Case Study 3 Interim Report: Who Votes in North Carolina?

Ethan Shen, Malavi Ravindran, Steven Herrera Tenorio, Anna Darwish

1 Introduction

In 2019, the Supreme Court asserted that federal courts were not allowed to rule cases on partisan gerrymandering, which allows parties that control the state legislature to draw district maps and influence election outcomes (Rucho v. Common Cause 2019) [1]. North Carolina was at the spotlight of this contested 5-4 ruling, cultivating national dialogue about the consequences of district map drawing, such as voter suppression. Our study aims to uncover demographic characteristics, such race, gender, age group, and county of residence, driving voting patterns in North Carolina. We ultimately seek to answer the following question– *who* votes in North Carolina? Answering this will not only allow for better predictions for the outcomes of Congressional elections, but will also illuminate which demographic groups in the state are hindered from civic engagement, whether by implicit or explicit means. This information can be instrumental to campaigns and activist groups seeking to bolster support and foster more politically minded communities.

Our study will use publicly available voter, registration, and census data in order to determine the number of likely and unlikely voters across various demographic groups in the 100 counties of North Carolina. After performing data munging and harmonization, we implement a multilevel logistic regression model with random county intercepts to make statements regarding an individual's likelihood of voting. In order to assess the performance of the model, we perform both 5-fold cross validation and out of sample validation using data from the 2012 elections. We conclude with a discussion of which demographic characteristics have a significant impact on the likelihood of voting, quantifications of these relationships, and limitations to our data and analysis.

2 Data Description

Data for this study utilizes NC voter history, NC voter registration information, and the census. Voting and registration data were used in conjunction with the census in order to not only understand who votes, but also who *does not*. Below we describe these various data sources as well as the munging and harmonization process used to obtain a singular, comprehensive dataset.

2.1 Voter History and Registration Data

The North Carolina voter history dataset shows how each registered voter in North Carolina, uniquely identified by a voter registration number, voted in the 2016 presidential and 2018 congressional elections. As not all voters participated in both elections, we designated only those who voted in 2016 as "likely" voters for the 2020 election, and disregarded those who voted in 2018. This decision was made because both 2016 and 2020 are presidential election years, while there is typically lower turnout for midterm elections [2].

The North Carolina voter registration dataset contains all legally available voter specific information for eligible voters in the state. Individuals for whom the most recent last voted date was greater than ten years were not included in the dataset. Information provided for each voter includes their county of residence, unique voter registration number, race, ethnicity, gender, and age. In this dataset, race could take on one of the following categories:

- Asian
- Black or African American
- American Indian or Alaska Native
- Two or more races
- Other
- Native Hawaiian or Pacific Islander
- Undesignated
- White

Ethnicity was one of the following:

- Hispanic or Latino
- Not Hispanic or Not Latino
- Undesignated

And gender belonged to one of:

- Male
- Female
- Undetermined

Recall that, in our study, a "likely" voter was one who voted in the 2016 election. Thus, not all individuals in the North Carolina registration dataset were necessarily "likely" voters. This is because certain individuals who were registered to vote and *did* vote within the last 10 years still may not have voted in the 2016 election. Thus, upon joining the voter and registration data by unique voter registration number, only those individuals who were both registered and voted in the 2016 election were preserved. In this sense, only "likely" voters remained in the merged data frame.

Next, certain modifications were made to facilitate analysis. Firstly, age was converted to a categorical variable, *age bin*, which belonged to one of five categories.

- 18-29
- 30-39
- 40-49
- 50-64
- 65+

In addition, the individual race and ethnic identities were combined into a *race and ethnicity* variable which was one of:

- Hispanic any race
- Non Hispanic White
- Non Hispanic Black
- Non Hispanic Other
- Non Hispanic Asian
- Non Hispanic Mixed
- Non Hispanic American Indian or Alaska Native
- Non Hispanic Native Hawaiian or Pacific Islander
- Undetermined

Note that all individuals of hispanic ethnicity were grouped into one category, regardless of racial identity. Once these modifications were made, the data was grouped by *county*, *race and ethnicity*, *age bin*, and *gender*. The size of each group, which is synonymous with the number of likely voters in that group, was included in the data frame as well. For example, one row of the grouped data frame would report the number of likely voters among 18-29 year old, hispanic women in Alamance county. This dataset will be referred to as “Voting and Registration” for the continuation of the paper.

2.2 Census Data

The data harmonization process described above finds the number of likely voters, by demographic, in the 2020 election. However, this process does not account for the population of North Carolinans who are not likely to vote. Thus, in order to properly address the question of who does and does not vote in North Carolina, census data was utilized. This data helped to provide a better sense of the denominator in our calculations– that is, what is the *total* size of each demographic group in each of the 100 counties in North Carolina?

The census data includes, for each county, census year, and age group, the population size by various demographic identities. Examples of these groups include females, hispanic males, and non hispanic females who identify as black alone or in combination. It is clear that the groups represented in the census data range from very broad (gender) to very granular (a combination of gender, ethnicity, and race). As these granular identities account for both race and ethnicity, they were consistent with the *race and ethnicity* variable defined in the “Voting and Registration” dataset. In contrast, age bins in the census were of size 5, which was smaller than those defined in “Voting and Registration.”

To join “Voting and Registration” with the census data, the census data was filtered for the most recent year (2020) and for age groups that were at or greater than the 15-19 bin. For those rows where the age group of interest fell within the 15-19 bin, each population size was divided by 2.5 in order to theoretically account for only those individuals who were 18 or above (and thus eligible to vote). Then, population information was summed over the smaller age groupings to create age bins consistent with those of “Voting and Registration.” For example, population counts from bins 4, 5, and 6 (which mapped to 18-19, 20-24, and 25-29, respectively) were summed together in order to represent counts for the 18-29 bin. Finally, using the census counts provided for each demographic group (by gender, age group, and race and ethnicity) within each county, a new dataframe was constructed with each row representing a county, racial and ethnic group, age bin, and gender and total group size. As can be seen, the format of this dataset was consistent with “Voting and Registration,” but with a column representing the *total* group size rather than the size of likely voters. This dataset will be referred to as “Census”.

The “Voting and Registration” and “Census” datasets were then merged by county, age bin, race and ethnicity, and gender. This final data frame reports, for each demographic group within each county, the number of likely voters (our numerator of interest) as well as the total size (our denominator of interest). This merging process was not perfect, as there were some inconsistencies between information provided by the “Census” and “Voting and Registration” datasets. For example, the census provided counts of only male and female populations. For this reason, those registered voters who specified a gender other than male or female were omitted in this dataset. Those with undesignated racial and ethnic identities were excluded for similar reasons. Similarly, we found that there were no instances of “Native Hawaiian or Pacific Islander” individuals who voted in 2016, necessating their removal from the dataset. The limitations of these decisions will be discussed more thoroughly in section ______.

3 Exploratory Data Analysis

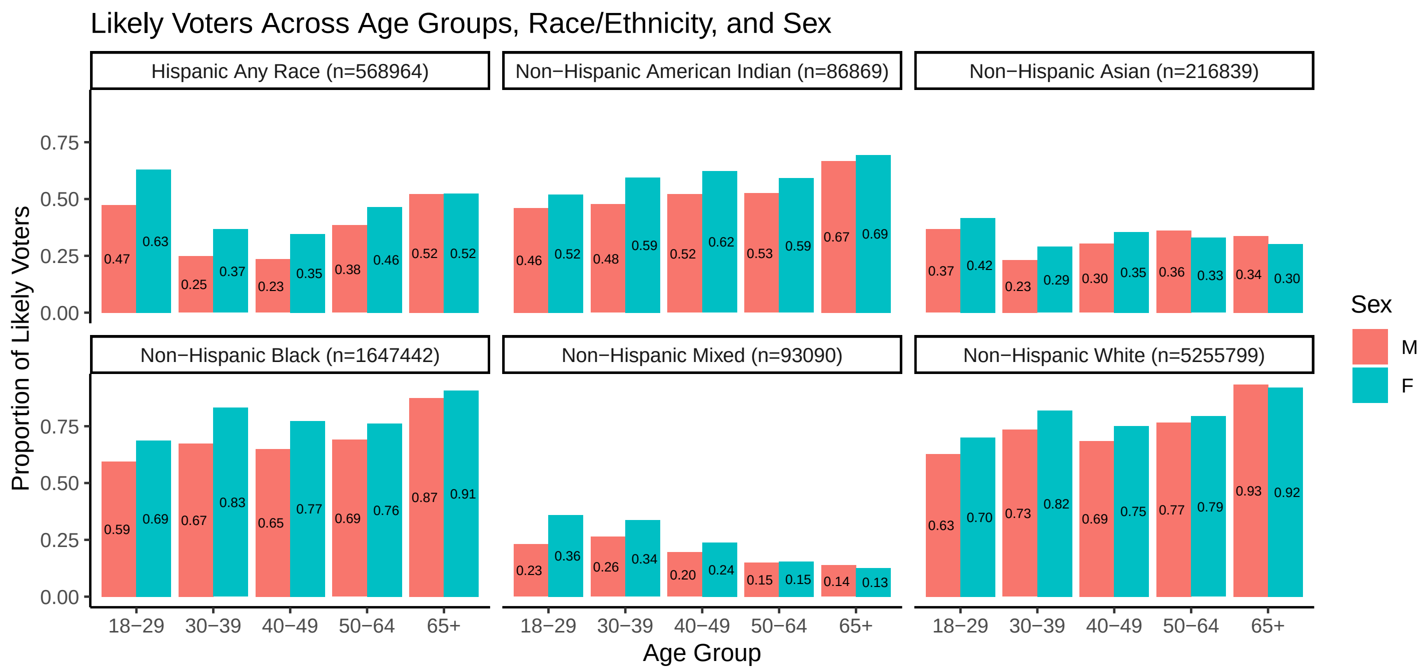


Figure 1

In Figure 1, we can see the proportion of likely voters across race-ethnicity, sex, and age group identities. There appears to be profound differences in likely voter proportion among the different race-ethnicity groups. *Hispanic Any Race*, *Non-Hispanic American Indian*, *Non-Hispanic Black*, and *Non-Hispanic White* groups generally have a much higher proportion of likely voters relative to other groups in North Carolina.

In general, we see an increasing trend in the proportion of likely voters with age for *Non-Hispanic American Indian*, *Non-Hispanic Black*, and *Non-Hispanic White* groups. This trend is actually reversed for *Non-Hispanic Mixed* groups. For *Hispanic Any Race* and *Non-Hispanic Asian* groups, we see a varying relationship that is slightly less profound. These trends point to a potential interaction effect between age and *race and ethnicity*.

In general, there appears to be a higher proportion of likely voters among women compared to men. However, we see this trend level off in the two oldest age groups.

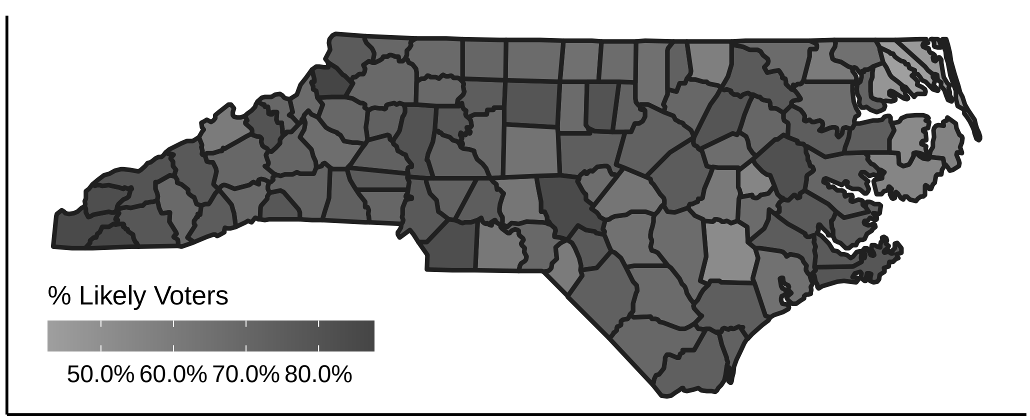


Figure 2a

In Figure 2, we see how the proportion of likely voters varies throughout North Carolina. Heterogeneity among the various counties motivates a consideration of random county intercepts in modelling the likelihood of voting. Some interesting trends can be seen in the state map. Cherokee county, which is the westernmost county in the state, has an extremely high proportion of likely voters. This county is 95% White and heavily Republican, with 75% of the vote going to Donald Trump in 2016. Watauga County, located slightly northeast of Cherokee, similarly shows a high proportion of likely voters. However, despite being similar in racial composition to Cherokee (96% White), growth of a younger voting population, brought forth by Appalachian State University, has turned Watauga into a hotly contested swing county.

4 Methods

To begin with our model development, we conducted exploratory data analysis to get a sense of which covariates were influential in determining likelihood to vote, along with any possible interactions. This led to our starting model:

$$log(\frac{P_{ij}(Likely)}{1 - P_{ij}(Likely)}) = \alpha_j + \alpha_1 * I(Gender_{ij} = Male) + \sum_{a=2}^{5} \alpha_{2a} * I(Age_{ij} = a) +$$
$$\sum_{r=2}^{6} \alpha_{3r} * I(Race/Ethnicity_{ij} = r) + \sum_{c=11}^{30} \alpha_{4c} * I(Age_{ij},\ Race/Ethnicity_{ij} = c) \quad (1)$$
$$\alpha_j \sim N(\alpha_0, \tau^2)$$

P_{ij} is the probability that individual i from county j will vote in the upcoming election, and $alpha_j$ represents the random intercept term for each county. For the *Age* term, we encode $a = 1$ as our baseline, 65+ year olds and then code the remaining in increasing age-bin order. For our *Race/Ethnicitiy* term, we encode $r = 1$, *Hispanic Any Race*, as our baseline group. The remaining *Race/Ethnicity* identities are ordered as follows: *{Non-Hispanic American Indian, Non-Hispanic Asian, Non-Hispanic Black, Non-Hispanic Mixed, Non-Hispanic White}*. Finally, as there are five age bins and six race/ethnicity groups, meaning thirty possible combinations, 10 of these involve the baseline terms, leaving 20 other combinations to appear as evaluated terms in our model.

We compared this model to one without an interaction between *Age* and *Race/Ethnicity*, ones with a random intercept for county and random effect slopes for the three demographic covariates, and other interaction effects. In order to compare the fits of these models, we examined accuracy, F_1-scores, confusion matrices, and ROC curves, along with binned residual plots to confirm that the underlying assumptions of each of these models were met. To compare the performance of these models, we performed 5-fold cross validation and conducted out-of-sample prediction with the 2012 presidential election.

We performed sensitivity analysis to validate the resulting estimates for our coefficients with a Bayesian approach, as there can be issues with convergence when developing random effect models with frequentist methods. Given the size of the dataset, we relied on relatively flat priors, and noticed that virtually all of the coefficients converged to be within two standard errors of our estimates from the frequentist approach. In addition to sensitivity analysis for our variable estimates, we also re-binned the age groups by collapsing them into the following groups: $18 - 39$, $40 - 64$, and 65+, recreated our final model, and observed similar estimates for the other covariates. The results of this model are in Table 2 in the Appendix.

5 Results and Discussion

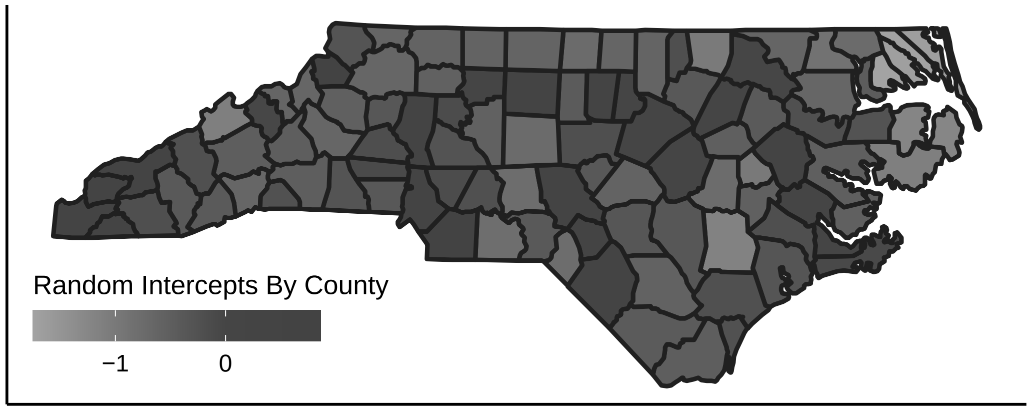
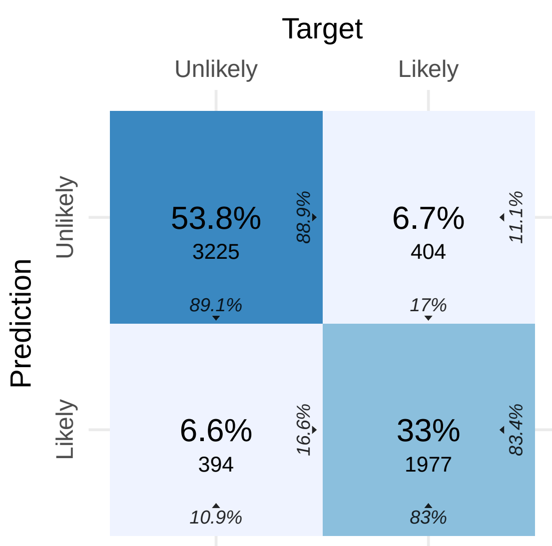


Figure 3

The results of our final model can be viewed in Table 1 in our Appendix. Figure 3 is a heatmap of the intercepts by county. While the intercepts appear to generally capture relative trends, these intercepts appear to be more homogeneous than the true percent of likely voters. This is likely a consequence of our assumption that they stem from the same distribution. The average accuracy and average F1-score on our test sets were 86.72% ± 0.71% and 89.02 ± 0.63 respectively. Our ROC curve also provides evidence this model fits well with an average AUC of 86.09 ± 0.78.

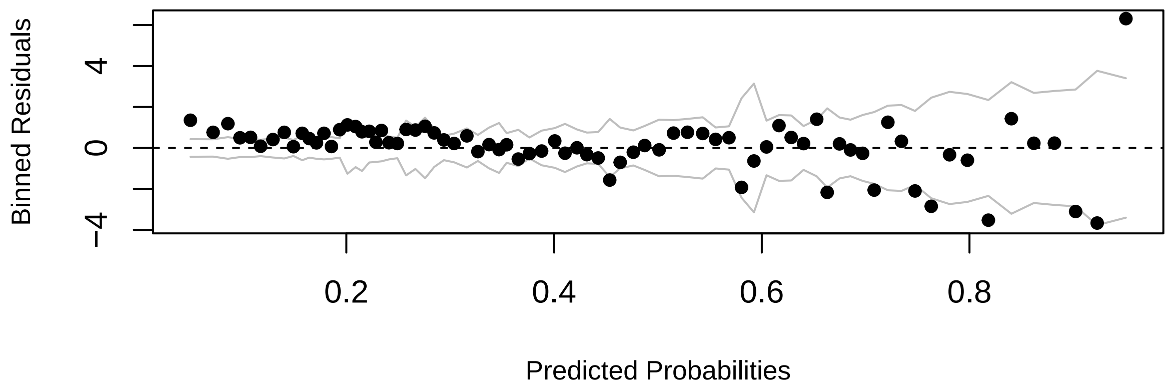
The results of our final model can be viewed in Table 1 in our Appendix, and Figure 3 is a heatmap of the random intercepts by county. While the intercepts appear to generally capture relative trends, these intercepts appear to be more homogeneous than the true percent of likely voters. This is likely a consequence of our assumption that they stem from the same distribution. The average accuracy and average F1-score on our test sets were 86.72% ± 0.71% and 89.02 ± 0.63 respectively. Our ROC curve also provides evidence this model fits well with an average AUC of 86.09 ± 0.78. These evaluation metrics indicate that our model does a reasonable job of predicting who votes in North Carolina.



From the confusion matrix, we can see a relatively even split of false negatives and false positives, suggesting our model isn't predisposed to making either mistake more than the other.

Our binned residuals, however, reveal some issues for our model. We appear to be consistently underestimating the likelihood of voting for groups that are less likely to vote, along with a fanning of a residuals as the likelihood increases. This may cause issues as groups that are less likely to vote have been more active in the past few elections [3] and is important to keep in mind. Another possible cause of this issue was the non-linear relationships between age bins and some Race/Ethnicity groups in our exploratory data analysis. This curvature is more clearly visualized in Figure 4, which is a binned residual plot resulting from the same model but without the interaction effect between Race/Ethnicity and age. Another issue to consider are the slight differences in the heatmaps, which suggested some differences in counties with very high and very low voter turnout. While we considered several transformations to better capture the nuances of this relationship, we experienced convergence issues and inflated standard errors.

Figure 4: Random County Intercept and Age Slope Model

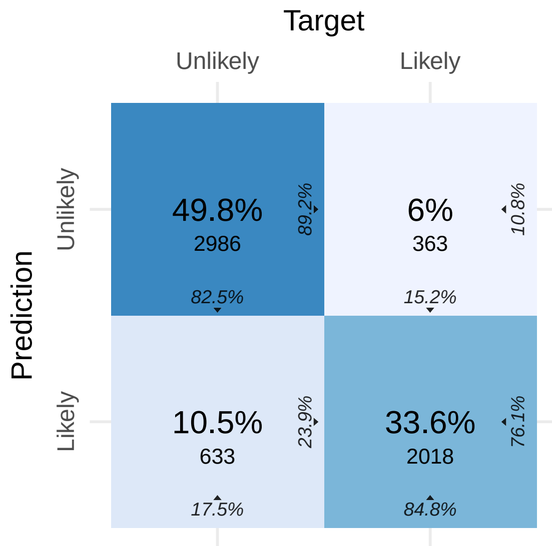


With respect to fixed effects, we see that, with all else held constant, the odds of a female voting in 2020 are 1.4 that of a male. Understanding the effects of race/ethnicity and age are not as straightfoward, as the interaction between the two terms indicates that one cannot be interpreted in isolation of the other. We will focus on an interesting group of voters, 18-29 year olds.

Table 1: Multiplicative odds of being a likely voter compared to Hispanic baseline (among 18-29 year olds)

Race/Ethnicity	Multiplicative Odds	95% Confidence Interval
White (Non Hispanic)	1.71	(1.63, 1.8)
Mixed (Non Hispanic)	0.32	(0.28, 0.37)
Black (Non Hispanic)	1.43	(1.35, 1.51)
Asian (Non Hispanic)	0.43	(0.4, 0.47)
American Indian (Non Hispanic)	0.64	(0.58, 0.7)

The table above shows, for 18-29 years old with county and gender held constant, the multiplicative odds of being a likely voter compared to the baseline of Hispanic (all races). As can be seen, for White and Black individuals, the odds of being a likely voter multiply by of 1.71 and 1.43 respectively. Thus, these individuals are more likely to vote than their hispanic counterparts. Conversely, for 18-29 year old Asians, American Indians, and Mixed individuals, the odds of being a likely voter multiple by factors of 0.43, 0.64 and 0.32 respective. Thus, these individuals are less likely to vote than their hispanic counterparts. It is important to note that these trends are specific to 18-29 year olds in North Carolina.



Now, we will consider our model's ability to predict on voting likelihood for the 2012 presidential election. For our out of sample predictions, we achieved an 83.4% accuracy and our confusion matrix is above. These results are quite similar to our previous confusion matrix, which provides evidence that our model is capable of predicting voter trends outside of the dataset's associated election years.

6 Conclusion and Limitations

Our goal in this case study were to gain a stronger understanding of the demographic composition of Noth Carolina voters relative to its overall population. By combining previous voter registration and history information with the 2020 census data, we were able to create a random effects model that accounted for the relationship between county, age, sex, and race-ethnicity and voter likelihood. By verifying our resulting model with cross validation and out of sample procedures, we were able to interpret and understand the strengths and limitations of our final model.

With respect to limitations, we will begin by noting that the census solely designates individuals as "Male" and "Female," severly limiting the scope of gender identity. For this reason, we had to exclude those individuals who did not report themselves as "Male" or "Female" from our analysis. Secondly, the racial and ethnic groups from these two sources also did not align perfectly. For example, individuals from the registration dataset could only report belonging to one racial and ethnic group. Conversely, individuals on the census could check multiple racial categories. Thus, mapping these identities between the two data sources was not perfect. Similarly, the voter registration data did not include any instances of Pacific Islanders and Native Hawaiians, so these individuals were removed from our analysis.

7 Appendix

7.1 Bibliography

1). Rucho v. Common Cause, No. 18-422, 588 U.S. ____ (2019)

2). DeSilver, Drew. “Voter Turnout Always Drops off for Midterm Elections, but Why?” Pew Research Center, Pew Research Center, 30 May 2020, www.pewresearch.org/fact-tank/2014/07/24/voter-turnout-always-drops-off-for-midterm-elections-but-why/.

3). Kennedy, Danielle Root and Liz. “Increasing Voter Participation in America.” Center for American Progress, www.americanprogress.org/issues/democracy/reports/2018/07/11/453319/increasing-voter-participation-america/.

7.2 Model Results (Frequentist)

Table 2: Frequentist Random County Intercept Model
(continued below)

	Estimates	Std.Err	P.Value
(Intercept)	-0.3872	0.05129	4.363e-14
age__bin18-29	0.1359	0.0128	2.406e-26
age__bin30-39	-0.9425	0.01313	0
age__bin40-49	-1.034	0.01346	0
age__bin50-64	-0.4105	0.01377	2.31e-195
race__ethnicityNon-Hispanic American Indian	0.5459	0.02232	4.084e-132
race__ethnicityNon-Hispanic Asian	-1.008	0.01908	0
race__ethnicityNon-Hispanic Black	2.181	0.01351	0
race__ethnicityNon-Hispanic Mixed	-1.984	0.03486	0
race__ethnicityNon-Hispanic White	2.635	0.01235	0
genderF	0.3295	0.001744	0
age__bin18-29:race__ethnicityNon-Hispanic American Indian	-0.9931	0.02669	4.882e-303
age__bin30-39:race__ethnicityNon-Hispanic American Indian	0.2893	0.02836	1.915e-24
age__bin40-49:race__ethnicityNon-Hispanic American Indian	0.5302	0.02841	1.063e-77
age__bin50-64:race__ethnicityNon-Hispanic American Indian	-0.1311	0.02696	1.163e-06
age__bin18-29:race__ethnicityNon-Hispanic Asian	0.1636	0.02163	3.893e-14
age__bin30-39:race__ethnicityNon-Hispanic Asian	0.6127	0.02214	1.361e-168
age__bin40-49:race__ethnicityNon-Hispanic Asian	1.052	0.02252	0
age__bin50-64:race__ethnicityNon-Hispanic Asian	0.5196	0.02297	2.716e-113
age__bin18-29:race__ethnicityNon-Hispanic Black	-1.824	0.01474	0

	Estimates	Std.Err	P.Value
age__bin30-39:race__ethnicityNon-Hispanic Black	-0.1712	0.01532	5.4e-29
age__bin40-49:race__ethnicityNon-Hispanic Black	-0.3045	0.01554	1.731e-85
age__bin50-64:race__ethnicityNon-Hispanic Black	-0.8052	0.01563	0
age__bin18-29:race__ethnicityNon-Hispanic Mixed	0.8526	0.03702	2.164e-117
age__bin30-39:race__ethnicityNon-Hispanic Mixed	1.943	0.03899	0
age__bin40-49:race__ethnicityNon-Hispanic Mixed	1.579	0.04125	0
age__bin50-64:race__ethnicityNon-Hispanic Mixed	0.5487	0.04225	1.451e-38
age__bin18-29:race__ethnicityNon-Hispanic White	-2.097	0.01345	0
age__bin30-39:race__ethnicityNon-Hispanic White	-0.423	0.01386	1.578e-204
age__bin40-49:race__ethnicityNon-Hispanic White	-0.6397	0.01411	0
age__bin50-64:race__ethnicityNon-Hispanic White	-0.897	0.01435	0

Table 3: Table continues below

	id
(Intercept)	(Intercept)
age__bin18-29	age__bin18-29
age__bin30-39	age__bin30-39
age__bin40-49	age__bin40-49
age__bin50-64	age__bin50-64
race__ethnicityNon-Hispanic American Indian	race__ethnicityNon-Hispanic American Indian
race__ethnicityNon-Hispanic Asian	race__ethnicityNon-Hispanic Asian
race__ethnicityNon-Hispanic Black	race__ethnicityNon-Hispanic Black
race__ethnicityNon-Hispanic Mixed	race__ethnicityNon-Hispanic Mixed
race__ethnicityNon-Hispanic White	race__ethnicityNon-Hispanic White
genderF	genderF
age__bin18-29:race__ethnicityNon-Hispanic American Indian	age__bin18-29:race__ethnicityNon-Hispanic American Indian
age__bin30-39:race__ethnicityNon-Hispanic American Indian	age__bin30-39:race__ethnicityNon-Hispanic American Indian
age__bin40-49:race__ethnicityNon-Hispanic American Indian	age__bin40-49:race__ethnicityNon-Hispanic American Indian
age__bin50-64:race__ethnicityNon-Hispanic American Indian	age__bin50-64:race__ethnicityNon-Hispanic American Indian

	id
age__bin18-29:race__ethnicityNon-Hispanic Asian	age__bin18-29:race__ethnicityNon-Hispanic Asian
age__bin30-39:race__ethnicityNon-Hispanic Asian	age__bin30-39:race__ethnicityNon-Hispanic Asian
age__bin40-49:race__ethnicityNon-Hispanic Asian	age__bin40-49:race__ethnicityNon-Hispanic Asian
age__bin50-64:race__ethnicityNon-Hispanic Asian	age__bin50-64:race__ethnicityNon-Hispanic Asian
age__bin18-29:race__ethnicityNon-Hispanic Black	age__bin18-29:race__ethnicityNon-Hispanic Black
age__bin30-39:race__ethnicityNon-Hispanic Black	age__bin30-39:race__ethnicityNon-Hispanic Black
age__bin40-49:race__ethnicityNon-Hispanic Black	age__bin40-49:race__ethnicityNon-Hispanic Black
age__bin50-64:race__ethnicityNon-Hispanic Black	age__bin50-64:race__ethnicityNon-Hispanic Black
age__bin18-29:race__ethnicityNon-Hispanic Mixed	age__bin18-29:race__ethnicityNon-Hispanic Mixed
age__bin30-39:race__ethnicityNon-Hispanic Mixed	age__bin30-39:race__ethnicityNon-Hispanic Mixed
age__bin40-49:race__ethnicityNon-Hispanic Mixed	age__bin40-49:race__ethnicityNon-Hispanic Mixed
age__bin50-64:race__ethnicityNon-Hispanic Mixed	age__bin50-64:race__ethnicityNon-Hispanic Mixed
age__bin18-29:race__ethnicityNon-Hispanic White	age__bin18-29:race__ethnicityNon-Hispanic White
age__bin30-39:race__ethnicityNon-Hispanic White	age__bin30-39:race__ethnicityNon-Hispanic White
age__bin40-49:race__ethnicityNon-Hispanic White	age__bin40-49:race__ethnicityNon-Hispanic White
age__bin50-64:race__ethnicityNon-Hispanic White	age__bin50-64:race__ethnicityNon-Hispanic White

	lower__bound	upper__bound
(Intercept)	-0.4878	-0.2867
age__bin18-29	0.1109	0.161
age__bin30-39	-0.9683	-0.9168
age__bin40-49	-1.061	-1.008
age__bin50-64	-0.4375	-0.3835
race__ethnicityNon-Hispanic American Indian	0.5022	0.5897
race__ethnicityNon-Hispanic Asian	-1.046	-0.9709
race__ethnicityNon-Hispanic Black	2.154	2.207
race__ethnicityNon-Hispanic Mixed	-2.052	-1.915
race__ethnicityNon-Hispanic White	2.611	2.659
genderF	0.3261	0.333
age__bin18-29:race__ethnicityNon-Hispanic American Indian	-1.045	-0.9408

	lower_bound	upper_bound
age_bin30-39:race_ethnicityNon-Hispanic American Indian	0.2338	0.3449
age_bin40-49:race_ethnicityNon-Hispanic American Indian	0.4745	0.5859
age_bin50-64:race_ethnicityNon-Hispanic American Indian	-0.1839	-0.07823
age_bin18-29:race_ethnicityNon-Hispanic Asian	0.1212	0.206
age_bin30-39:race_ethnicityNon-Hispanic Asian	0.5693	0.6561
age_bin40-49:race_ethnicityNon-Hispanic Asian	1.008	1.096
age_bin50-64:race_ethnicityNon-Hispanic Asian	0.4746	0.5646
age_bin18-29:race_ethnicityNon-Hispanic Black	-1.853	-1.795
age_bin30-39:race_ethnicityNon-Hispanic Black	-0.2013	-0.1412
age_bin40-49:race_ethnicityNon-Hispanic Black	-0.3349	-0.274
age_bin50-64:race_ethnicityNon-Hispanic Black	-0.8358	-0.7746
age_bin18-29:race_ethnicityNon-Hispanic Mixed	0.7801	0.9252
age_bin30-39:race_ethnicityNon-Hispanic Mixed	1.867	2.02
age_bin40-49:race_ethnicityNon-Hispanic Mixed	1.499	1.66
age_bin50-64:race_ethnicityNon-Hispanic Mixed	0.4659	0.6315
age_bin18-29:race_ethnicityNon-Hispanic White	-2.123	-2.07
age_bin30-39:race_ethnicityNon-Hispanic White	-0.4501	-0.3958
age_bin40-49:race_ethnicityNon-Hispanic White	-0.6674	-0.612
age_bin50-64:race_ethnicityNon-Hispanic White	-0.9251	-0.8689

7.3 Model Results (Bayesian)

Table 5: Bayesian Random County Intercept Model

	Estimate	Std.Err
age_bin18M29	0.1367	0.014
age_bin18M29:race_ethnicityNonMHispanicAmericanIndian	-0.9922	0.0283
age_bin18M29:race_ethnicityNonMHispanicAsian	0.1608	0.0221
age_bin18M29:race_ethnicityNonMHispanicBlack	-1.824	0.0161
age_bin18M29:race_ethnicityNonMHispanicMixed	0.8457	0.0376
age_bin18M29:race_ethnicityNonMHispanicWhite	-2.097	0.0144
age_bin30M39	-0.9418	0.0142
age_bin30M39:race_ethnicityNonMHispanicAmericanIndian	0.2897	0.0295
age_bin30M39:race_ethnicityNonMHispanicAsian	0.6109	0.0225
age_bin30M39:race_ethnicityNonMHispanicBlack	-0.1713	0.0169
age_bin30M39:race_ethnicityNonMHispanicMixed	1.935	0.039
age_bin30M39:race_ethnicityNonMHispanicWhite	-0.4237	0.0148
age_bin40M49	-1.034	0.0148
age_bin40M49:race_ethnicityNonMHispanicAmericanIndian	0.5306	0.0288
age_bin40M49:race_ethnicityNonMHispanicAsian	1.05	0.0227
age_bin40M49:race_ethnicityNonMHispanicBlack	-0.3044	0.017
age_bin40M49:race_ethnicityNonMHispanicMixed	1.572	0.0418
age_bin40M49:race_ethnicityNonMHispanicWhite	-0.6401	0.0154
age_bin50M64	-0.4095	0.0146

	Estimate	Std.Err
age_bin50M64:race_ethnicityNonMHispanicAmericanIndian	-0.132	0.0282
age_bin50M64:race_ethnicityNonMHispanicAsian	0.5165	0.0227
age_bin50M64:race_ethnicityNonMHispanicBlack	-0.8058	0.017
age_bin50M64:race_ethnicityNonMHispanicMixed	0.5418	0.0431
age_bin50M64:race_ethnicityNonMHispanicWhite	-0.8978	0.0151
genderF	0.3295	0.0018
Intercept	-0.3932	0.0514
race_ethnicityNonMHispanicAmericanIndian	0.5454	0.0235
race_ethnicityNonMHispanicAsian	-1.006	0.0195
race_ethnicityNonMHispanicBlack	2.181	0.015
race_ethnicityNonMHispanicMixed	-1.977	0.036
race_ethnicityNonMHispanicWhite	2.635	0.0134

7.4 Likely Voter Count by Count

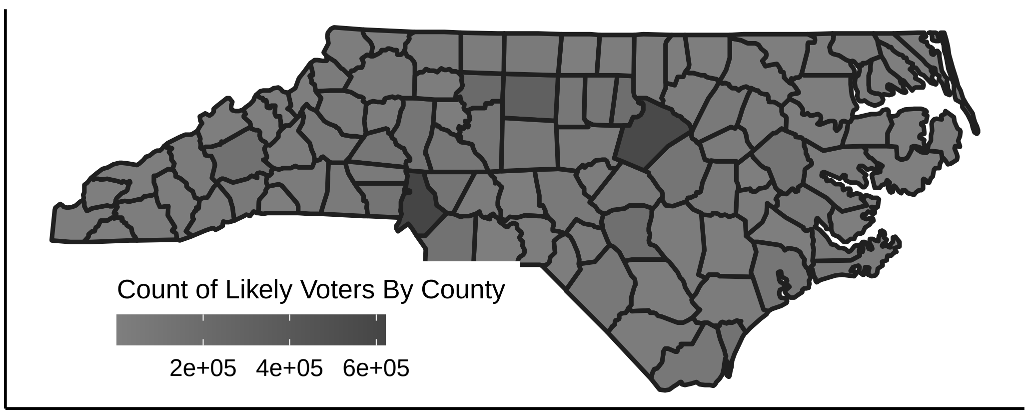


Figure 2b

7.5 Model Results

Table 6: Random County Intercept Model (continued below)

	Estimates	Std.Err	P.Value
(Intercept)	-0.3872	0.05129	4.363e-14
age_bin18-29	0.1359	0.0128	2.406e-26
age_bin30-39	-0.9425	0.01313	0
age_bin40-49	-1.034	0.01346	0
age_bin50-64	-0.4105	0.01377	2.31e-195
race_ethnicityNon-Hispanic American Indian	0.5459	0.02232	4.084e-132
race_ethnicityNon-Hispanic Asian	-1.008	0.01908	0
race_ethnicityNon-Hispanic Black	2.181	0.01351	0
race_ethnicityNon-Hispanic Mixed	-1.984	0.03486	0
race_ethnicityNon-Hispanic White	2.635	0.01235	0
genderF	0.3295	0.001744	0

	Estimates	Std.Err	P.Value
age__bin18-29:race__ethnicityNon-Hispanic American Indian	-0.9931	0.02669	4.882e-303
age__bin30-39:race__ethnicityNon-Hispanic American Indian	0.2893	0.02836	1.915e-24
age__bin40-49:race__ethnicityNon-Hispanic American Indian	0.5302	0.02841	1.063e-77
age__bin50-64:race__ethnicityNon-Hispanic American Indian	-0.1311	0.02696	1.163e-06
age__bin18-29:race__ethnicityNon-Hispanic Asian	0.1636	0.02163	3.893e-14
age__bin30-39:race__ethnicityNon-Hispanic Asian	0.6127	0.02214	1.361e-168
age__bin40-49:race__ethnicityNon-Hispanic Asian	1.052	0.02252	0
age__bin50-64:race__ethnicityNon-Hispanic Asian	0.5196	0.02297	2.716e-113
age__bin18-29:race__ethnicityNon-Hispanic Black	-1.824	0.01474	0
age__bin30-39:race__ethnicityNon-Hispanic Black	-0.1712	0.01532	5.4e-29
age__bin40-49:race__ethnicityNon-Hispanic Black	-0.3045	0.01554	1.731e-85
age__bin50-64:race__ethnicityNon-Hispanic Black	-0.8052	0.01563	0
age__bin18-29:race__ethnicityNon-Hispanic Mixed	0.8526	0.03702	2.164e-117
age__bin30-39:race__ethnicityNon-Hispanic Mixed	1.943	0.03899	0
age__bin40-49:race__ethnicityNon-Hispanic Mixed	1.579	0.04125	0
age__bin50-64:race__ethnicityNon-Hispanic Mixed	0.5487	0.04225	1.451e-38
age__bin18-29:race__ethnicityNon-Hispanic White	-2.097	0.01345	0
age__bin30-39:race__ethnicityNon-Hispanic White	-0.423	0.01386	1.578e-204
age__bin40-49:race__ethnicityNon-Hispanic White	-0.6397	0.01411	0
age__bin50-64:race__ethnicityNon-Hispanic White	-0.897	0.01435	0

Table 7: Table continues below

	id
(Intercept)	(Intercept)
age__bin18-29	age__bin18-29
age__bin30-39	age__bin30-39

	id
age__bin40-49	age__bin40-49
age__bin50-64	age__bin50-64
race__ethnicityNon-Hispanic American Indian	race__ethnicityNon-Hispanic American Indian
race__ethnicityNon-Hispanic Asian	race__ethnicityNon-Hispanic Asian
race__ethnicityNon-Hispanic Black	race__ethnicityNon-Hispanic Black
race__ethnicityNon-Hispanic Mixed	race__ethnicityNon-Hispanic Mixed
race__ethnicityNon-Hispanic White	race__ethnicityNon-Hispanic White
genderF	genderF
age__bin18-29:race__ethnicityNon-Hispanic American Indian	age__bin18-29:race__ethnicityNon-Hispanic American Indian
age__bin30-39:race__ethnicityNon-Hispanic American Indian	age__bin30-39:race__ethnicityNon-Hispanic American Indian
age__bin40-49:race__ethnicityNon-Hispanic American Indian	age__bin40-49:race__ethnicityNon-Hispanic American Indian
age__bin50-64:race__ethnicityNon-Hispanic American Indian	age__bin50-64:race__ethnicityNon-Hispanic American Indian
age__bin18-29:race__ethnicityNon-Hispanic Asian	age__bin18-29:race__ethnicityNon-Hispanic Asian
age__bin30-39:race__ethnicityNon-Hispanic Asian	age__bin30-39:race__ethnicityNon-Hispanic Asian
age__bin40-49:race__ethnicityNon-Hispanic Asian	age__bin40-49:race__ethnicityNon-Hispanic Asian
age__bin50-64:race__ethnicityNon-Hispanic Asian	age__bin50-64:race__ethnicityNon-Hispanic Asian
age__bin18-29:race__ethnicityNon-Hispanic Black	age__bin18-29:race__ethnicityNon-Hispanic Black
age__bin30-39:race__ethnicityNon-Hispanic Black	age__bin30-39:race__ethnicityNon-Hispanic Black
age__bin40-49:race__ethnicityNon-Hispanic Black	age__bin40-49:race__ethnicityNon-Hispanic Black
age__bin50-64:race__ethnicityNon-Hispanic Black	age__bin50-64:race__ethnicityNon-Hispanic Black
age__bin18-29:race__ethnicityNon-Hispanic Mixed	age__bin18-29:race__ethnicityNon-Hispanic Mixed
age__bin30-39:race__ethnicityNon-Hispanic Mixed	age__bin30-39:race__ethnicityNon-Hispanic Mixed
age__bin40-49:race__ethnicityNon-Hispanic Mixed	age__bin40-49:race__ethnicityNon-Hispanic Mixed
age__bin50-64:race__ethnicityNon-Hispanic Mixed	age__bin50-64:race__ethnicityNon-Hispanic Mixed
age__bin18-29:race__ethnicityNon-Hispanic White	age__bin18-29:race__ethnicityNon-Hispanic White
age__bin30-39:race__ethnicityNon-Hispanic White	age__bin30-39:race__ethnicityNon-Hispanic White
age__bin40-49:race__ethnicityNon-Hispanic White	age__bin40-49:race__ethnicityNon-Hispanic White

	id
age_bin50-64:race_ethnicityNon-Hispanic White	age_bin50-64:race_ethnicityNon-Hispanic White

	lower_bound	upper_bound
(Intercept)	-0.4878	-0.2867
age_bin18-29	0.1109	0.161
age_bin30-39	-0.9683	-0.9168
age_bin40-49	-1.061	-1.008
age_bin50-64	-0.4375	-0.3835
race_ethnicityNon-Hispanic American Indian	0.5022	0.5897
race_ethnicityNon-Hispanic Asian	-1.046	-0.9709
race_ethnicityNon-Hispanic Black	2.154	2.207
race_ethnicityNon-Hispanic Mixed	-2.052	-1.915
race_ethnicityNon-Hispanic White	2.611	2.659
genderF	0.3261	0.333
age_bin18-29:race_ethnicityNon-Hispanic American Indian	-1.045	-0.9408
age_bin30-39:race_ethnicityNon-Hispanic American Indian	0.2338	0.3449
age_bin40-49:race_ethnicityNon-Hispanic American Indian	0.4745	0.5859
age_bin50-64:race_ethnicityNon-Hispanic American Indian	-0.1839	-0.07823
age_bin18-29:race_ethnicityNon-Hispanic Asian	0.1212	0.206
age_bin30-39:race_ethnicityNon-Hispanic Asian	0.5693	0.6561
age_bin40-49:race_ethnicityNon-Hispanic Asian	1.008	1.096
age_bin50-64:race_ethnicityNon-Hispanic Asian	0.4746	0.5646
age_bin18-29:race_ethnicityNon-Hispanic Black	-1.853	-1.795
age_bin30-39:race_ethnicityNon-Hispanic Black	-0.2013	-0.1412
age_bin40-49:race_ethnicityNon-Hispanic Black	-0.3349	-0.274
age_bin50-64:race_ethnicityNon-Hispanic Black	-0.8358	-0.7746
age_bin18-29:race_ethnicityNon-Hispanic Mixed	0.7801	0.9252
age_bin30-39:race_ethnicityNon-Hispanic Mixed	1.867	2.02
age_bin40-49:race_ethnicityNon-Hispanic Mixed	1.499	1.66
age_bin50-64:race_ethnicityNon-Hispanic Mixed	0.4659	0.6315
age_bin18-29:race_ethnicityNon-Hispanic White	-2.123	-2.07
age_bin30-39:race_ethnicityNon-Hispanic White	-0.4501	-0.3958
age_bin40-49:race_ethnicityNon-Hispanic White	-0.6674	-0.612
age_bin50-64:race_ethnicityNon-Hispanic White	-0.9251	-0.8689

7.6 No Interaction Effect

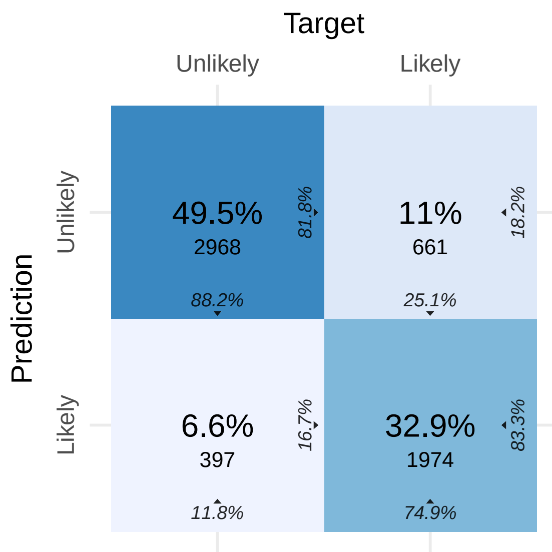
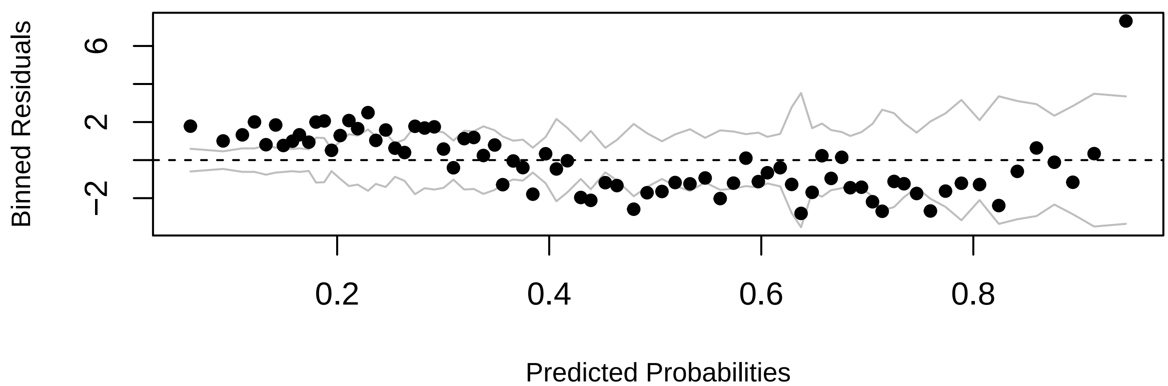
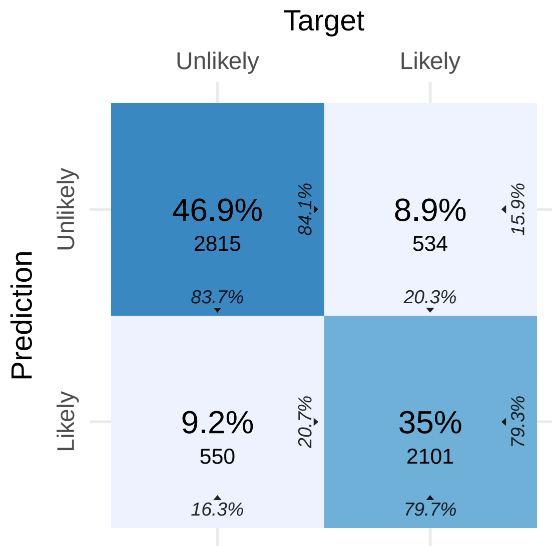


Figure 4: Random County Intercept and Age Slope Model





7.7 Age Bin Sensitivity

Table 9: Age Bin Sensitivity Analysis Model

	Estimates	Std.Err	P.Value
(Intercept)	-0.3816	0.05045	3.912e-14
age_bin18-39	-0.3414	0.01236	5.931e-168
age_bin40-64	-0.7625	0.01272	0
race_ethnicityNon-Hispanic American Indian	0.5471	0.02231	7.589e-133
race_ethnicityNon-Hispanic Asian	-1.005	0.01907	0
race_ethnicityNon-Hispanic Black	2.177	0.01351	0
race_ethnicityNon-Hispanic Mixed	-1.984	0.03486	0
race_ethnicityNon-Hispanic White	2.629	0.01234	0
genderF	0.3321	0.001733	0
age_bin18-39:race_ethnicityNon-Hispanic American Indian	-0.4286	0.02481	7.335e-67
age_bin40-64:race_ethnicityNon-Hispanic American Indian	0.2374	0.0249	1.523e-21
age_bin18-39:race_ethnicityNon-Hispanic Asian	0.336	0.02047	1.451e-60
age_bin40-64:race_ethnicityNon-Hispanic Asian	0.8239	0.02096	0
age_bin18-39:race_ethnicityNon-Hispanic Black	-1.124	0.01421	0

	Estimates	Std.Err	P.Value
age__bin40-64:race__ethnicityNon-Hispanic Black	-0.4999	0.01454	3.969e-259
age__bin18-39:race__ethnicityNon-Hispanic Mixed	1.337	0.03628	3.155e-297
age__bin40-64:race__ethnicityNon-Hispanic Mixed	1.109	0.0384	1.783e-183
age__bin18-39:race__ethnicityNon-Hispanic White	-1.369	0.01295	0
age__bin40-64:race__ethnicityNon-Hispanic White	-0.6871	0.01328	0