

My title*

My subtitle if needed

Talia Fabregas Fatimah Yunusa Aamishi Sundeep Avarsekar

March 10, 2024

The 2024 U.S. Presidential Election ...

1 Introduction

The 2024 U.S. Presidential Election will take place on Tuesday November 5 2024. Incumbent President Joseph R. Biden Jr. will seek a second term. Former President Donald J. Trump GOP nominee.

Delegates to secure the nomination, Nikki Haley has dropped out of GOP You can and should cross-reference sections and sub-sections. We use R Core Team (2023) and

The remainder of this paper is structured as follows. Section 2 discusses the survey and post-stratification data used.

Clear gap that needs to be filled ... what is the research gap and why is this important?

- 2024 US Presidential Election
- how consequential this election is
- what is on the ballot
- women's rights
- trump's project 2025 thing

If someone's done it before not on this data set not in this context

Previous groups have looked at ...

R Core Team (2023) was used

*Code and data are available at: <https://github.com/taliafabs/US-Election-Forecast-2024.git>

2 Data

```
library(arrow)
library(dplyr)
survey_analysis_data <- arrow::read_parquet("/Users/talia/US-Election-Forecast/data/analysis")
# poststrat_analysis_data <- read_parquet("/Users/talia/US-Election-Forecast/data/analysis")
print(survey_analysis_data)

# A tibble: 924 x 21
  vote24      vote_biden pid7 presvote16post presvote20post ideo5 birthyr age
  <chr>      <dbl> <chr> <chr>          <chr>          <chr> <dbl> <dbl>
1 Joe Biden      1 Not ~ Hillary Clint~ Joe Biden      Mode~      1993    31
2 Donald Tr~      0 Lean~ Donald Trump  Jo Jorgensen  Cons~      1993    31
3 Donald Tr~      0 Stro~ Donald Trump  Donald Trump  Very~      1983    41
4 Joe Biden      1 Lean~ Did not vote ~ Joe Biden      Mode~      1983    41
5 Joe Biden      1 Lean~ Hillary Clint~ Joe Biden      Libe~      1980    44
6 Donald Tr~      0 Not ~ Did not vote ~ Jo Jorgensen  Cons~      1998    26
7 Donald Tr~      0 Lean~ Donald Trump  Donald Trump  Cons~      1974    50
8 Joe Biden      1 Stro~ Hillary Clint~ Joe Biden      Very~      1997    27
9 Donald Tr~      0 Stro~ Did not vote ~ Donald Trump  Very~      2002    22
10 Joe Biden      1 Lean~ Did not vote ~ Did not vote ~ Not ~      1994    30
# i 914 more rows
# i 13 more variables: age_bracket <chr>, sex <fct>, races <fct>,
#   race_white <dbl>, race_asian <dbl>, race_black <dbl>, race_hispanic <dbl>,
#   race_native <dbl>, marstat <chr>, education_level <fct>, faminc_new <fct>,
#   state <fct>, urban <chr>

proportion_biden <- survey_analysis_data %>%
  filter(vote_biden == 1) %>%
  summarise(proportion = n() / nrow(survey_analysis_data))

proportion_biden_by_state <- survey_analysis_data %>%
  group_by(state) %>%
  summarise(proportion_biden = mean(vote_biden == 1, na.rm = TRUE))

proportion_by_state <- survey_analysis_data %>%
  group_by(state) %>%
  summarise(proportion = n() / nrow(survey_analysis_data))

proportion_biden_by_state
```

```
# A tibble: 50 x 2
  state                proportion_biden
  <fct>                <dbl>
1 Alabama              0.625
2 Alaska               0.333
3 Arizona              0.391
4 Arkansas              0.667
5 California            0.670
6 Colorado              0.533
7 Connecticut          0.857
8 Delaware              0.5
9 District of Columbia 1
10 Florida              0.523
# i 40 more rows
```

```
proportion_by_state
```

```
# A tibble: 50 x 2
  state                proportion
  <fct>                <dbl>
1 Alabama              0.00866
2 Alaska               0.00325
3 Arizona              0.0249
4 Arkansas              0.00974
5 California            0.102
6 Colorado              0.0162
7 Connecticut          0.00758
8 Delaware              0.00433
9 District of Columbia 0.00216
10 Florida              0.0931
# i 40 more rows
```

```
# only includes 7 observations from hawaii
# Hawaii is historically a Democratic stronghold
```

2.1 Survey Data

Survey data is from Iyengar, Lelkes, and Westwood (2024) This section should talk about the survey data set

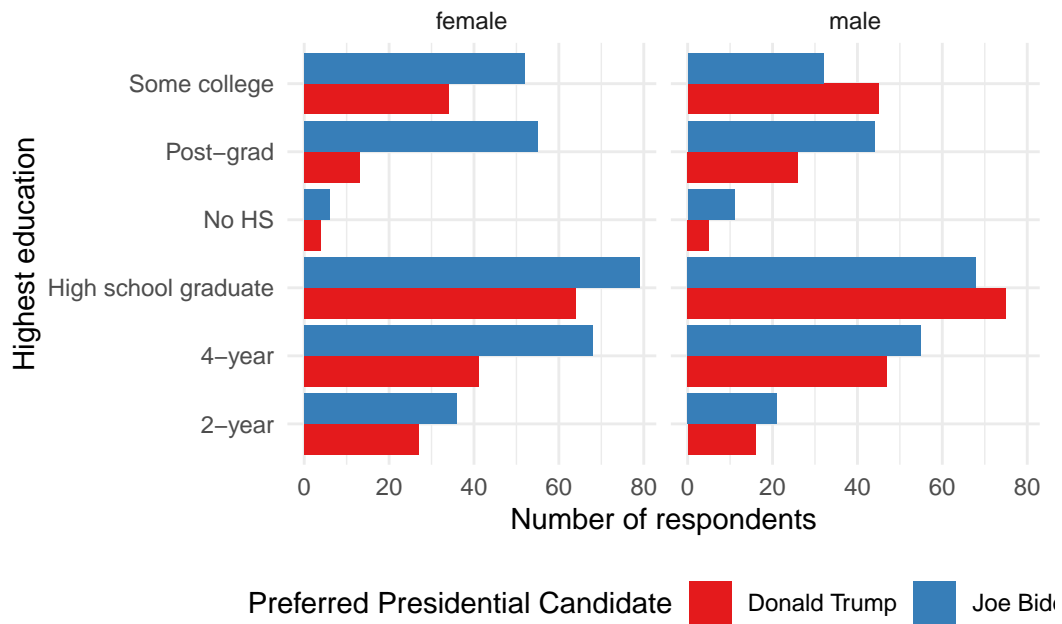


Figure 1: The Distribution of Presidential Preferences, by gender and highest level of education

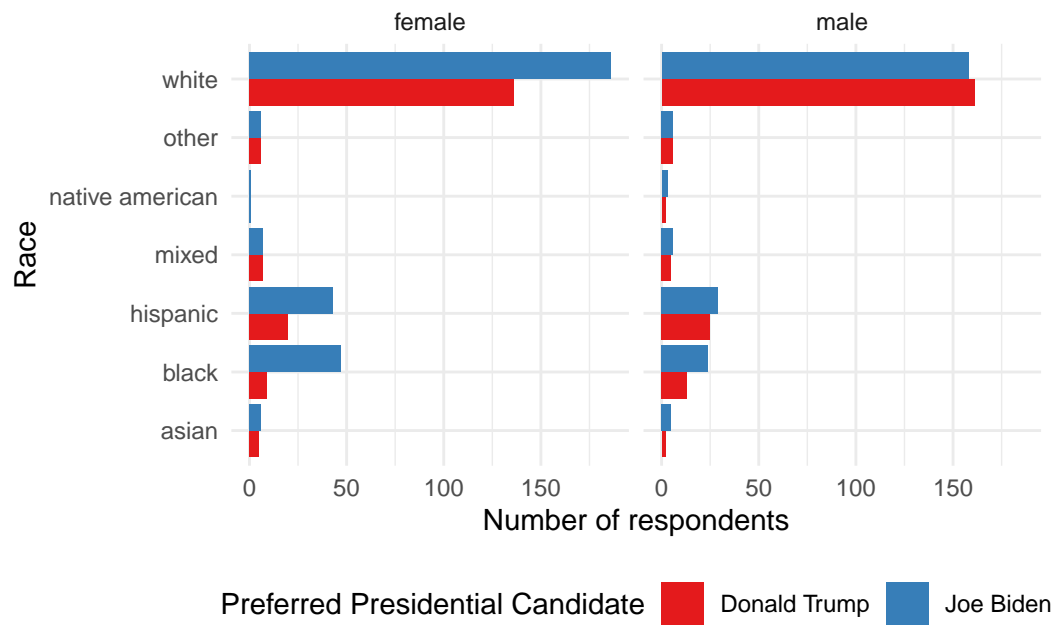


Figure 2: The Distribution of Presidential Preferences, by gender and race

2.1.1 Survey Data figures and tables go here

2.2 Post Stratification Data

Steven Ruggles and Schouweiler (2024)

Talk more about it.

And also planes (?@fig-planes). (You can change the height and width, but don't worry about doing that until you have finished every other aspect of the paper - Quarto will try to make it look nice and the defaults usually work well once you have enough text.)

Talk way more about it.

3 Model

logistic regression

binary

predicts support for trump or biden

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

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

The logistic regression used to predict ... is as follows:

$$y_i = \beta_0 + \beta_1 \times x_{sex} + \beta_2 \times x_{age_bracket} + \beta_3 \times x_{races} + \beta_3 \times x_{education_level} + \beta_4 \times x_{state} + \beta_5 \times x_{urban} + \beta_5 \times x_{\dots} \quad (1)$$

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

We use the `predict()` function in R (R Core Team (2023)) to apply our logistic regression model to the ACS Census data Steven Ruggles and Schouweiler (2024).

Model weaknesses: logistic regression output is binary so it does not include the options of not voting or voting for a third-party candidate.

3.1.1 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 θ .

4 Results

Our results are summarized in Table [1](#).

5 Discussion

5.1 Popular Vote Projection

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 Electoral College Projection

5.3 Polarization and America's Urban-Rural Divide

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Table 1: Explanatory models of presidential vote based on ...

	Presidential Vote Model
(Intercept)	1.75 (3.75)
sexmale	−0.61 (0.16)
age_bracket30-44	−0.52 (0.29)
age_bracket45-59	−0.69 (0.30)
age_bracket60+	−0.68 (0.26)
raceshispanic	0.00 (3.61)
racemixed	−0.30 (3.58)
racenative american	0.17 (3.76)
racesoother	−0.37 (3.66)
racesswhite	−0.36 (3.57)
race_black	1.14 (3.62)
education_level4-year	0.54 (0.29)
education_levelHigh school graduate	−0.16 (0.27)
education_levelNo HS	0.61 (0.52)
education_levelPost-grad	1.21 (0.33)
education_levelSome college	−0.01 (0.30)
stateAlaska	−0.85 (1.71)
stateArizona	−1.48 (0.98)
stateArkansas	−0.25 (1.15)
stateCalifornia	−0.26 (0.89)
stateColorado	−0.79 (1.04)
stateConnecticut	1.29 (1.55)
stateDelaware	−1.50 (1.45)
stateDistrict of Columbia	34.35 (31.62)
stateFlorida	−0.91 (0.91)

Appendix

A Additional data details

B Model details

B.1 Correlation Map

In `?@fig-ppcheckandposteriorvsprior-1` we implement a posterior predictive check. This shows...

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

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

Figure 3: `?(caption)`

Model plot (still need to fix this)

B.2 Diagnostics

`?@fig-stanareyouokay-1` is a trace plot. It shows... This suggests...

`?@fig-stanareyouokay-2` is a Rhat plot. It shows... This suggests...

Checking the convergence of the MCMC
algorithm

Figure 4: `?(caption)`

References

- Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. “Rstanarm: Bayesian Applied Regression Modeling via Stan.” <https://mc-stan.org/rstanarm/>.
- Iyengar, Shanto, Yphtach Lelkes, and Sean Westwood. 2024. *America’s Political Pulse*. <https://polarizationresearchlab.org/americas-political-pulse/>.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Steven Ruggles, Matthew Sobek, Sarah Flood, and Megan Schouweiler. 2024. *IPUMS USA: Version 15.0 [ACS 2022]*. Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D010.V15.0>.