Interim Report: Who Votes in NC?

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

20 October 2020

Introduction

The United States is getting closer to the 2020 Congressional Elections on November 3, 2020. With the current polarizing political landscape, the congressional election outcomes are significant to determine the next stage of this country. As it becomes critical for statisticians to help build forecasting models to predict the election outcomes, we need to first understand the patterns of voter turnout. Voting has special importance in democratic systems, but only about half of the eligible U.S. citizenry votes, and there are real political consequences when voters differ systematically from nonvoters (Uhlaner et al.). There has been abundant literature proving that variation in voter turnout will have electoral consequences (Hansford, et al.), in a number of ways. First of all, the media conventional wisdom is that “higher turnout benefits Democrats,” although there has been mixed evidence about this theory (Weinschenk, 2019). Second, literature has proven certain demographic factors to statistically significantly benefit one party over the other, in both congressional elections and presidential races. For example, gender, race and party registration could help profile voting patterns for congressional elections (Uhlaner et al.). Election prediction models need the baseline population of voters to predict the potential outcomes, and the demographic composition of voters will directly determine the forecasting results.
Among all states, North Carolina has been as a swing state in presidential and congressional elections for decades. In 2008, Obama won the state narrowly, but lost it narrowly after 4 years in 2012. Since 1996, the Republican statewide vote share in congressional elections has varied “from a low of 45% in 2008 to a high of 55% in 2014 (Perrin et al.).” It makes North Carolina an interesting battleground in which voter demographic changes could potentially lead to significant implications of election outcomes and “an excellent site for those interested in partisan voting trends (Perrin et al.).” This report seeks to understand the voter turnout of North Carolina for 2020 NC Congressional Elections, predicting who will vote in 2020.

Data Description

We are using public data provided by the NC State Board of Elections, which can all be accessed directly at the link https://dl.ncsbe.gov/list.html. The database contains voter history information for elections within the past 10 years in the ncvhis files, and all legally available voter specific information ion the ncvoter files. The ncvoter files contain point-in-time snapshot voter registration data. For privacy concerns, names, birth dates and drivers license are not included, but the two types of files could be matched by North Carolina identification (NCID) number. The database was last updated on September 9, 2020. While we understand that voters might register later than that as the voter registration deadline for North Carolina is October 9, 2020, we believe it is sufficient to

represent the majority of NC potential voters.
From the ncvhis files, we only kept the voters that voted for the 2016 general election for our analysis. Studies have shown that presidential elections help mobilize voters, so voter turnout in presidential election years are significantly higher. In recent elections, voter turnout during presidential election years is around 60%, and only about 40% during midterm elections (FairVote.org). For North Carolina, voter turnout data in 2018 is also inappropriate to use because neither of North Carolina's U.S. senators nor the governor was up for reelection, further demotivated voters (Perrin et al.). From the ncvoter files, we filtered demographic factors that are supported by existing literature to be significant in understanding voting patterns, including gender, race, party registration, and age (Kim et al.). We also have their county and congressional district information available.
Additionally, we found relevant literature proving the relationship between voter turnout and wage (Charles et al.), so we found county-level median household income data from Economic Research Service under United States Department of Agriculture (https://www.ers.usda.gov/data-products/county-level-data-sets)

Data Munging

TO DO: more justification on grouping

After binding ncvhis files and ncvoter files by NCID and binding NC median household income by county, we started to process data for analysis. First of all, we identified those data points older than 116 years old and removed them as the oldest person in NC is 116 years old and anyone older should be wrong data points. Many data points are also missing congressional district information. We imputed the missing districts by matching the voter's registered county with congressional district. We removed the 4% of voters who reside in counties that span across more than one county. In the combined data set, there are party registrations for all parties, including The Libertarian Party and The Green Party. Because we are interested primarily only in the Republican Party and the Democratic Party and there are concerning class imbalance issues as the two parties take up the majority of registered voter population, we binded other parties as the third category `Other` for `Party`. Similarly, because of class imbalance, we binded the races other than White and African Americans as `Other` for `Race` as well. For those missing `Gender` information, we binded them with `Unspecified`.
Because we have eight million data points available, running models in a one-line-per-voter data set will be very computationally expensive. We instead decided to group data points by gender, race, party registration, county median income, and age, so that we can run models for the data set in a collapsed format. We divided (1) median county household income into four levels by the 25th, 50th, and 75th quantiles; (2) age into four levels for 18-29, 30-44, 45-59, and older than 60 years old, as it is a common way to analyze voter ages (McDonald, 2020); (3) gender into three categories, Female, Male and Other, and (4) race into three categories, Black, White, and Other.

EDA

Method

We will take a Bayesian approach to not only predict if a voter with a certain profile would vote, but also understand quantitatively how the geographic and demographic information of a registered voter is associated with his or her likelihood of actually casting a ballout. To model the binary outcome (vote vs not vote), we will first fit a simple logistic regression model with selected variables as a baseline for comparison. Then motivated by Y. Ghitza and A. Gelman's idea of grouping

(2013), we divide the population into mutually exclusive categories according to their demographic and geographic characteristics and fit a Bayesian model with group-level predictors as well as their interactions. With poststratification we can get average estimates for each of the subgroups.

add some Bayesian justification add priors (look at sensitivity analysis rmd)

The model takes the following form: *add latex? check Amy's slide* https://amy-herring.github.io/STA440/decks/glmm_01_deck.html#/section-18

$$logit(Vote) = \beta_0 + \beta_1 I(Median\ Income\ > 64,509) + ...$$

In a later section, we will compare this Bayesian model with two additional models: one is a frequentist logistic regression model with the same predictors and interactions and the other a similar Bayesian model with additional random effect at the congressional district level. In this way we hope to assess if the Bayesian framework is superior than a frequentist approach when predicting voter turnout and if there is any salient unexplained variation within each congressional district. *fit with whole dataset and run 5-cv for the main model.*

TODO: talk about interactions – lit review justification

```
binary_model <-
  brm(data = voter_grouped, family = binomial,
      votes | trials(n) ~ 1 + med_inc_binned + gender_code + race_code + age_binned + party_cd
      iter = 2500, warmup = 500, cores = 2, chains = 2,
      seed = 10)
```

```
## Compiling Stan program...
```

```
## Trying to compile a simple C file
```

```
## Running /Library/Frameworks/R.framework/Resources/bin/R CMD SHLIB foo.c
## clang -mmacosx-version-min=10.13 -I"/Library/Frameworks/R.framework/Resources/include" -DNDI
## In file included from <built-in>:1:
## In file included from /Users/cathylee/Library/R/4.0/library/StanHeaders/include/stan/math/p
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Dense:1
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Core:88
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/src/Core/util/Macros.h:613:1:
## namespace Eigen {
## ^
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/src/Core/util/Macros.h:613:16
## namespace Eigen {
##                ^
##                ;
## In file included from <built-in>:1:
## In file included from /Users/cathylee/Library/R/4.0/library/StanHeaders/include/stan/math/p
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Dense:1
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Core:96:10: fatal error: 'com
## #include <complex>
##          ^~~~~~~~~
## 3 errors generated.
## make: *** [foo.o] Error 1
```

```
## Start sampling
```

```
#summary(binary_model)
#saveRDS(binary_model, "grouped_model_no_randeff_whole_dataset.rds")
```

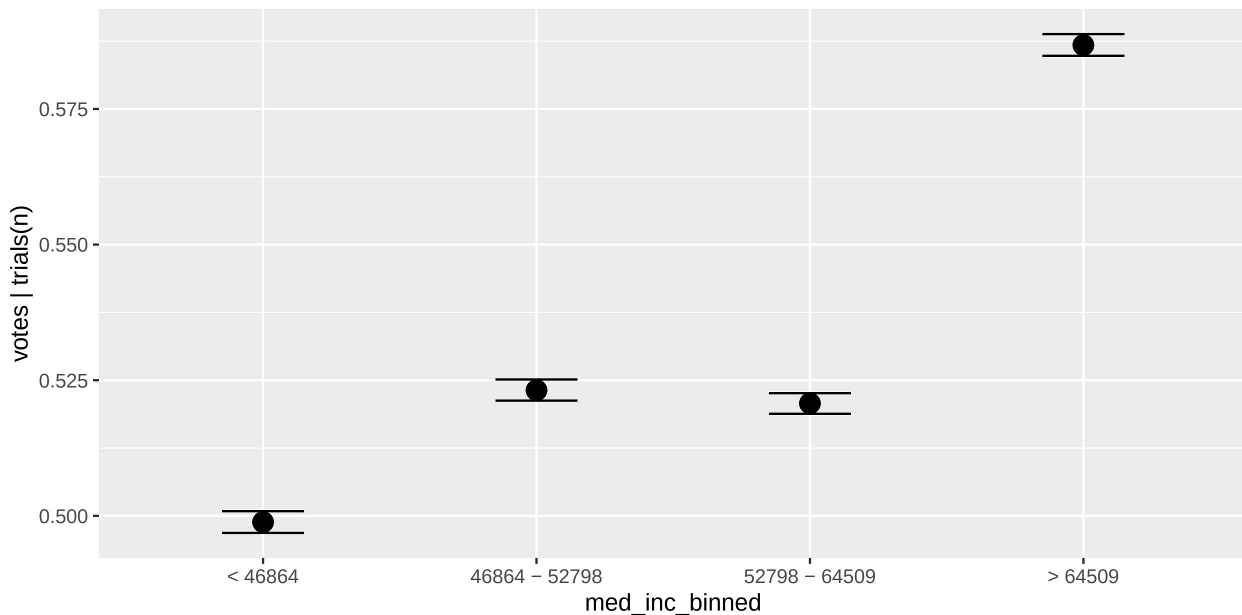
Results and Interpretations

TO DO: try to make this side by side

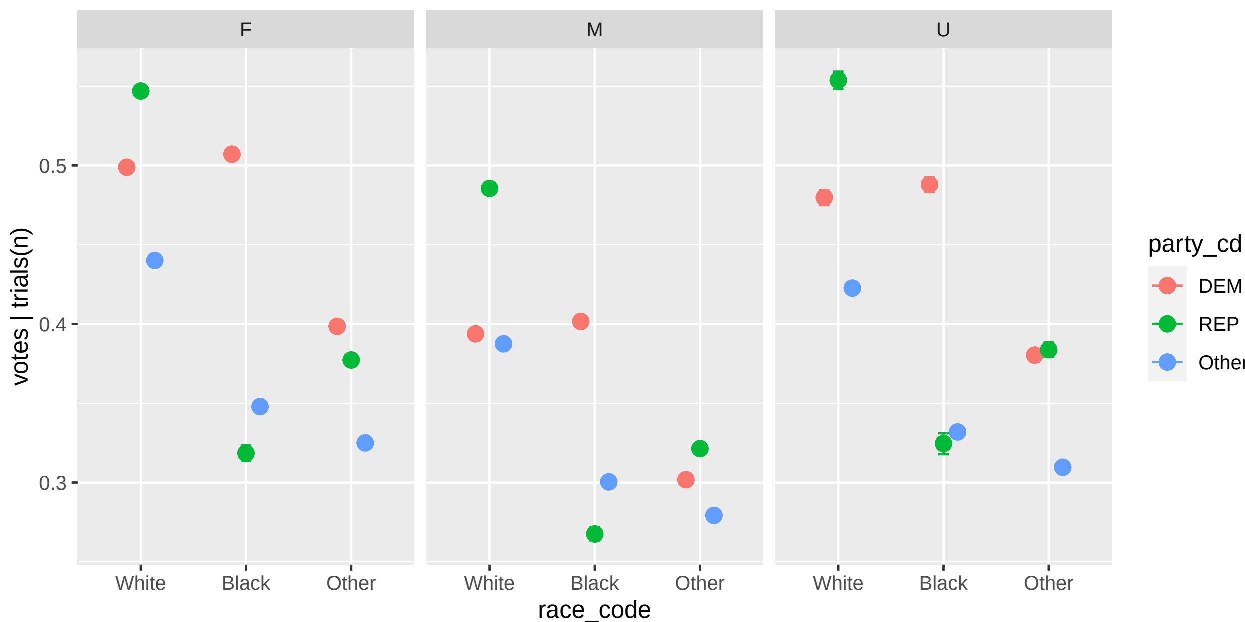
	Estimate	Std. Error	2.5% Quantile	97.5% Quantile
Intercept	0.00	0.00	-0.01	0.00
Median Income > 64,509	0.10	0.00	0.09	0.10
Median Income 46,864-52,798	0.09	0.00	0.08	0.09
Median Income 52,798-64,509	0.36	0.00	0.35	0.36
Gender Male	-0.43	0.00	-0.44	-0.42
Gender Unspecified	-0.08	0.01	-0.09	-0.06
Race Black	0.03	0.00	0.03	0.04
Race Other	-0.41	0.00	-0.42	-0.40
Age 30-44	0.62	0.00	0.61	0.63
Age 45-59	1.08	0.00	1.07	1.09
Age 60+	0.94	0.00	0.93	0.95
Party Republican	0.19	0.01	0.18	0.20
Party Other	-0.24	0.00	-0.25	-0.23
Gender Male:Party Republican	0.18	0.00	0.17	0.19
Gender Unspecified:Party Republican	0.10	0.01	0.08	0.13
Gender Male:Party Other	0.21	0.00	0.20	0.22
Gender Unspecified:Party Other	0.01	0.01	-0.01	0.02
Race Black:Party Republican	-0.98	0.01	-1.00	-0.96
Race Other:Party Republican	-0.28	0.01	-0.30	-0.27
Race Black:Party Other	-0.42	0.01	-0.43	-0.41
Race Other:Party Other	-0.08	0.01	-0.10	-0.07
Gender Male:Age 30-44	0.02	0.00	0.01	0.03
Gender Unspecified:Age 30-44	-0.41	0.01	-0.43	-0.40
Gender Male:Age 45-59	0.11	0.00	0.10	0.12
Gender Unspecified:Age 45-59	-0.60	0.01	-0.62	-0.58
Gender Male:Age 60+	0.25	0.00	0.24	0.26
Gender Unspecified:Age 60+	-0.47	0.01	-0.50	-0.45
Age 30-44:Party Republican	0.04	0.01	0.03	0.05
Age 45-59:Party Republican	-0.05	0.01	-0.06	-0.04
Age 60+:Party Republican	-0.06	0.01	-0.07	-0.05
Age 30-44:Party Other	0.05	0.01	0.04	0.06
Age 45-59:Party Other	0.01	0.01	0.00	0.02
Age 60+:Party Other	0.27	0.01	0.25	0.28

TODO: talk about small SE (no identifiability issues!! goodness of fit) and say which are significant (credible interval doesn't contain 0), "other" level is not super informative anyway

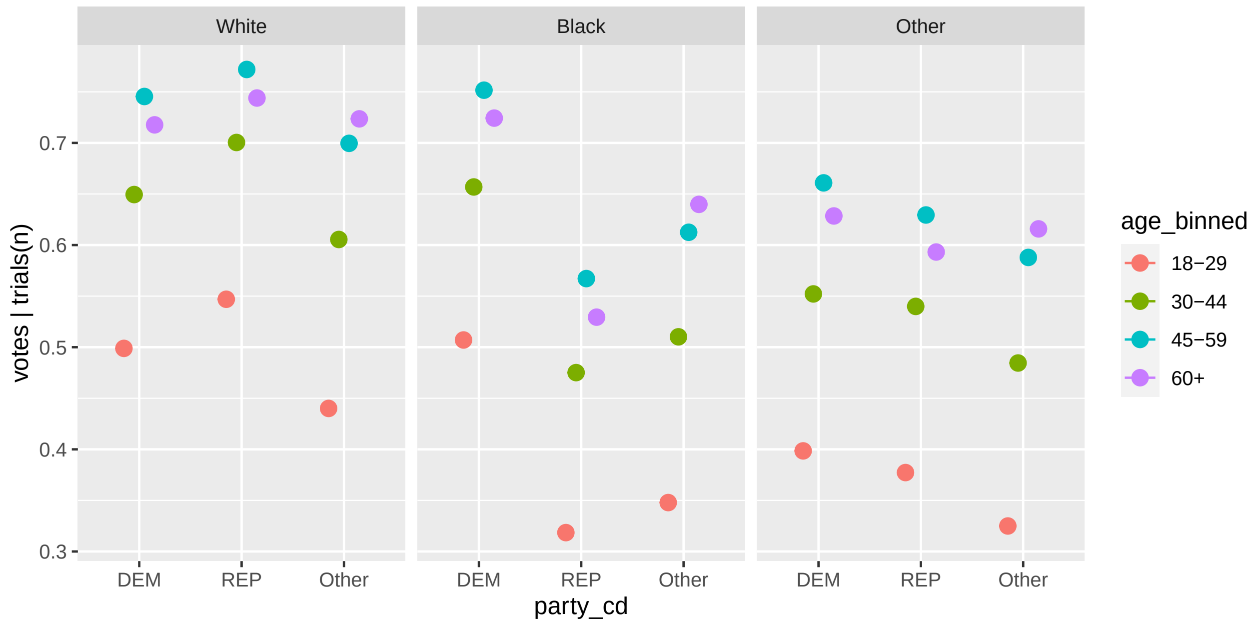
TODO: label graphs



From the plot above, we see that the expected probability of voting is generaly greater that 50% for all median household income levels, but tends to increase as median household income increases, holding all other attributes constant (age, gender, race, party).



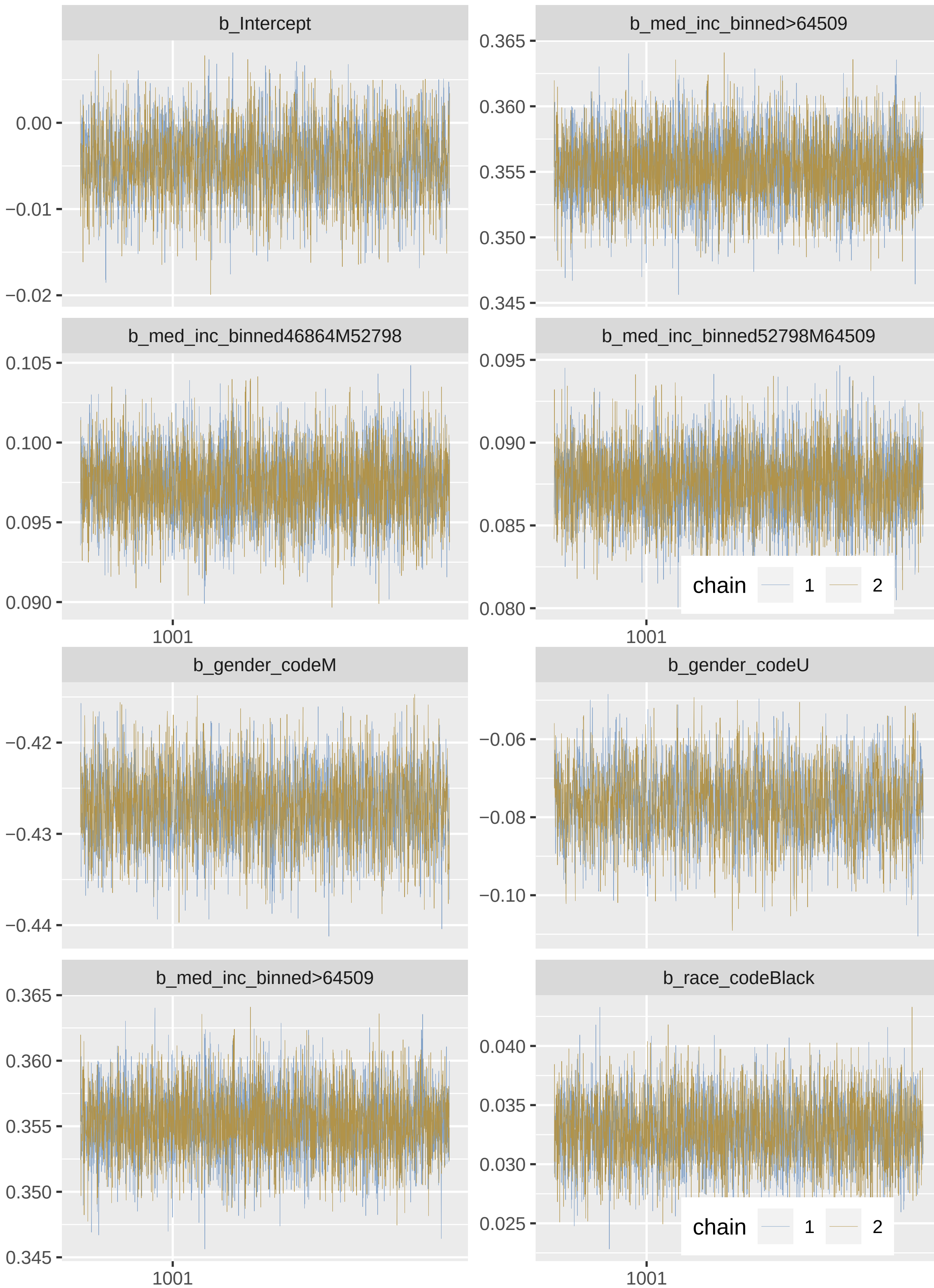
We can interpret each point in the plot above as follows: holding median household income at baseline (less than $46,864) and age at baseline (ages 18-29), the y-axis value is the expected probability that a person of a particular race (x-axis), party (color), and gender (facet) votes. For example, the expected probability that a black, male, Democrat votes is 0.4, whereas the expected probability that a black, male, Republican votes is approximately 0.28. We can also see that women, regardless of race and party, are expected to be more likely to vote than men.

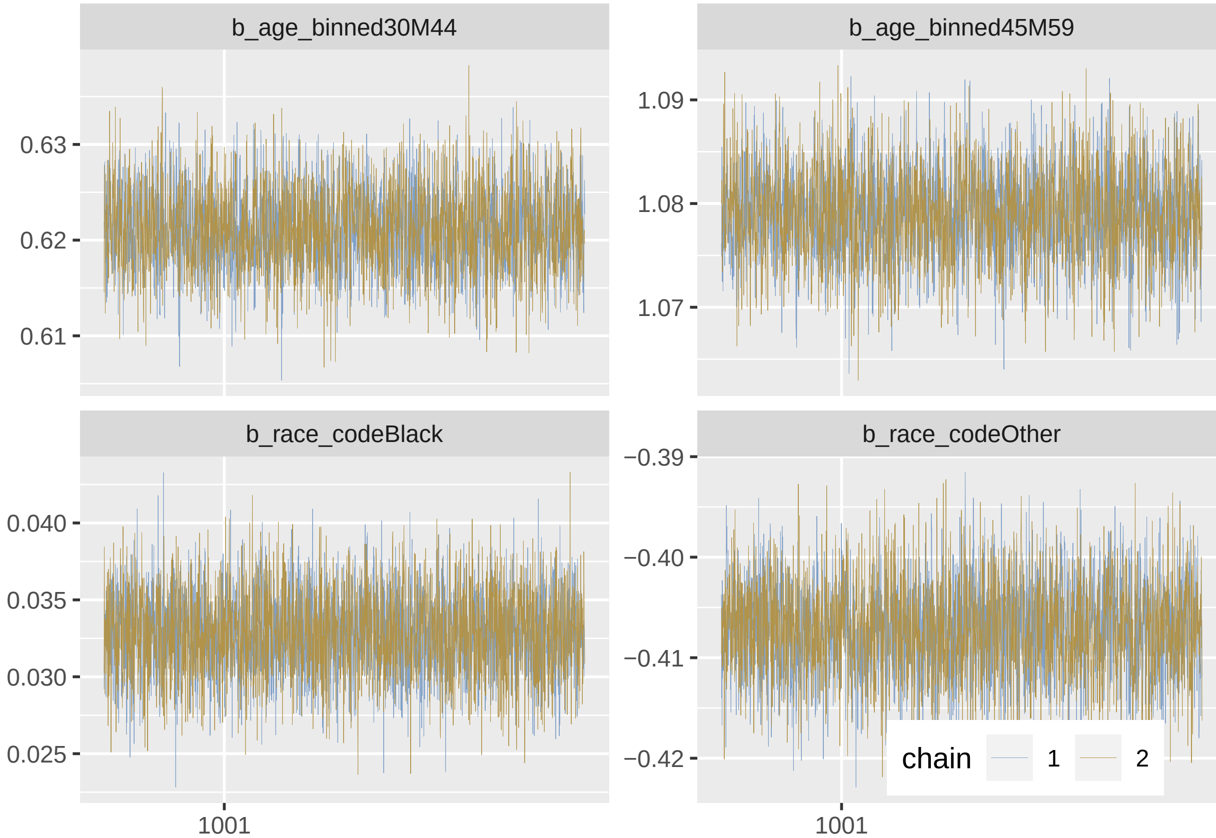


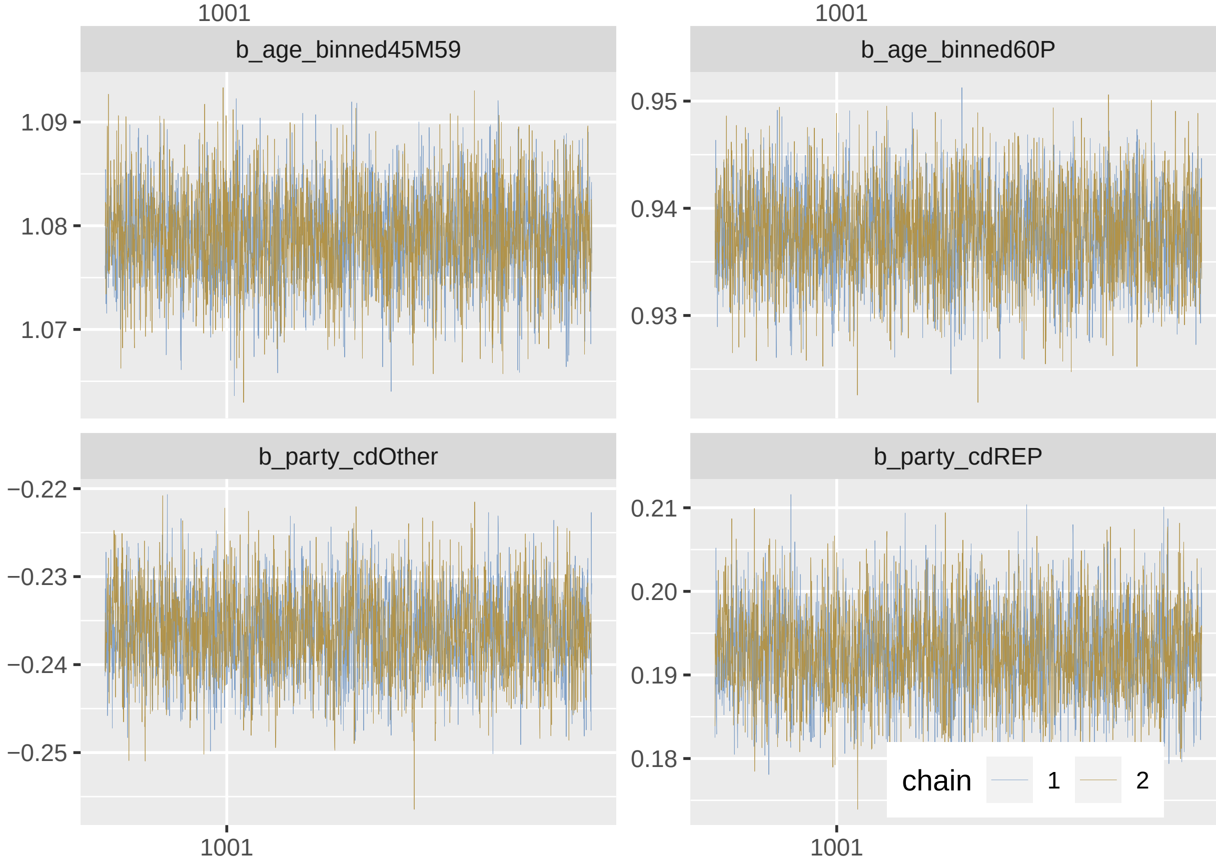
In the plot above, holding median household income at baseline (less than $46,864) and gender at baseline (female), the expected probability of voting for white Democrats across all age groups is less than that for white Republicans. However, the expected probability of voting among black Democrats across all age groups is higher than that among black Republicans.

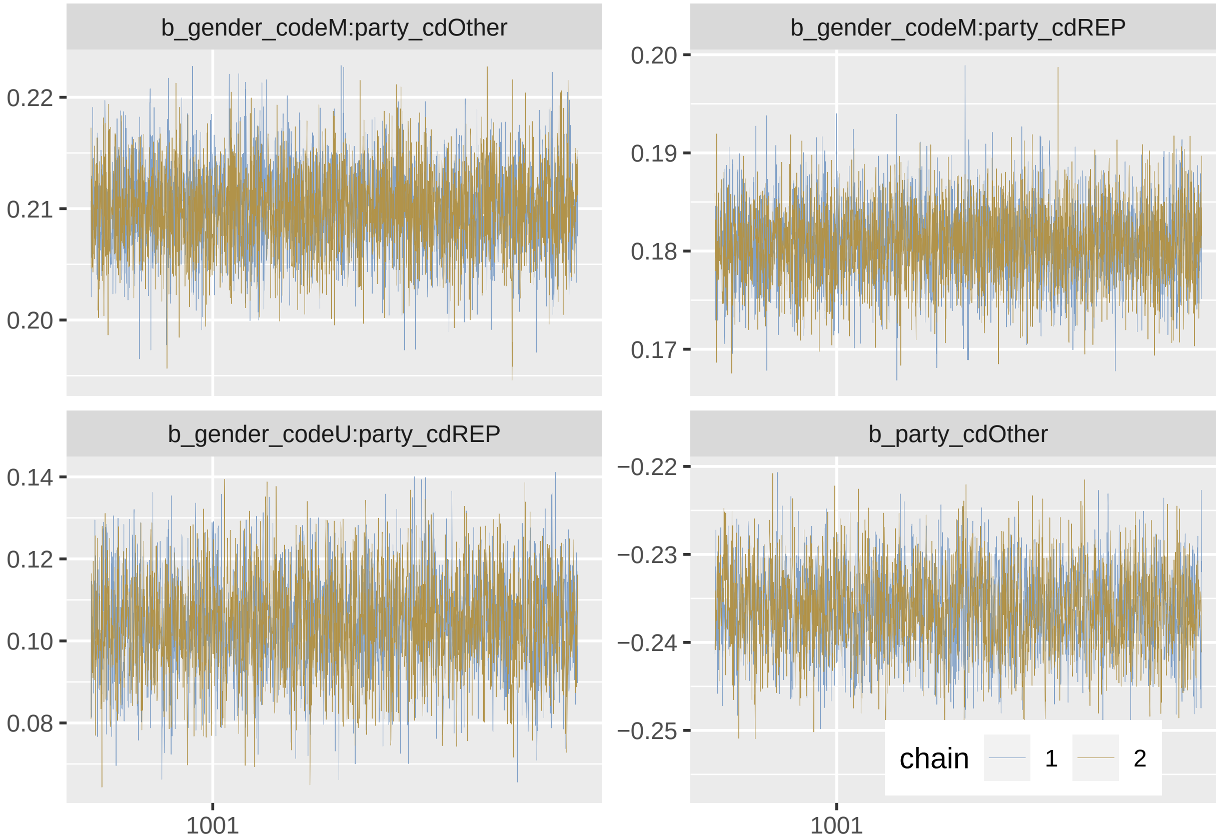
Model Validation

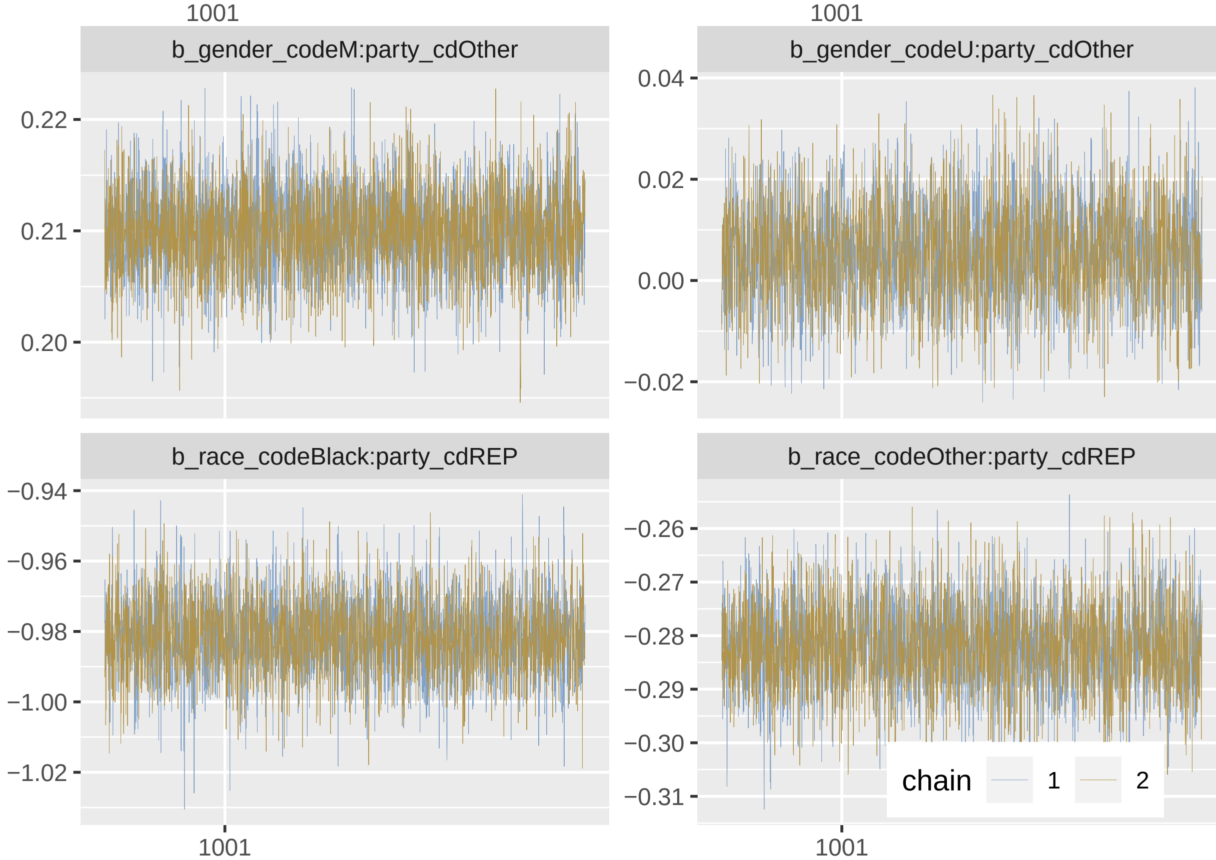
Most of the standardized residuals are within +/- 2, but there are some points that have somewhat larger values. This means that for the majority of the groupings, the model predicts fairly well.

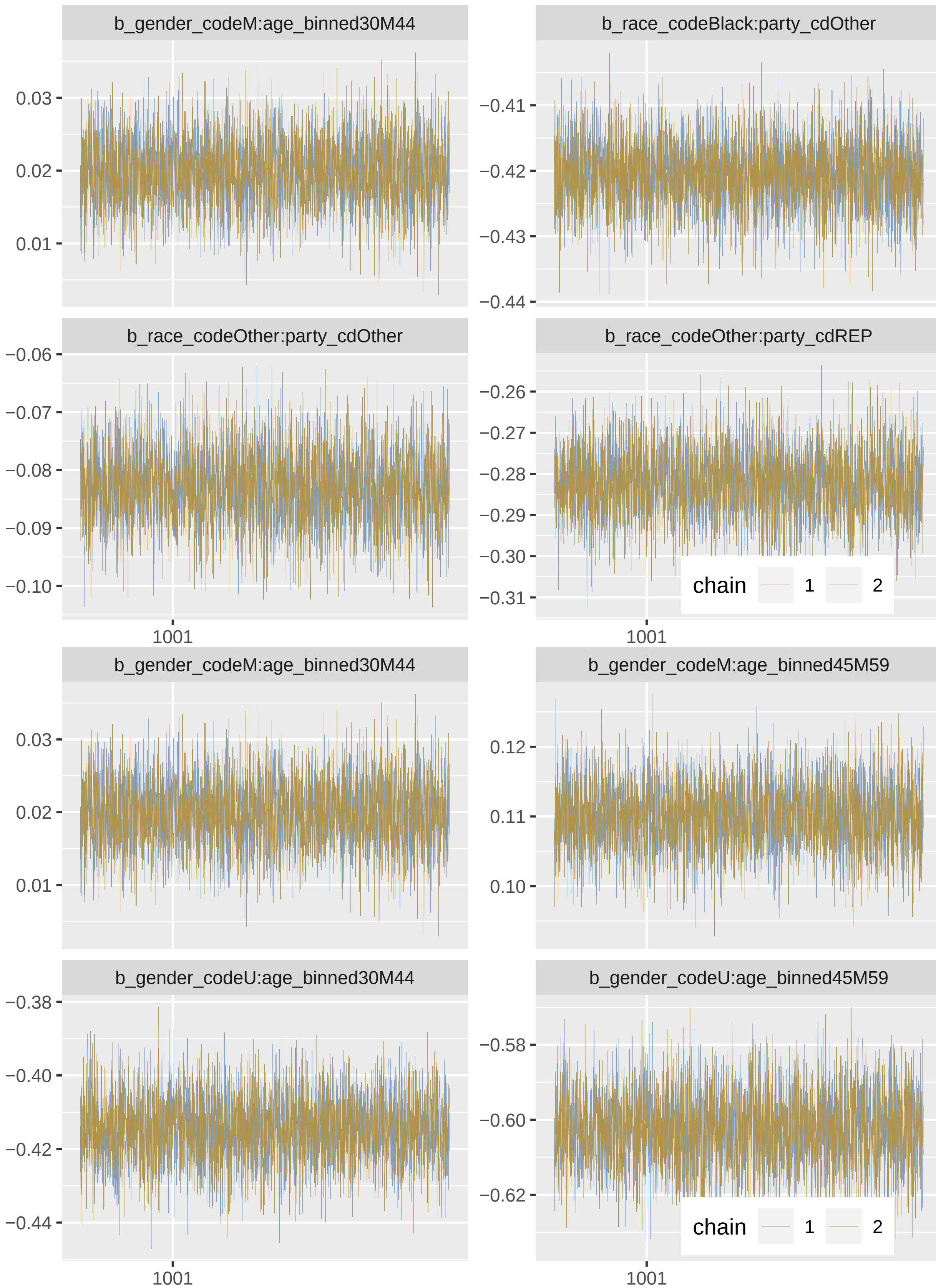

b_Intercept
0.00
−0.01
−0.02
b_med_inc_binned>64509
0.365
0.360
0.355
0.350
0.345
b_med_inc_binned46864M52798
0.105
0.100
0.095
0.090
1001
b_med_inc_binned52798M64509
0.095
0.090
0.085
0.080
chain
1
2
1001
b_gender_codeM
−0.42
−0.43
−0.44
b_gender_codeU
−0.06
−0.08
−0.10
b_med_inc_binned>64509
0.365
0.360
0.355
0.350
0.345
1001
b_race_codeBlack
0.040
0.035
0.030
0.025
chain
1
2
1001

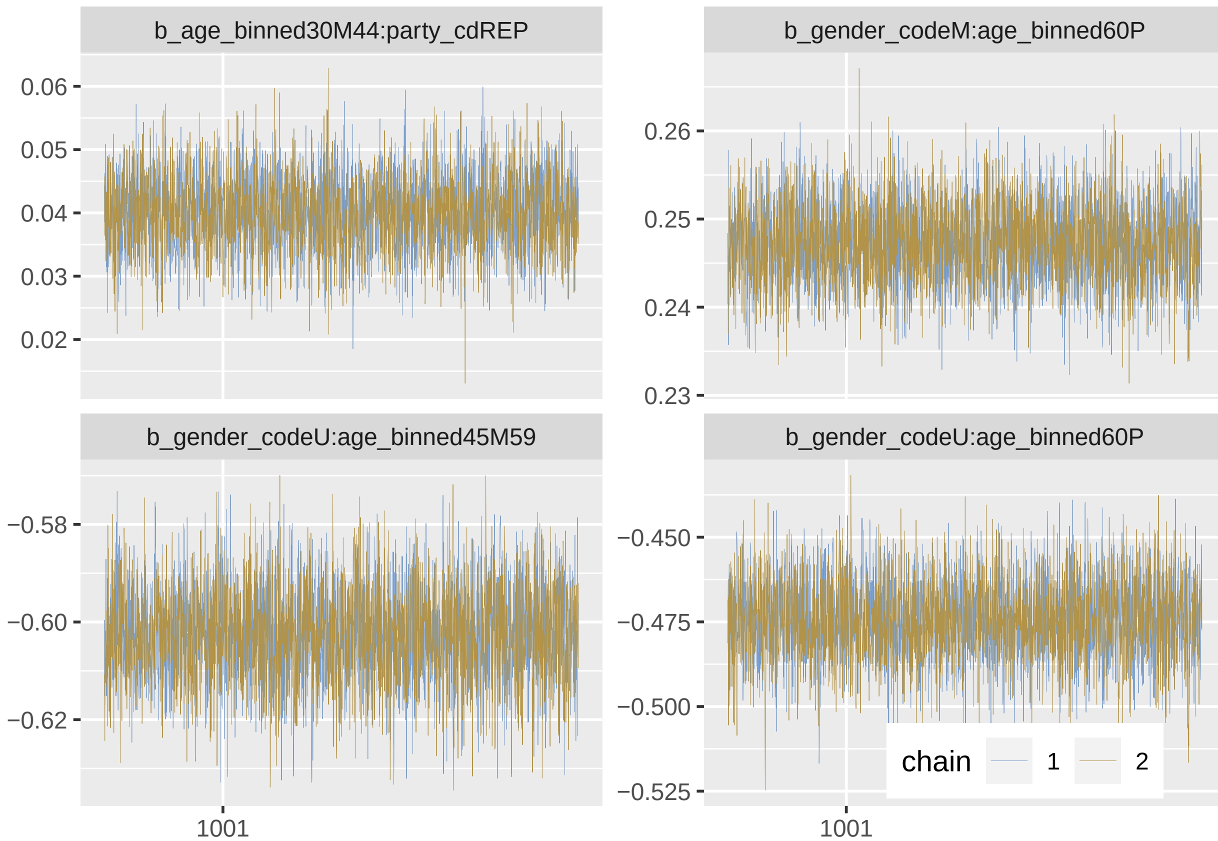

b_age_binned30M44
b_age_binned45M59
b_race_codeBlack
b_race_codeOther
0.63
0.62
0.61
1.09
1.08
1.07
0.040
0.035
0.030
0.025
−0.39
−0.40
−0.41
−0.42
1001
1001
chain
1
2

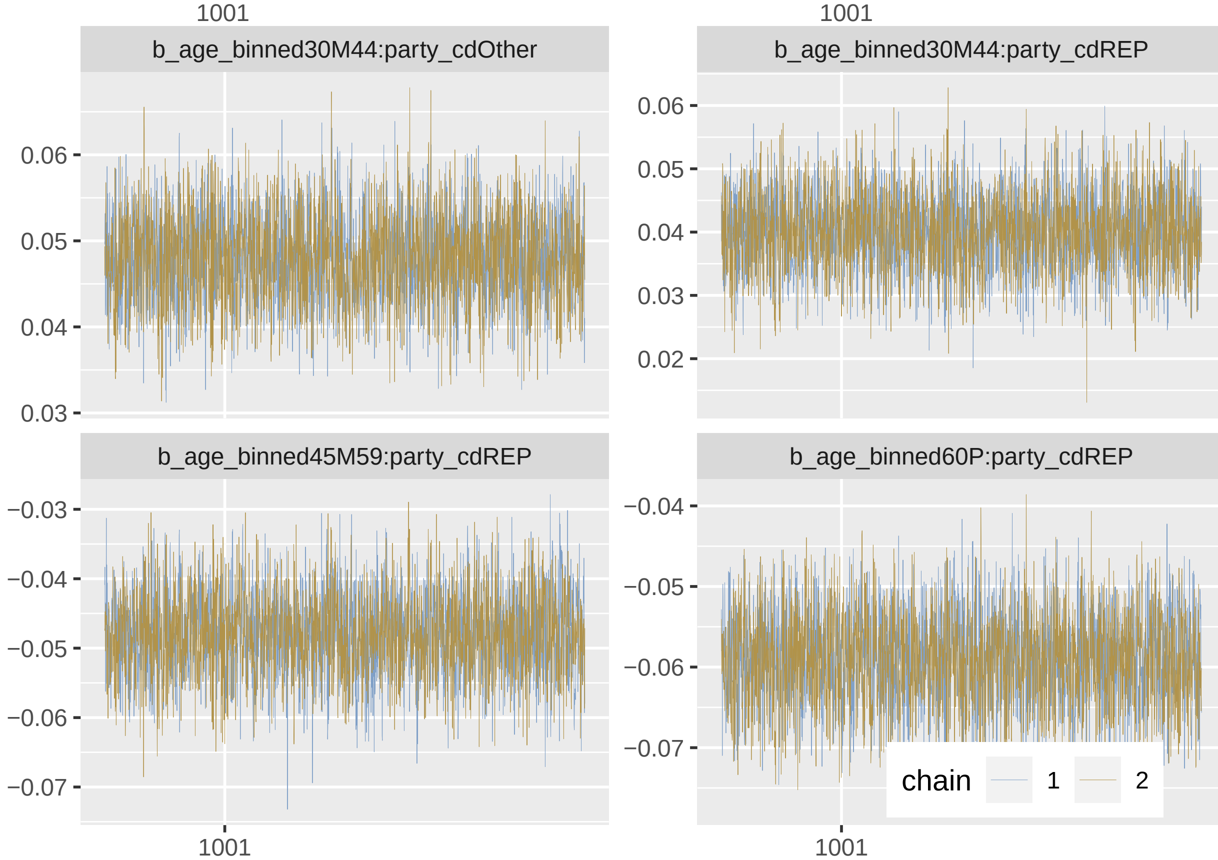

b_age_binned45M59
b_age_binned60P
b_party_cdOther
b_party_cdREP
1.09
1.08
1.07
0.95
0.94
0.93
−0.22
−0.23
−0.24
−0.25
0.21
0.20
0.19
0.18
1001
1001
chain
1
2


b_gender_codeM:party_cdOther
b_gender_codeM:party_cdREP
b_gender_codeU:party_cdREP
b_party_cdOther
0.22
0.21
0.20
0.19
0.18
0.17
0.14
0.12
0.10
0.08
−0.22
−0.23
−0.24
−0.25
1001
chain
1
2


b_gender_codeM:party_cdOther
b_gender_codeU:party_cdOther
b_race_codeBlack:party_cdREP
b_race_codeOther:party_cdREP
0.22
0.21
0.20
0.04
0.02
0.00
−0.02
−0.94
−0.96
−0.98
−1.00
−1.02
−0.26
−0.27
−0.28
−0.29
−0.30
−0.31
1001
chain
1
2


b_gender_codeM:age_binned30M44
0.03
0.02
0.01
b_race_codeBlack:party_cdOther
−0.41
−0.42
−0.43
−0.44
b_race_codeOther:party_cdOther
−0.06
−0.07
−0.08
−0.09
−0.10
1001
b_race_codeOther:party_cdREP
−0.26
−0.27
−0.28
−0.29
−0.30
−0.31
chain
1
2
1001
b_gender_codeM:age_binned30M44
0.03
0.02
0.01
b_gender_codeM:age_binned45M59
0.12
0.11
0.10
b_gender_codeU:age_binned30M44
−0.38
−0.40
−0.42
−0.44
1001
b_gender_codeU:age_binned45M59
−0.58
−0.60
−0.62
chain
1
2
1001


b_age_binned30M44:party_cdREP
b_gender_codeM:age_binned60P
b_gender_codeU:age_binned45M59
b_gender_codeU:age_binned60P
0.06
0.05
0.04
0.03
0.02
0.26
0.25
0.24
0.23
−0.58
−0.60
−0.62
−0.450
−0.475
−0.500
−0.525
1001
1001
chain
1
2


b_age_binned30M44:party_cdOther
b_age_binned30M44:party_cdREP
b_age_binned45M59:party_cdREP
b_age_binned60P:party_cdREP
0.06
0.05
0.04
0.03
0.06
0.05
0.04
0.03
0.02
−0.03
−0.04
−0.05
−0.06
−0.07
−0.04
−0.05
−0.06
−0.07
1001
1001
chain
1
2

Sensitivity Analysis

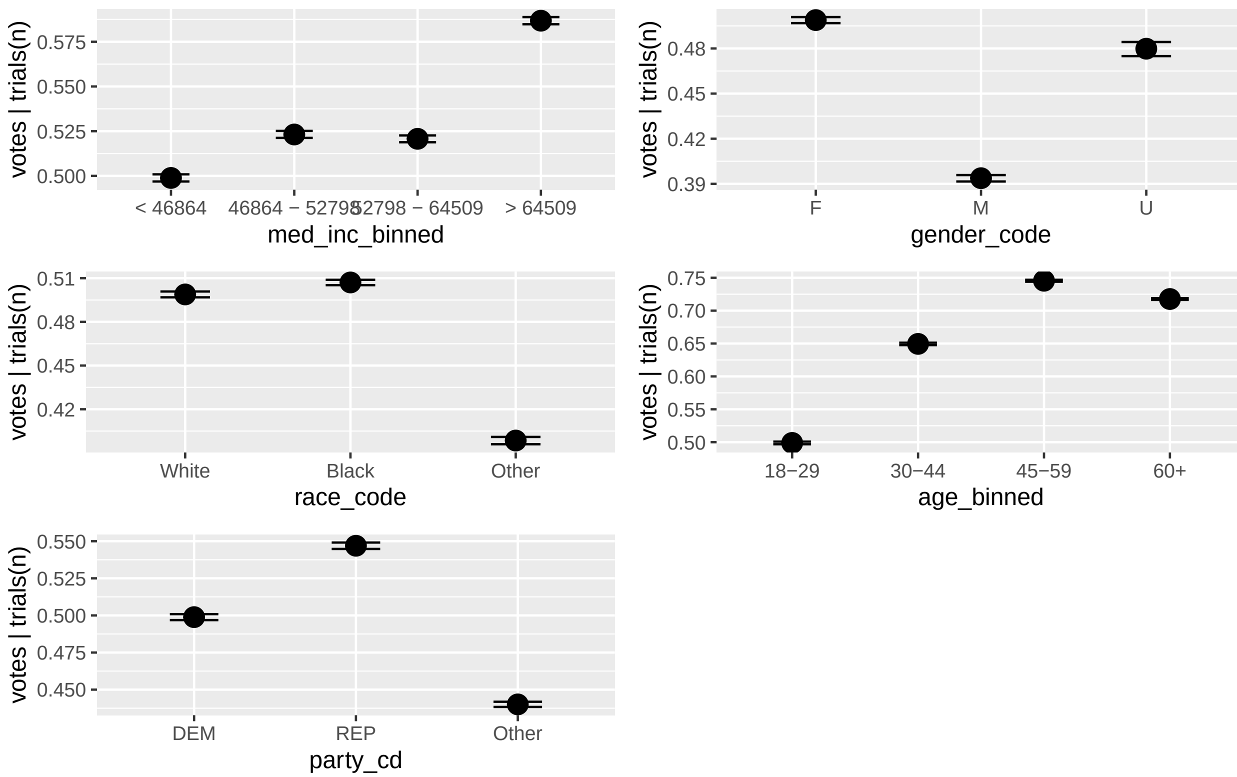
```
## [1] 25039.1
```

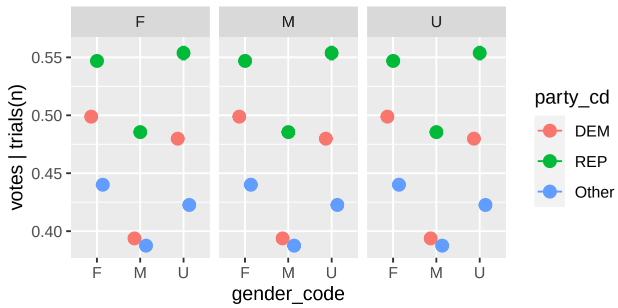
```
## [1] 83471.66
```

TODO: add some sentences

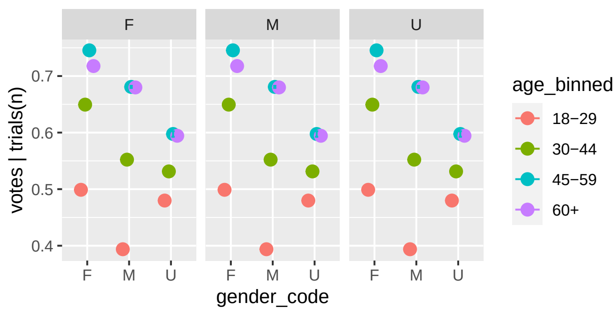
Appendix

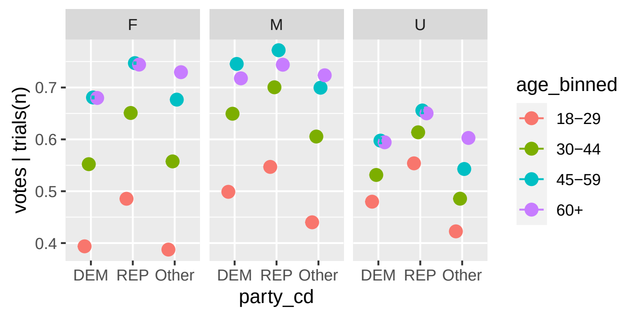
Plots

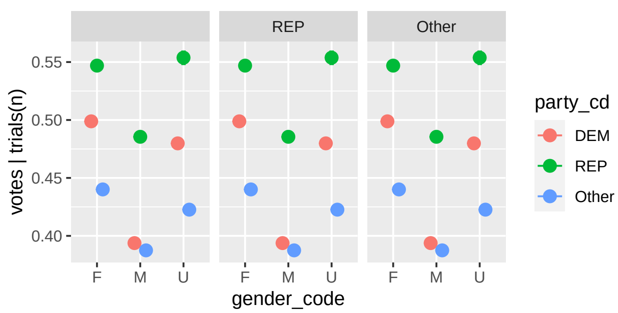


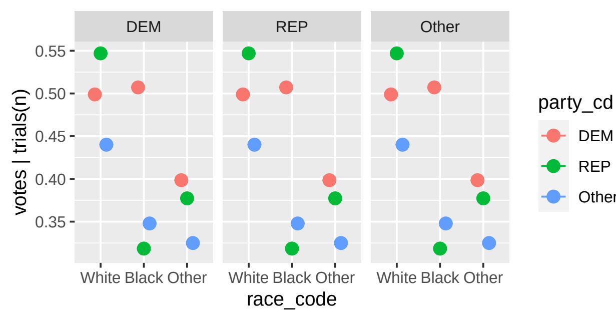

F
M
U
votes | trials(n)
0.55
0.50
0.45
0.40
F M U
gender_code
party_cd
DEM
REP
Other

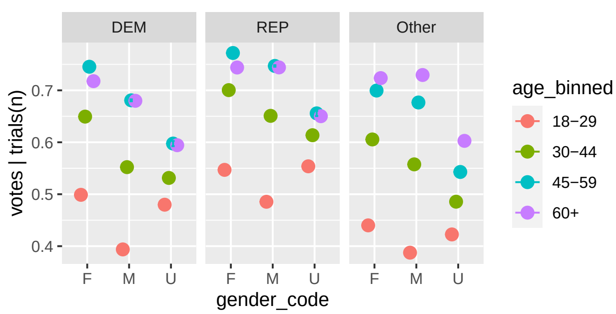

F
M
U
votes | trials(n)
0.5
0.4
0.3
White Black Other
race_code
party_cd
DEM
REP
Other

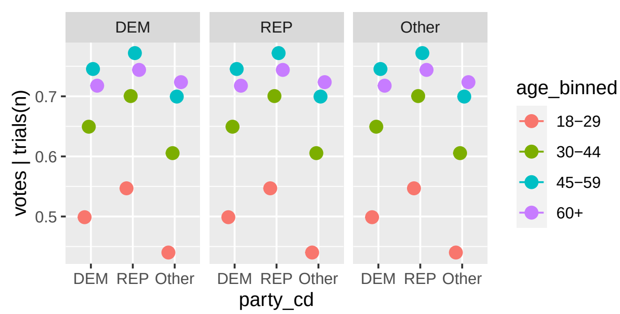

F
M
U
votes | trials(n)
0.7
0.6
0.5
0.4
F M U
gender_code
age_binned
18−29
30−44
45−59
60+

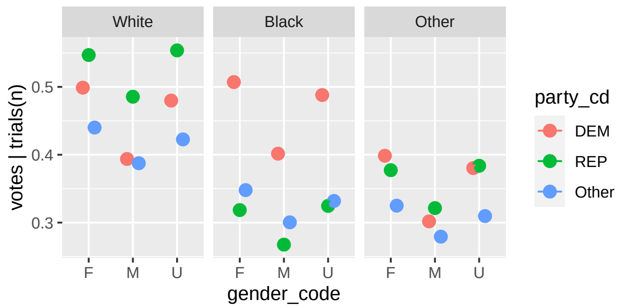

F
M
U
votes | trials(n)
0.7
0.6
0.5
0.4
DEM REP Other
party_cd
age_binned
18−29
30−44
45−59
60+

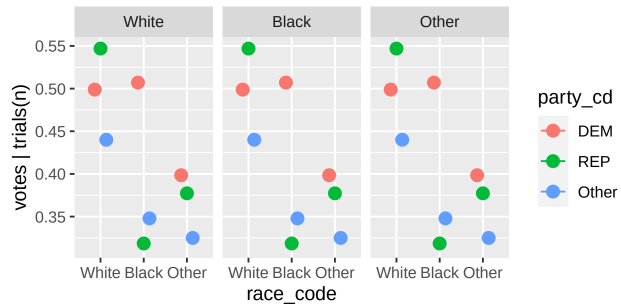

REP
Other
votes | trials(n)
0.55
0.50
0.45
0.40
F M U
gender_code
party_cd
DEM
REP
Other

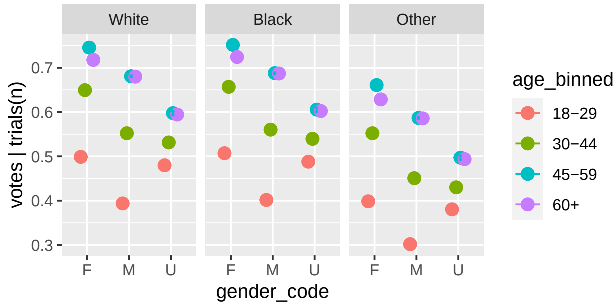

DEM
REP
Other
votes | trials(n)
0.55
0.50
0.45
0.40
0.35
White Black Other
race_code
party_cd
DEM
REP
Other

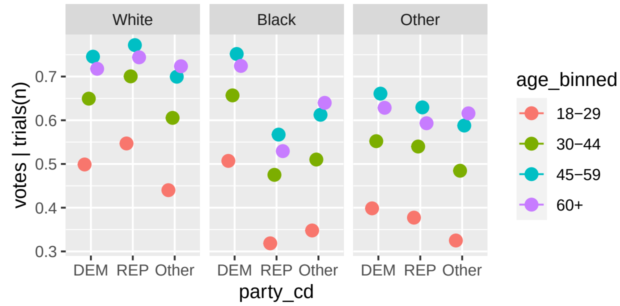

DEM
REP
Other
votes | trials(n)
0.7
0.6
0.5
0.4
F M U
gender_code
age_binned
18−29
30−44
45−59
60+


DEM
REP
Other
votes | trials(n)
0.7
0.6
0.5
DEM REP Other
party_cd
age_binned
18−29
30−44
45−59
60+









Sensitivity Analysis

```
priors <- c(set_prior("normal(0,3)", class = "Intercept"),
            set_prior("normal(0,3)", class = "b"),
            set_prior("normal(0.5,3)", class = "b", coef = "age_binned30M44"),
            set_prior("normal(0.75,3)", class = "b", coef = "age_binned45M59"),
            set_prior("normal(1,3)", class = "b", coef = "age_binned60P"),
            set_prior("normal(-0.5,3)", class = "b", coef = "gender_codeM" ),
            set_prior("normal(-0.5,3)", class = "b", coef = "race_codeBlack" ),
            set_prior("normal(-1,3)", class = "b", coef = "race_codeOther" ),
            set_prior("normal(-0.5,3)", class = "b", coef = "party_cdOther" ),
            set_prior("normal(0.5,3)", class = "b", coef = "med_inc_binned46864M52798" ),
            set_prior("normal(0.5,3)", class = "b", coef = "med_inc_binned52798M64509" ),
            set_prior("normal(1,3)", class = "b", coef = "med_inc_binned>64509" ))


binary_model_newpriors <-
  brm(data = voter_grouped, family = binomial,
      votes | trials(n) ~ 1 + med_inc_binned + gender_code + race_code + age_binned + party_cd
      iter = 2500, warmup = 500, cores = 2, chains = 2,
      seed = 10,
      prior=priors)
```

```
## Compiling Stan program...

## Trying to compile a simple C file

## Running /Library/Frameworks/R.framework/Resources/bin/R CMD SHLIB foo.c
## clang -mmacosx-version-min=10.13 -I"/Library/Frameworks/R.framework/Resources/include" -DNDI
## In file included from <built-in>:1:
```

```
## In file included from /Users/cathylee/Library/R/4.0/library/StanHeaders/include/stan/math/pr
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Dense:1
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Core:88
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/src/Core/util/Macros.h:613:1:
## namespace Eigen {
## ^
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/src/Core/util/Macros.h:613:16
## namespace Eigen {
##                ^
##                ;
## In file included from <built-in>:1:
## In file included from /Users/cathylee/Library/R/4.0/library/StanHeaders/include/stan/math/pr
## In file included from /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Dense:1
## /Users/cathylee/Library/R/4.0/library/RcppEigen/include/Eigen/Core:96:10: fatal error: 'comp
## #include <complex>
##          ^~~~~~~~~
## 3 errors generated.
## make: *** [foo.o] Error 1

## Start sampling
```

```
summary(binary_model_newpriors)
```

```
##  Family: binomial
##   Links: mu = logit
## Formula: votes | trials(n) ~ 1 + med_inc_binned + gender_code + race_code + age_binned + par
##    Data: voter_grouped (Number of observations: 432)
## Samples: 2 chains, each with iter = 2500; warmup = 500; thin = 1;
##          total post-warmup samples = 4000
##
## Population-Level Effects:
##                              Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                       -0.00      0.00    -0.01     0.00 1.00
## med_inc_binned46864M52798        0.10      0.00     0.09     0.10 1.00
## med_inc_binned52798M64509        0.09      0.00     0.08     0.09 1.00
## med_inc_binned>64509             0.36      0.00     0.35     0.36 1.00
## gender_codeM                    -0.43      0.00    -0.44    -0.42 1.00
## gender_codeU                    -0.08      0.01    -0.09    -0.06 1.00
## race_codeBlack                   0.03      0.00     0.03     0.04 1.00
## race_codeOther                  -0.41      0.00    -0.42    -0.40 1.00
## age_binned30M44                  0.62      0.00     0.61     0.63 1.00
## age_binned45M59                  1.08      0.00     1.07     1.09 1.00
## age_binned60P                    0.94      0.00     0.93     0.95 1.00
## party_cdREP                      0.19      0.01     0.18     0.20 1.00
## party_cdOther                   -0.24      0.00    -0.25    -0.23 1.00
## gender_codeM:party_cdREP         0.18      0.00     0.17     0.19 1.00
## gender_codeU:party_cdREP         0.10      0.01     0.08     0.13 1.00
## gender_codeM:party_cdOther       0.21      0.00     0.20     0.22 1.00
## gender_codeU:party_cdOther       0.01      0.01    -0.01     0.03 1.00
```

```
## race_codeBlack:party_cdREP       -0.98      0.01    -1.00    -0.96 1.00
## race_codeOther:party_cdREP       -0.28      0.01    -0.30    -0.27 1.00
## race_codeBlack:party_cdOther     -0.42      0.01    -0.43    -0.41 1.00
## race_codeOther:party_cdOther     -0.08      0.01    -0.10    -0.07 1.00
## gender_codeM:age_binned30M44      0.02      0.00     0.01     0.03 1.00
## gender_codeU:age_binned30M44     -0.41      0.01    -0.43    -0.40 1.00
## gender_codeM:age_binned45M59      0.11      0.00     0.10     0.12 1.00
## gender_codeU:age_binned45M59     -0.60      0.01    -0.62    -0.58 1.00
## gender_codeM:age_binned60P        0.25      0.00     0.24     0.26 1.00
## gender_codeU:age_binned60P       -0.47      0.01    -0.50    -0.45 1.00
## age_binned30M44:party_cdREP       0.04      0.01     0.03     0.05 1.00
## age_binned45M59:party_cdREP      -0.05      0.01    -0.06    -0.04 1.00
## age_binned60P:party_cdREP        -0.06      0.01    -0.07    -0.05 1.00
## age_binned30M44:party_cdOther     0.05      0.01     0.04     0.06 1.00
## age_binned45M59:party_cdOther     0.01      0.01    -0.00     0.02 1.00
## age_binned60P:party_cdOther       0.27      0.01     0.25     0.28 1.00
##                              Bulk_ESS Tail_ESS
## Intercept                        2036     3104
## med_inc_binned46864M52798        5855     3291
## med_inc_binned52798M64509        5677     3436
## med_inc_binned>64509             6194     3170
## gender_codeM                     2907     3038
## gender_codeU                     1838     2443
## race_codeBlack                   4241     3443
## race_codeOther                   2351     2598
## age_binned30M44                  2764     3113
## age_binned45M59                  2703     3098
## age_binned60P                    2375     3003
## party_cdREP                      2243     2787
## party_cdOther                    2114     3196
## gender_codeM:party_cdREP         5297     3278
## gender_codeU:party_cdREP         2348     2683
## gender_codeM:party_cdOther       5240     3574
## gender_codeU:party_cdOther       2107     2633
## race_codeBlack:party_cdREP       4066     2973
## race_codeOther:party_cdREP       2560     3029
## race_codeBlack:party_cdOther     4980     2779
## race_codeOther:party_cdOther     2479     2838
## gender_codeM:age_binned30M44     3465     3404
## gender_codeU:age_binned30M44     3147     3088
## gender_codeM:age_binned45M59     3461     3211
## gender_codeU:age_binned45M59     3459     3050
## gender_codeM:age_binned60P       3491     3360
## gender_codeU:age_binned60P       3785     2765
## age_binned30M44:party_cdREP      2858     2912
## age_binned45M59:party_cdREP      3078     2974
## age_binned60P:party_cdREP        2593     2890
## age_binned30M44:party_cdOther    3087     3210
```

```
## age_binned45M59:party_cdOther     2845     3075
## age_binned60P:party_cdOther       2619     2528
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```
#prior_summary(binary_model_newpriors)
#saveRDS(binary_model_newpriors, "grouped_model_no_randeff_newpriors_whole_dataset.rds")
```

```
summary(randeff_model)
```

```
##  Family: binomial
##   Links: mu = logit
## Formula: votes | trials(n) ~ 1 + med_inc_binned + gender_code + race_code + age_binned + pa
##    Data: voter_grouped_sa (Number of observations: 3752)
## Samples: 2 chains, each with iter = 4500; warmup = 500; thin = 1;
##          total post-warmup samples = 8000
##
## Group-Level Effects:
## ~cong_dist_abbrv (Number of levels: 13)
##               Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sd(Intercept)     0.16      0.04     0.11     0.25 1.00     1268     2043
##
## Population-Level Effects:
##                              Estimate Est.Error l-95% CI u-95% CI Rhat
## Intercept                        0.03      0.05    -0.06     0.12 1.00
## med_inc_binned46864_52798        0.02      0.00     0.02     0.03 1.00
## med_inc_binned52798_64509        0.09      0.00     0.09     0.10 1.00
## med_inc_binned>64509             0.23      0.00     0.22     0.24 1.00
## gender_codeM                    -0.43      0.00    -0.44    -0.42 1.00
## gender_codeU                    -0.08      0.01    -0.10    -0.07 1.00
## race_codeBlack                   0.01      0.00     0.01     0.02 1.00
## race_codeOther                  -0.41      0.00    -0.42    -0.41 1.00
## age_binned30M44                  0.63      0.00     0.62     0.64 1.00
## age_binned45M59                  1.09      0.00     1.08     1.10 1.00
## age_binned60P                    0.96      0.00     0.95     0.97 1.00
## party_cdOther                   -0.22      0.00    -0.23    -0.21 1.00
## party_cdREP                      0.21      0.01     0.20     0.22 1.00
## gender_codeM:party_cdOther       0.21      0.00     0.20     0.22 1.00
## gender_codeU:party_cdOther       0.01      0.01    -0.01     0.03 1.00
## gender_codeM:party_cdREP         0.18      0.00     0.17     0.19 1.00
## gender_codeU:party_cdREP         0.10      0.01     0.08     0.13 1.00
## race_codeBlack:party_cdOther    -0.42      0.01    -0.43    -0.41 1.00
## race_codeOther:party_cdOther    -0.09      0.01    -0.10    -0.08 1.00
## race_codeBlack:party_cdREP      -0.98      0.01    -1.00    -0.96 1.00
## race_codeOther:party_cdREP      -0.27      0.01    -0.29    -0.26 1.00
## gender_codeM:age_binned30M44     0.02      0.00     0.01     0.03 1.00
## gender_codeU:age_binned30M44    -0.41      0.01    -0.42    -0.39 1.00
```

```
## gender_codeM:age_binned45M59      0.11      0.00     0.10     0.12 1.00
## gender_codeU:age_binned45M59     -0.59      0.01    -0.61    -0.57 1.00
## gender_codeM:age_binned60P        0.25      0.00     0.24     0.26 1.00
## gender_codeU:age_binned60P       -0.46      0.01    -0.49    -0.44 1.00
## age_binned30M44:party_cdOther     0.05      0.01     0.03     0.06 1.00
## age_binned45M59:party_cdOther     0.00      0.01    -0.01     0.01 1.00
## age_binned60P:party_cdOther       0.27      0.01     0.26     0.28 1.00
## age_binned30M44:party_cdREP       0.04      0.01     0.02     0.05 1.00
## age_binned45M59:party_cdREP      -0.06      0.01    -0.07    -0.05 1.00
## age_binned60P:party_cdREP        -0.08      0.01    -0.09    -0.06 1.00
##                               Bulk_ESS Tail_ESS
## Intercept                         1025     1824
## med_inc_binned46864_52798        11878     7053
## med_inc_binned52798_64509        11027     6531
## med_inc_binned>64509             11415     5995
## gender_codeM                      6304     5774
## gender_codeU                      4267     5025
## race_codeBlack                    8513     7074
## race_codeOther                    5986     5971
## age_binned30M44                   4903     5442
## age_binned45M59                   4843     5678
## age_binned60P                     4813     5598
## party_cdOther                     4616     5837
## party_cdREP                       4437     5583
## gender_codeM:party_cdOther       10123     5857
## gender_codeU:party_cdOther        4914     5237
## gender_codeM:party_cdREP         10170     5624
## gender_codeU:party_cdREP          5209     6028
## race_codeBlack:party_cdOther      8613     6428
## race_codeOther:party_cdOther      6368     5780
## race_codeBlack:party_cdREP        8085     5355
## race_codeOther:party_cdREP        6167     5850
## gender_codeM:age_binned30M44      6490     6022
## gender_codeU:age_binned30M44      6527     5825
## gender_codeM:age_binned45M59      6816     5924
## gender_codeU:age_binned45M59      7111     6226
## gender_codeM:age_binned60P        6899     6144
## gender_codeU:age_binned60P        7054     5635
## age_binned30M44:party_cdOther     5101     5844
## age_binned45M59:party_cdOther     5379     6185
## age_binned60P:party_cdOther       5422     6462
## age_binned30M44:party_cdREP       5300     6127
## age_binned45M59:party_cdREP       4888     5740
## age_binned60P:party_cdREP         4919     5673
##
## Samples were drawn using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

```r
freq_model = glm(cbind(votes, n-votes) ~ med_inc_binned + gender_code + race_code + age_binned
summary(freq_model)
```

```
##
## Call:
## glm(formula = cbind(votes, n - votes) ~ med_inc_binned + gender_code +
##     race_code + age_binned + party_cd + gender_code:party_cd +
##     race_code:party_cd + gender_code:age_binned + party_cd:age_binned,
##     family = "binomial", data = voter_grouped)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -26.1320   -3.2901   -0.1832    3.1721   21.6650
##
## Coefficients:
##                              Estimate Std. Error  z value Pr(>|z|)
## (Intercept)                 -0.004655   0.004044   -1.151    0.250
## med_inc_binned46864 - 52798  0.097287   0.002260   43.055  < 2e-16 ***
## med_inc_binned52798 - 64509  0.087593   0.002144   40.858  < 2e-16 ***
## med_inc_binned> 64509        0.355307   0.002584  137.505  < 2e-16 ***
## gender_codeM                -0.427075   0.004135 -103.295  < 2e-16 ***
## gender_codeU                -0.076579   0.009172   -8.349  < 2e-16 ***
## race_codeBlack               0.032890   0.002849   11.543  < 2e-16 ***
## race_codeOther              -0.407169   0.004838  -84.162  < 2e-16 ***
## age_binned30-44              0.621153   0.004370  142.143  < 2e-16 ***
## age_binned45-59              1.079045   0.004457  242.121  < 2e-16 ***
## age_binned60+                0.937577   0.004250  220.611  < 2e-16 ***
## party_cdREP                  0.192797   0.005219   36.939  < 2e-16 ***
## party_cdOther               -0.236224   0.004684  -50.432  < 2e-16 ***
## gender_codeM:party_cdREP     0.180748   0.004127   43.798  < 2e-16 ***
## gender_codeU:party_cdREP     0.104340   0.012238    8.526  < 2e-16 ***
## gender_codeM:party_cdOther   0.210030   0.004001   52.489  < 2e-16 ***
## gender_codeU:party_cdOther   0.005784   0.010015    0.577    0.564
## race_codeBlack:party_cdREP  -0.981962   0.011422  -85.972  < 2e-16 ***
## race_codeOther:party_cdREP  -0.282242   0.008327  -33.896  < 2e-16 ***
## race_codeBlack:party_cdOther -0.420549  0.005216  -80.624  < 2e-16 ***
## race_codeOther:party_cdOther -0.083115  0.006562  -12.667  < 2e-16 ***
## gender_codeM:age_binned30-44  0.019943  0.004774    4.178 2.94e-05 ***
## gender_codeU:age_binned30-44 -0.414602  0.009339  -44.396  < 2e-16 ***
## gender_codeM:age_binned45-59  0.110012  0.004851   22.679  < 2e-16 ***
## gender_codeU:age_binned45-59 -0.602838  0.010334  -58.335  < 2e-16 ***
## gender_codeM:age_binned60+    0.247269  0.004762   51.922  < 2e-16 ***
## gender_codeU:age_binned60+   -0.474441  0.012201  -38.887  < 2e-16 ***
## age_binned30-44:party_cdREP   0.040026  0.006088    6.575 4.87e-11 ***
## age_binned45-59:party_cdREP  -0.048105  0.005980   -8.044 8.69e-16 ***
## age_binned60+:party_cdREP    -0.059120  0.005802  -10.189  < 2e-16 ***
## age_binned30-44:party_cdOther 0.047938  0.005296    9.052  < 2e-16 ***
```

```
## age_binned45-59:party_cdOther  0.007212   0.005547    1.300    0.194
## age_binned60+:party_cdOther    0.265494   0.005650   46.988   < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 547088  on 431  degrees of freedom
## Residual deviance:  18738  on 399  degrees of freedom
## AIC: 22590
##
## Number of Fisher Scoring iterations: 3
```

```r
confint(freq_model)
```

```
## Waiting for profiling to be done...
```

```
##                                      2.5 %       97.5 %
## (Intercept)                    -0.012580887  0.003271585
## med_inc_binned46864 - 52798     0.092857831  0.101715345
## med_inc_binned52798 - 64509     0.083390913  0.091794588
## med_inc_binned> 64509           0.350242632  0.360371565
## gender_codeM                   -0.435179396 -0.418972345
## gender_codeU                   -0.094557946 -0.058602900
## race_codeBlack                  0.027305673  0.038474958
## race_codeOther                 -0.416650724 -0.397686413
## age_binned30-44                 0.612588502  0.629718264
## age_binned45-59                 1.070311151  1.087780841
## age_binned60+                   0.929247951  0.945907334
## party_cdREP                     0.182567942  0.203027631
## party_cdOther                  -0.245404309 -0.227043457
## gender_codeM:party_cdREP        0.172659271  0.188836408
## gender_codeU:party_cdREP        0.080355176  0.128326133
## gender_codeM:party_cdOther      0.202187403  0.217872600
## gender_codeU:party_cdOther     -0.013846738  0.025412840
## race_codeBlack:party_cdREP     -1.004353225 -0.959579957
## race_codeOther:party_cdREP     -0.298561289 -0.265920746
## race_codeBlack:party_cdOther   -0.430773162 -0.410325971
## race_codeOther:party_cdOther   -0.095975454 -0.070254725
## gender_codeM:age_binned30-44    0.010586898  0.029299569
## gender_codeU:age_binned30-44   -0.432905359 -0.396298369
## gender_codeM:age_binned45-59    0.100504390  0.119519467
## gender_codeU:age_binned45-59   -0.623090454 -0.582581831
## gender_codeM:age_binned60+      0.237935226  0.256603297
## gender_codeU:age_binned60+     -0.498349024 -0.450523254
## age_binned30-44:party_cdREP     0.028094598  0.051958072
## age_binned45-59:party_cdREP    -0.059825765 -0.036383822
## age_binned60+:party_cdREP      -0.070492595 -0.047747709
## age_binned30-44:party_cdOther   0.037558466  0.058317554
```

```
## age_binned45-59:party_cdOther -0.003659305  0.018082926
## age_binned60+:party_cdOther    0.254420621  0.276569407
```

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