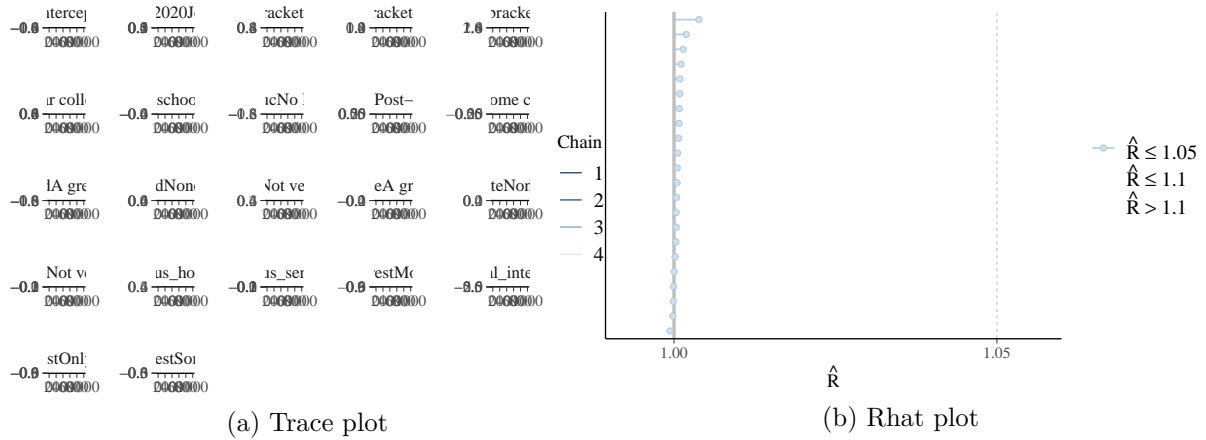


Figure 4: Checking the convergence of the MCMC algorithm

## References

- Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. <https://tellingstorieswithdata.com/>.
- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. “rstanarm: Bayesian applied regression modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- 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>.