Final Report: Election Prediction

Alexander Bendeck, Lynn Fan, Cathy Lee, Alice Liao, Justina Zou

28 October 2020

Introduction

Data Description

Senate poll/House poll: from https://data.fivethirtyeight.com/ States_cov_matrix: from the Economist paper Economist prediction polls: from the Economist paper (pres) partisan lean (prior for pres): from 538 Abramowitz data: from the Economist paper (pres)

Exploratory Data Analysis

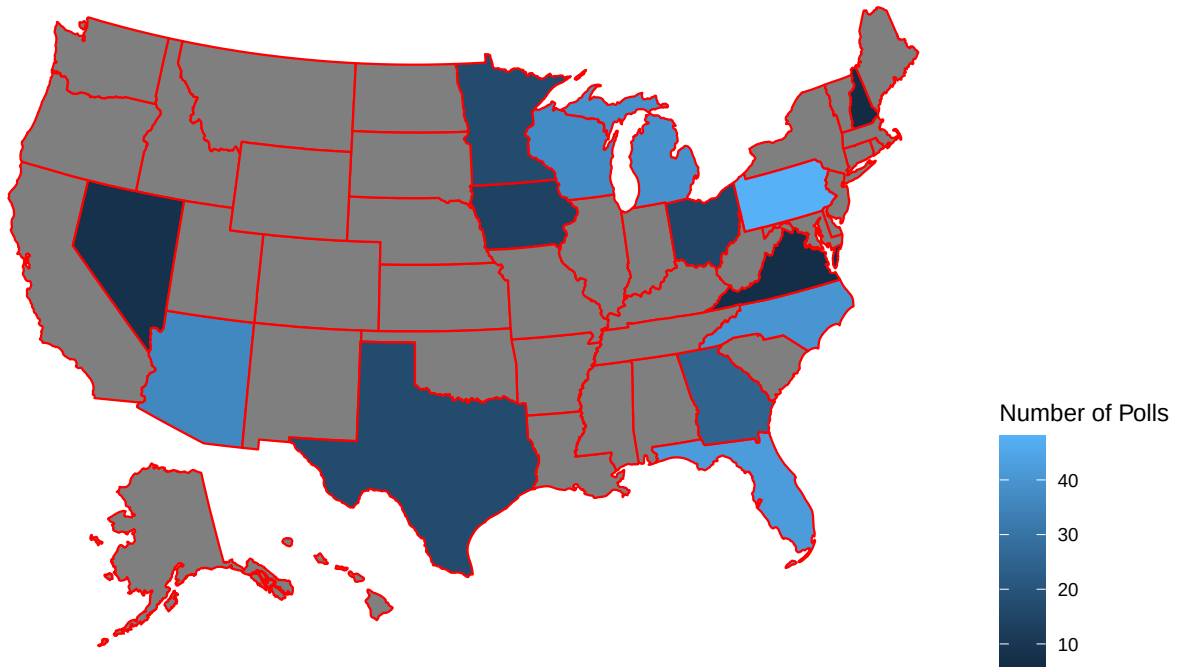
```
pres_eda_df = readRDS("cleaned_pres_polls.rds")
econ_eda_df = readRDS("cleaned_econ_polls.rds")
econ_state_df = econ_eda_df %>%
  group_by(state) %>%
  summarise(n=n(),avg=mean(y))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

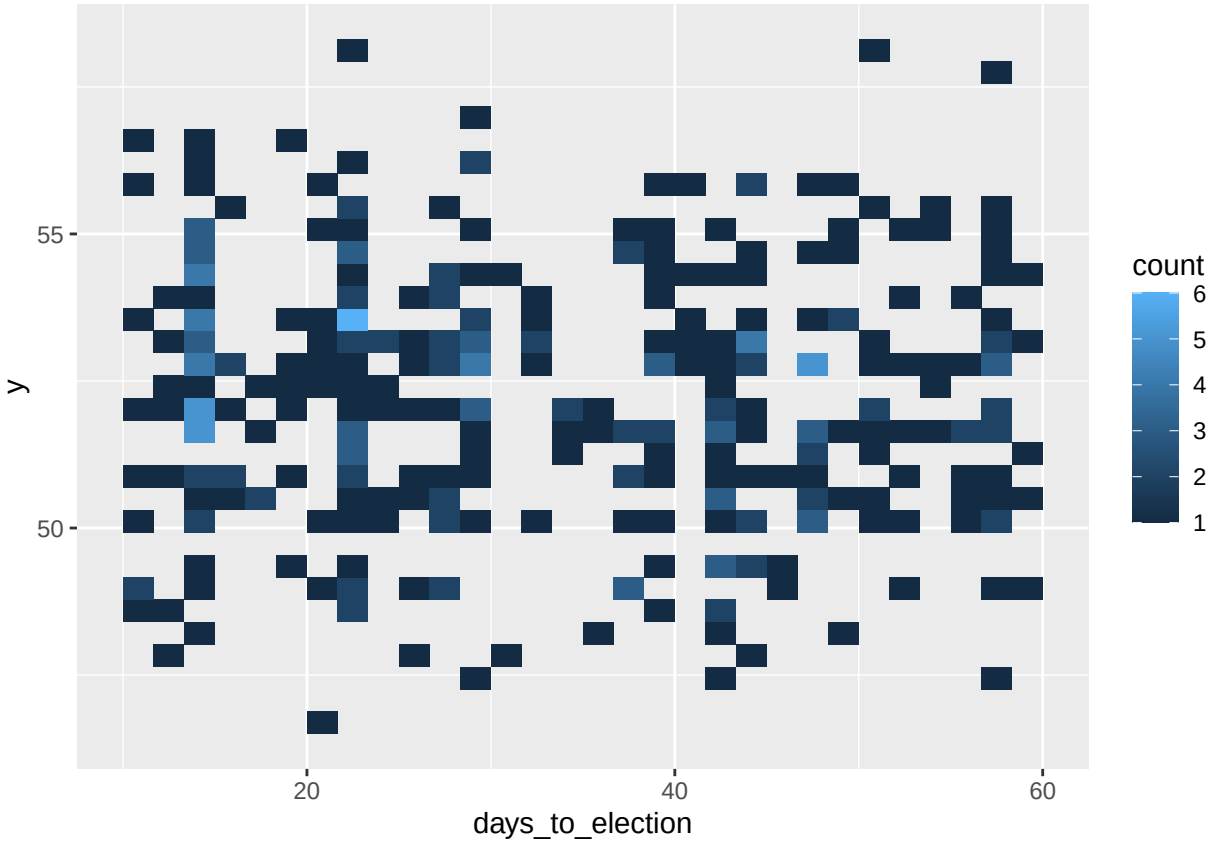
```
house_eda_df = readRDS("cleaned_house_polls.rds")
senate_eda_df = readRDS("cleaned_senate_polls.rds")
senate_state_df = senate_eda_df %>%
  group_by(state) %>%
  summarise(n=n(),avg=mean(y))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

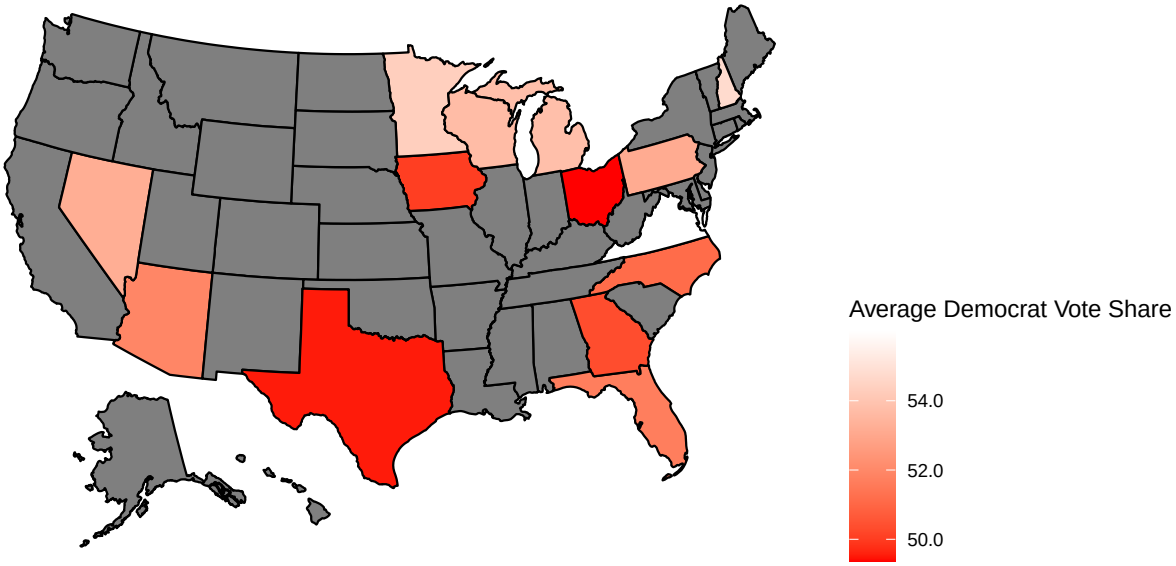
```
#number of polls for each state in the economist paper data set
plot_usmap(data = econ_state_df, values = "n", color = "red") +
  scale_fill_continuous(name = "Number of Polls", label = scales::comma) +
  theme(legend.position = "right")
```



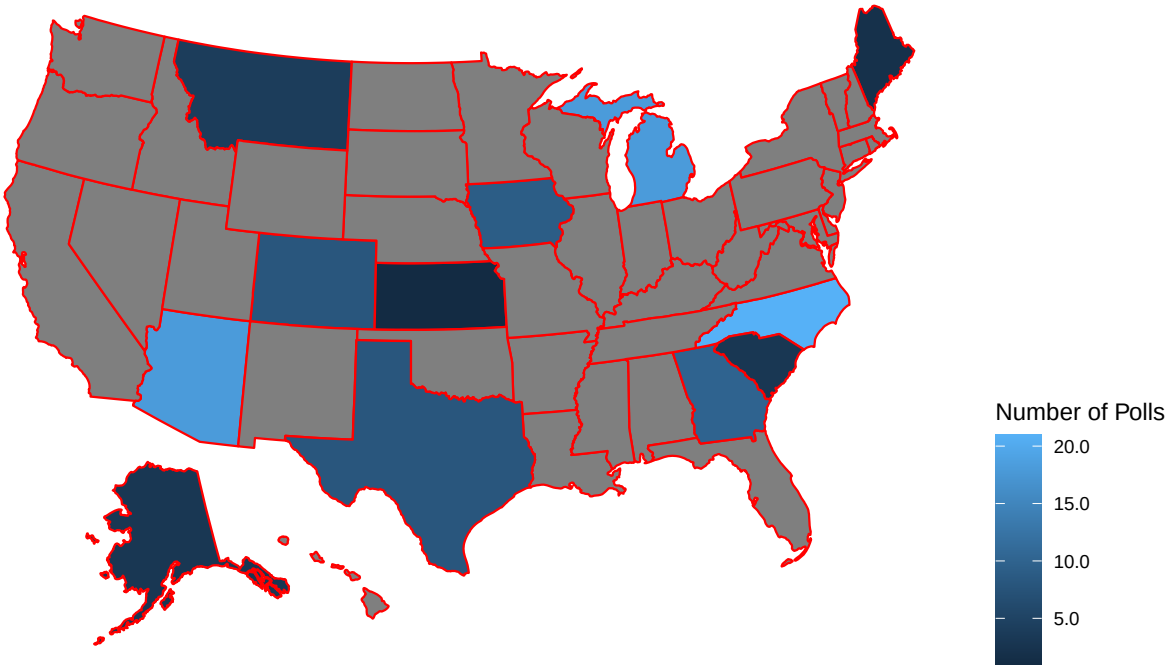

```
#democrat percentage in polling data
ggplot(econ_eda_df, aes(days_to_election, y)) + geom_bin2d()
```



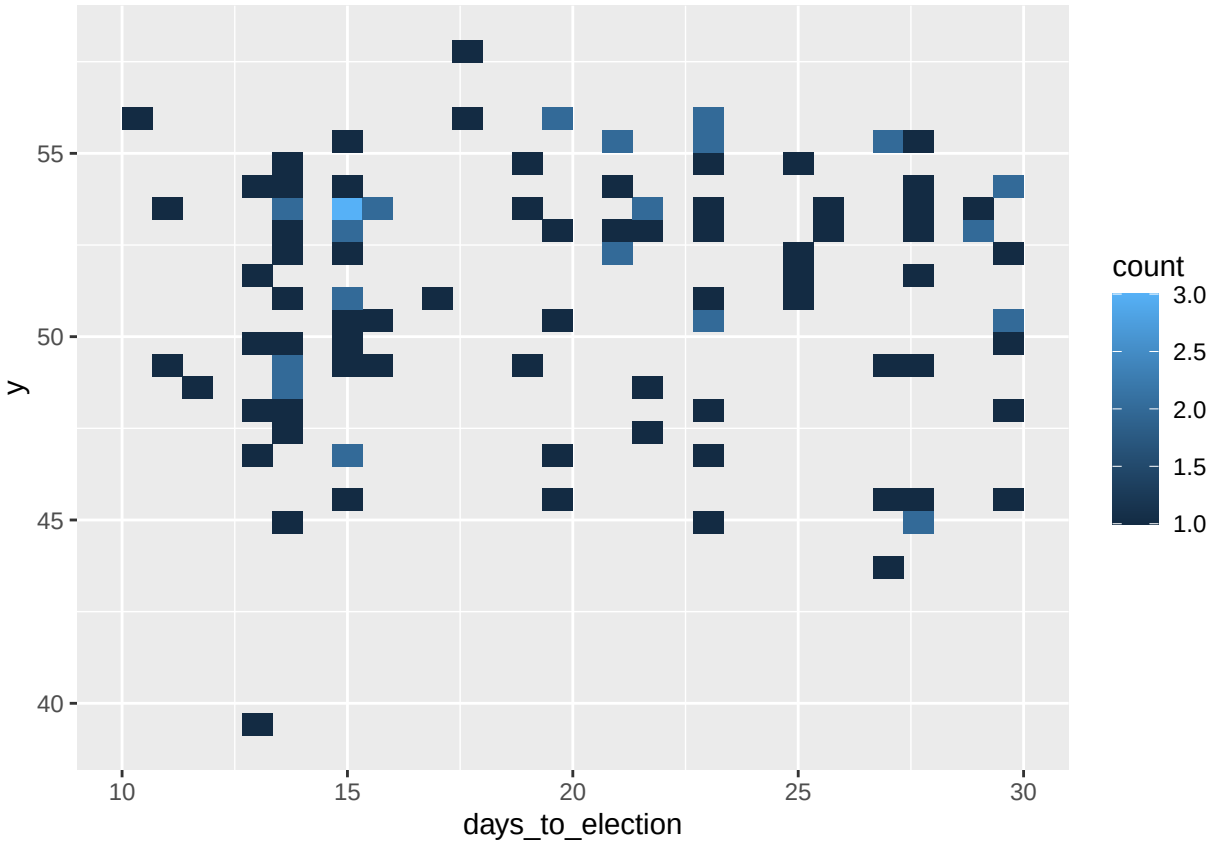

```
#democrat percentage by states
plot_usmap(data = econ_state_df, values = "avg", color = "black") +
  scale_fill_continuous(low = "red", high = "white",name = "Average Democrat Vote Share", label = sc
  theme(legend.position = "right")
```



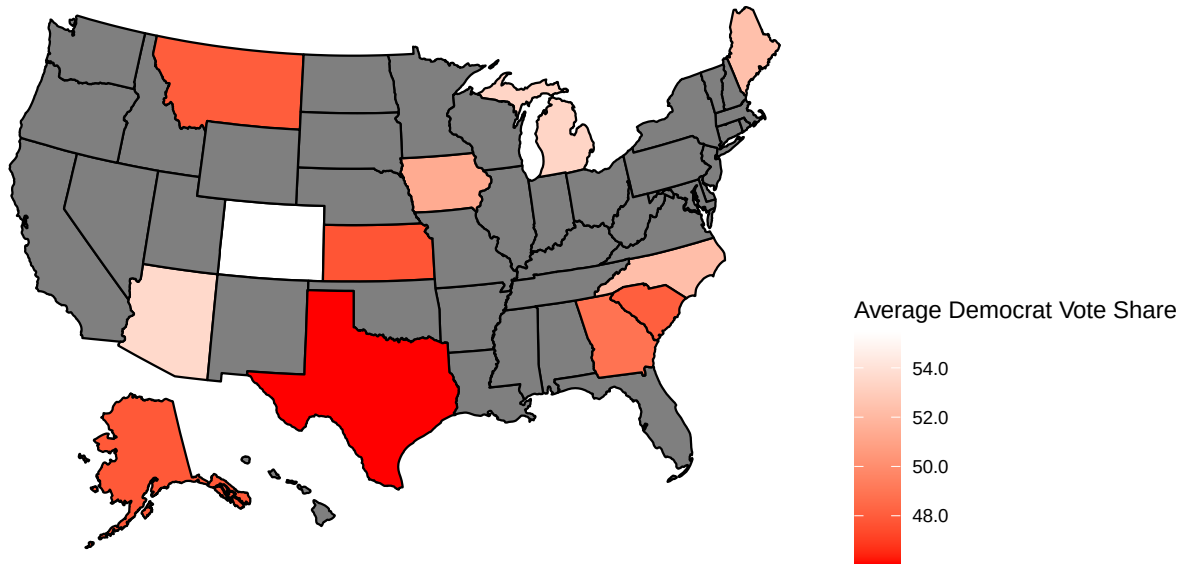

```
#number of polls for senate
plot_usmap(data = senate_state_df, values = "n", color = "red") +
  scale_fill_continuous(name = "Number of Polls", label = scales::comma) +
  theme(legend.position = "right")
```



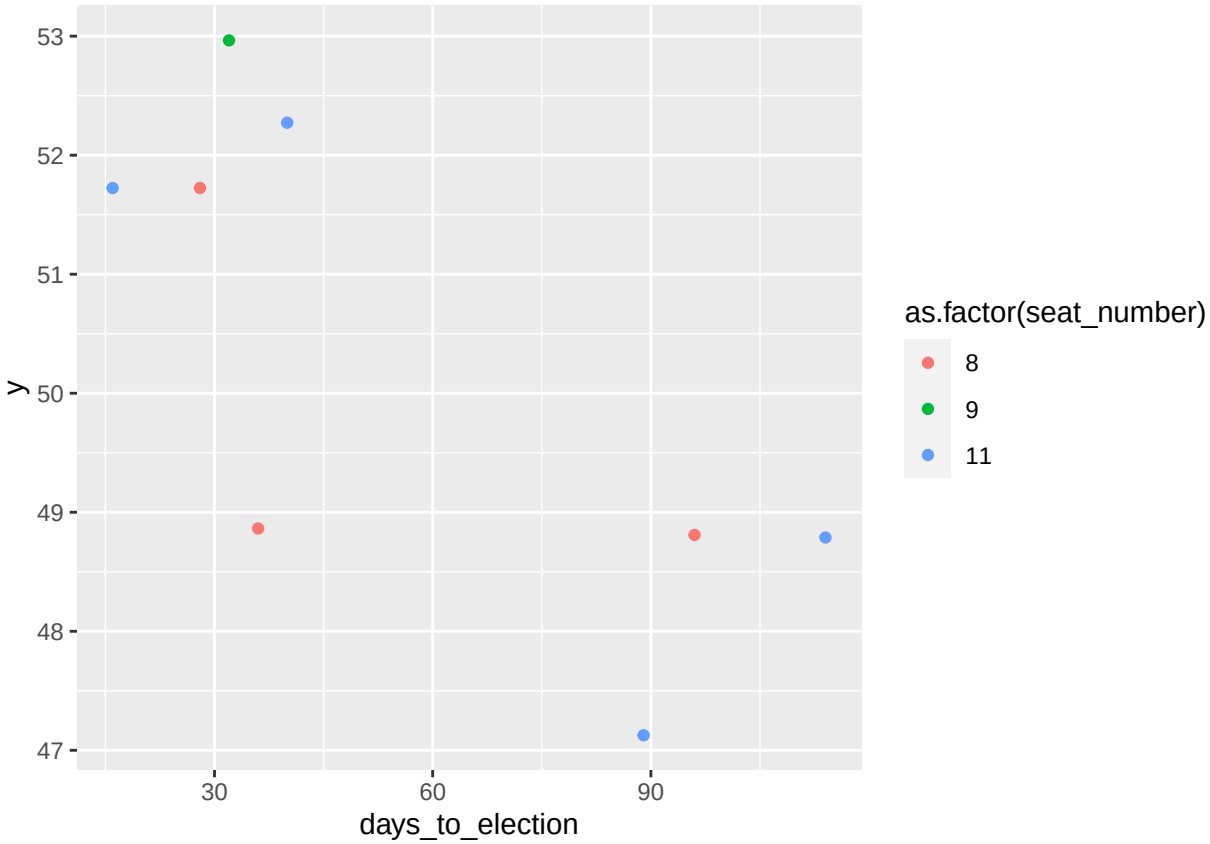

```
#biden percentage in polling data
ggplot(senate_eda_df, aes(days_to_election, y)) + geom_bin2d()
```




```
#democrat percentage by states
plot_usmap(data = senate_state_df, values = "avg", color = "black") +
  scale_fill_continuous(low = "red", high = "white",name = "Average Democrat Vote Share", label = sc
  theme(legend.position = "right")
```




```
#biden percentage in polling data
ggplot(house_eda_df, aes(x=days_to_election, y=y,color=as.factor(seat_number))) + geom_point()
```



Methods

Presidential Election Model

The response we will use is Y_k, Biden's share of the two-party vote using the following model where the poll k (of K=*fill in* polls) ended t days (of $T = 60$ days) before the election and was conducted on state j (of J=14 states of interest).

$Y_k \sim N(\theta_{jt}, \sigma^2_{yj})$
$\theta_{\cdots t} \sim MVN(\theta_{\cdots t-1}, \Sigma)$
$\Sigma \sim Wishart(S, J+1)$ where S is the state covariance matrix and J is the number of states in the model.
$\sigma^2_{yj} \sim InvGamma(\nu_u, \nu_y \tau_y)$
$\nu_y \sim Uniform(0, 100)$ and $\tau_y \sim Uniform(0, 100)$.
$\theta_{j1} \sim N(\mu_j, \sigma^2)$
$\sigma^2 \sim InvGamma(0.5, 0.5)$
$\mu_j \sim N(h_j, 7.5^2)$
$h_j = 0.1\text{Presidental Fundamentals}_{2020} + 0.9 * \text{Vote Share from Partisan Lean}_j$
$\text{Presidental Fundamentals}_i = \gamma_0 + \gamma_1 \text{June Approval Ratings for Incumbent Party}_i + \gamma_2 \text{Three Month Stock Growth}_i + \gamma_3 \text{2nd Quarter Real Income Growth}_i + \epsilon_i$
$\epsilon_i \sim N(0, \phi^2)$

where explain variables later the S&P stock growth three months prior to the election and also explain i where i is each year in that Abramowitz data.

explain how partisan lean was calculated from 538 and how it was turned into Vote Share from Partisan lean

Senate Election Model

The response we will use is Y_k, the Democrat candidate's share of the two-party vote using the following model where the poll k (of K=*fill in* polls) ended t days (of $T = 60$ days) before the election and was conducted on state j (of J=14 states of interest).

$Y_k \sim N(\beta_{jt}, \sigma^2_{yj})$
$\beta_{jt} \sim N(\beta_{jt-1}, \sigma^2_{\beta j})$
$\sigma^2_{yj} \sim InvGamma(\nu_u, \nu_y \tau_y)$
$\nu_y \sim Uniform(0, 100)$ and $\tau_y \sim Uniform(0, 100)$.
$\sigma^2_{\beta j} \sim InvGamma(\nu_\beta, \nu_\beta \tau_\beta))$
$\nu_\beta \sim Uniform(0, 100)$ and $\tau_\beta \sim Uniform(0, 100)$.
$\beta_{j1} \sim N(\mu_j, \sigma^2)$
$\sigma^2 \sim InvGamma(0.5, 0.5)$
$\mu_j \sim N(h_j, 7.5^2)$
$h_j = \text{Vote Share from Partisan Lean}_j + \text{Incumbency Advantage}_j$

House Election Model

The response we will use is Y_k, the Democrat candidate's share of the two-party vote using the following model where the poll k (of K=*fill in* polls) ended t days (of $T = 60$ days) before the election and was conducted on district j (where j is in districts $1, \cdots, 11, 13$ of North Carolina). For district 12, the vote share is coded as 100 because there is only one candidate, and she is a Democrat.

$Y_k \sim N(\beta_{jt}, \sigma^2_{yj})$
$\beta_{jt} \sim N(\beta_{jt-1}, \sigma^2_{\beta j})$
$\sigma^2_{yj} \sim InvGamma(\nu_u, \nu_y \tau_y)$
$\nu_y \sim Uniform(0, 100)$ and $\tau_y \sim Uniform(0, 100)$.
$\sigma^2_{\beta j} \sim InvGamma(\nu_\beta, \nu_\beta \tau_\beta))$
$\nu_\beta \sim Uniform(0, 100)$ and $\tau_\beta \sim Uniform(0, 100)$.
$\beta_{j1} \sim N(h_j, 7.5)$
$h_j = 0.9 * \text{Vote Share from Partisan Lean}_j + 0.1 * \text{Expected Vote Share from Voter Turnout}_j$

Expected Vote Share from Voter Turnout$_j$ was calculated from the model in the Interim Report *insert citation whatever that ends up being.* Please see Appendix A for more details.

explain how partisan lean was calculated from 538 and how it was turned into Vote Share from Partisan lean

Results

Model Validation and Sensitivity Checks

Diagnostics

Appendix A

Appendix B

Appendix C (plots and other miscellaneous output)

References

1. Linzer, D. A. (2013). Dynamic Bayesian Forecasting of Presidential Elections in the States. Journal of the American Statistical Association, 108(501), 124-134. doi:10.1080/01621459.2012.737735

2. Hansford, T. G., & Gomez, B. T. (2010). Estimating the Electoral Effects of Voter Turnout. American Political Science Review, 104(2), 268-288. doi:10.1017/s0003055410000109

3. 2020 Election. (2020, October 20). Retrieved October 20, 2020, from https://fivethirtyeight.com/politics/elections/

4. Park, D. K., Gelman, A., & Bafumi, J. (2006). State-Level Opinions from National Surveys:. Public Opinion in State Politics, 209-228. doi:10.2307/j.ctvr33bdg.17

5. Mahler, V. A., Jesuit, D. K., & Paradowski, P. R. (2013). Electoral Turnout and State Redistribution. Political Research Quarterly, 67(2), 361-373. doi:10.1177/1065912913509306

6. Uhlaner, C. J., & Scola, B. (2015). Collective Representation as a Mobilizer. State Politics & Policy Quarterly, 16(2), 227-263. doi:10.1177/1532440015603576

7. Godbout, J. (2012). Turnout and presidential coattails in congressional elections. Public Choice, 157(1-2), 333-356. doi:10.1007/s11127-012-9947-7

8. Kim, S. S., Alvarez, R. M., & Ramirez, C. M. (2020). Who Voted in 2016? Using Fuzzy Forests to Understand Voter Turnout. doi:10.33774/apsa-2020-xzx29

9. Weinschenk, A. C. (2019) That's Why the Lady Lost to the Trump: Demographics and the 2016 Presidential Election, Journal of Political Marketing, 18:1-2, 69-91, DOI: 10.1080/15377857.2018.1478657

10. Charles, K. K., & Stephens, M. (2011). Employment, Wages and Voter Turnout. doi:10.3386/w17270

11. Hills, M. (2020, September 25). US election 2020: A really simple guide. Retrieved October 20, 2020, from https://www.bbc.com/news/election-us-2020-53785985

12. Railey, K. (2016). Federal Judges Let Stand North Carolina,'s New Congressional Map. The Hotline. https://link.gale.com/apps/doc/A498010836/ITOF?u=duke_perkins&sid=ITOF&xid=119d6ad9

13. Perrin, A. J., & Ifatunji, M. A. (2020). Race, Immigration, and Support for Donald Trump: Evidence From the 2018 North Carolina Election. Sociological Forum, 35(S1), 941-953. doi:10.1111/socf.12600

14. Redistricting in North Carolina. (2020). Retrieved October 21, 2020, from https://ballotpedia.org/Redistricting_in_North_Carolina

15. Â§ 132-1. Public Records. https://www.ncleg.gov/EnactedLegislation/Statutes/PDF/BySection/Chapter_132/GS_132-1.pdf

16. Â§ 163-82.10. Official record of voter registration. https://www.ncleg.gov/EnactedLegislation/Statutes/PDF/BySection/Chapter_163/GS_163-82.10.pdf

17. FairVote.org. (n.d.). Voter Turnout. Retrieved October 22, 2020, from https://www.fairvote.org/voter_turnout

18. McDonald, M. P. (2020). Voter Turnout Demographics. Retrieved October 23, 2020, from http://www.electproject.org/home/voter-turnout/demographics

19. Y. Ghitza, A. Gelman's (2013). Deep interactions with MRP: Election Turnout and Voting Patterns

Among Small Electoral Subgroups. American Journal of Political Science 57 762– 776.

20. Walker, H.L., Herron, M.C. & Smith, D.A. Early Voting Changes and Voter Turnout: North Carolina in the 2016 General Election. Polit Behav 41, 841–869 (2019). https://doi-org.proxy.lib.duke.edu/10.1007/s11109-018-9473-5

21. Pew Research Center (2020). Men and women in the U.S. continue to differ in voter turnout rate, party identification. https://www.pewresearch.org/fact-tank/2020/08/18/men-and-women-in-the-u-s-continue-to-differ-in-voter-turnout-rate-party-identification/

22. Pew Research Center (2016). Party affiliation among voters: 1992-2016. https://www.pewresearch.org/politics/2016/09/13/2-party-affiliation-among-voters-1992-2016/

23. J. Misra (2019). Voter Turnout Rates Among All Voting Age and Major Racial and Ethnic Groups Were Higher Than in 2014. https://www.pewresearch.org/fact-tank/2020/08/18/men-and-women-in-the-u-s-continue-to-differ-in-voter-turnout-rate-party-identification/

24. Killian, L. (2012). Five Myths About Independent Voters. Retrieved October 22, 2020, from https://www.washingtonpost.com/opinions/five-myths-about-independent-voters/2012/05/17/gIQAZmGyWU_story.html

25. Igielnik, R. (2020). Men and Women in the U.S. Continue to Differ in Voter Turnout Rate. Retrieved October 22, 2020, from https://www.pewresearch.org/fact-tank/2020/08/18/men-and-women-in-the-u-s-continue-to-differ-in-voter-turnout-rate-party-identification/

26. Akee, R. (2019). Voting and Income. Retrieved October 22, 2020, from https://econofact.org/voting-and-income