My title*

My subtitle if needed

Jeongwoo Kim   Jiwon Choi

March 10, 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....

2 Data

We have used two datasets for this study. One is the U.S election survey data of Democracy Fund + UCLA Nationscape dataset from the Voter Study Group, conducted on October 3, 2019. Second is the census data from IPUMS America Census Service, which is used as the post-stratification data for the survey data to adjust the weight.

2.1 Survey Data

This survey data is an 18-month election study conducted by UCLA researchers with roughly 6250 online interviews each from from July 2019 to February 2021 (Tausanovitch and Vavreck (2020)). The sample is weighted to represent the U.S. adult population (Tausanovitch and Vavreck (2020)). Nationscape groups weight on the following important factors: gender, the four major census regions, race, Hispanic ethnicity, household income, education, age, language spoken at home, nativity, 2016 presidential vote, and the urban-rural mix of the respondent's ZIP code (Tausanovitch and Vavreck (2020)). According to the data, Male make up 48.3%

---

*Code and data are available at: https://github.com/Kjeongwoo99/STA302H_Paper3

while female make up 51.3% (Tausanovitch and Vavreck (2020)). 74.2% of the respondents are White, 6.8% are Asian/Pacific, 12% are Black (Tausanovitch and Vavreck (2020)). 20.4% are those between 18-29, 33.4% are 30-49, 32.4% are 50-69, 3.3% are 70+ (Tausanovitch and Vavreck (2020)). On average, 5.1 percent declined immediately among those who are selected for the survey. 16.7 percent of the respondents did not complete the survey. Another 5.9 percent were categorized as speeding or straight-line which means they completed the survey in less than 6 minutes or selected the same response for every question in the three policy question batteries. Leaving these out leave 72.4 percent of the original sample for the analysis.

The Nationscape survey's strength lies in its methodological rigor - the effectiveness in collecting large samples from the U.S. citizen and its weighting strategy desinged to mirror the U.S. adult population by including weight factors such as age, gender, race and income and more. As they filter out inaccurate or missing data, it makes sure that the data collected are accurate and ensures data integrity. While other datasets such as the General Social Survey (GSS) and the American National Election Studies (ANES) are available, the Nationscape dataset's frequency (surveys collected every week) give it an advantage in analyzing electoral trends and shifts in real-time. Its' extensive sample size also justifies the choice of this dataset.

For our analysis, we decided to focus on five demographics: age, gender, education, race and state. Age is important because in general, voters tend to become more conservative as they get older. To account for the age difference, we divided the age group into four categories: 18-29, 30-49, 50-69 and 70+.

Gender is also an important category because in general, men tend be more conservative and women tend to be more liberal. Recently, gender issues are growing social issues and this may affect the election, hence we wanted to explore how this affects our model.

Education is also an interesting factor. In the past, non-college white voters used to support Democrats while college-educated white voters supported Republicans (Harris (2018)). However, there has been a switch in this trend as 61 percent of non-college white voters showed their support wheres just 45 percent of college-educated white voters did in the exit polls (Harris (2018)). Only 37 percent of those without a degree cast their votes for Democrats while 53 percent with a degree did so (Harris (2018)). We categorized education into four categories: 'High school or less', 'Some college', 'College degree', 'Postgrad'.

Race also needs some attention because normally non-white groups are highly in favour of Democrats regardless of candidates and white swing by depending on candidates. According to the statistics collected in 2016, 93% of black, 71% of Latino, 68% of Asian support democrats while only 41% of white support democrats (Prokop (2021)). As white voters make up 74% of the voting population, it is really important for both parties to attain this demographic group.

Lastly, states are very important as some states historically favor conservatives while some states vote for democrats. In general, the west and the east coasts are democrat supporters whereas south are conservative supporters.

2.2 Post-stratification Data

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.)
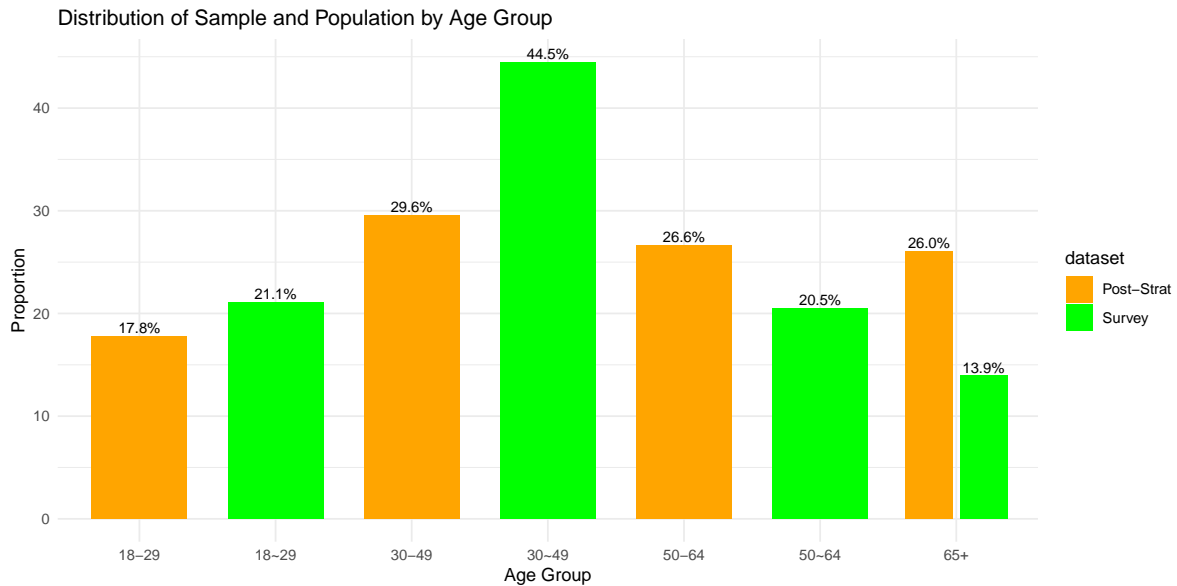
Distribution of Sample and Population by Age Group

Figure 1: Distribution of Sample and Population by Age Group

Table 1: Voters Intention to Support Trump

Response	Number of Respondents	Proportion (%)
Yes	2052	34.39
No	3290	55.15
Other	624	10.46

Table 2: Voters Intention of Their Primary Party

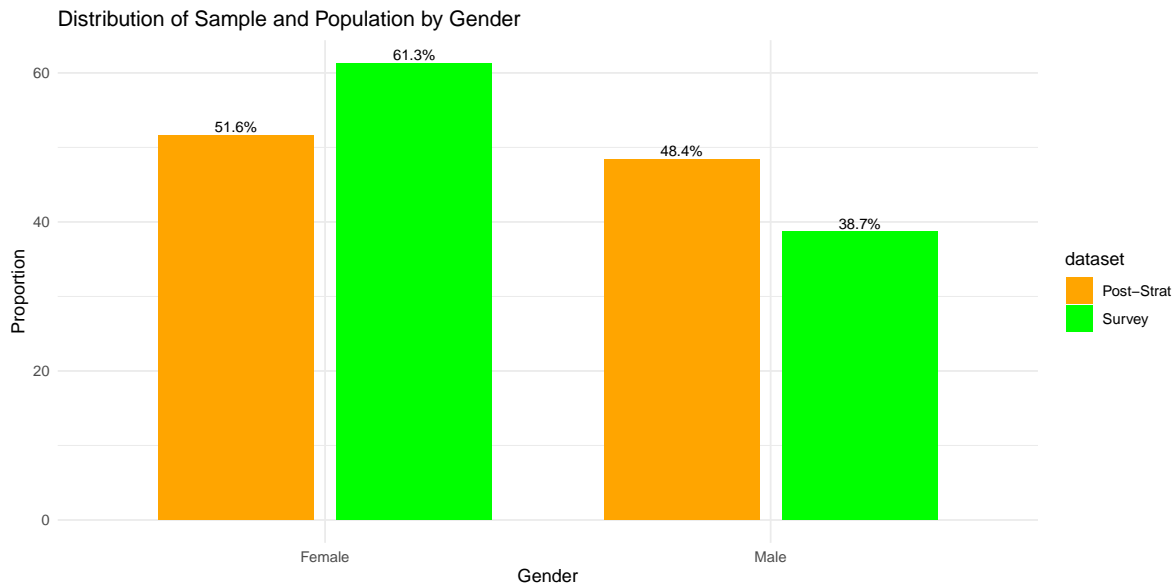| Party Preference	Number of Respondents	Proportion (%)
Democratic	2349	39.37
Republican	1657	27.77
Other	1960	32.85

Figure 2: Distribution of Sample and Population by Gender

3 Model

The goal of our modelling strategy is twofold. Firstly,…

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.

$$y_i|\mu_i,\sigma \sim \text{Normal}(\mu_i,\sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0,2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0,2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0,2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

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`.

4 Results

Our results are summarized in **?@tbl-modelresults**.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be 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

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

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)**

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/.

Harris, Adam. 2018. "AMERICA IS DIVIDED BY EDUCATION." *The Atlantic*. https://www.theatlantic.com/education/archive/2018/11/education-gap-explains-american-politics/575113/.

Prokop, Andrew. 2021. "A New Report Complicates Simplistic Narratives about Race and the 2020 Election." *Vox*, May. https://www.vox.com/2021/5/10/22425178/catalist-report-2020-election-biden-trump-demographics.

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

Tausanovitch, Chris, and Lynn Vavreck. 2020. *Democracy Fund + UCLA Nationscape*. October 10-17, 2019 (version 20200814). https://www.voterstudygroup.org/publication/nationscape-data-set.

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.