Political Leaning Predictions

Richa Chaturvedi, Kartik Papatla, Mirai Shah, and Ramtin Talebi

I. Introduction

In the weeks and months leading up to the 2016 election, a number of pundits and political writers wrote off the possibility of a Trump victory due to the Trump campaign's underdeveloped voter analytics team. Following the election, however, the narrative flipped – suddenly experts agreed that it was the Democrats who had spent too much time looking at data, and not enough time talking to voters. The truth actually lies somewhere in between. Analytics can inform, not replace, political campaigns on both the national and the local level.

Thus, to gain a better sense of political sentiment, our goal in this project is to predict whether an individual voter would support Democrats in the 2016 election. Obtained from BlueLabs, the original dataset has 47 different variables, ranging from education level to whether or not the person in question plays golf. Apart from the education predictor, in which 10% of its values were missing, the dataset contains mostly complete observations. Only 5 out of these 47 variables are continuous and quantitative (ppi, median_census_income, cnty_pct_evangelical, cnty_pcy_religious), while the other 42 are categorical or binary.

Preliminary visualizations of the quantitative variables (*Appendix A)* show some noteworthy trends. For instance, the age distribution of voters supporting Democrats vs. the age distribution of those not supporting Democrats in the 2016 election are quite different, with younger voters tending to support Democrats over older ones. In addition, voters living in more religious communities tended not to support Democrats while voters in less religious communities tended to support Democrats. Lastly, it appears that individuals living in areas with lower median incomes were more likely to support Democrats. We also visualized the relationship between pairs of predictor variables to gain a sense of the best interaction terms (*Appendix B)*. For instance, although we initially assumed a majority of veterans would support guns, we found that many veterans in fact had no interest in guns.

We used these preliminary data explorations and visualizations to guide our idea of which main effects and interaction terms our model should include. Moreover, as we constructed various types of models, we identified variables that could be transformed, reduced, or refactored as well as different methods to deal with missing data. Ultimately, we constructed a GAM (General Additive Model) with carefully selected main effects, tuned parameters, and significant interaction terms for our final model, yielding a log-loss score of 0.58096.

II. Initial Models and Insights

We first built a wide array of models, from basic GLMs (General Linear Models) with all 47 predictors to complex random forests with transformed variables and tuned parameters. Although likelihood ratio tests and analyses of deviance were often useful in determining the best predictors to include in the model, these tests generally indicate which models are better for inference, not necessarily for prediction. Thus, in order to determine which models would yield the best predictions on the test set, we performed cross-validation, using the log-loss evaluation metric provided in the project description. This entailed fitting the model on a randomly chosen 90% of the training data, using the model to predict on the remaining 10% of the data, and then calculating the log-loss. To ensure that we were obtaining robust results, we ran 100 simulations of this procedure to find the mean log-loss for the model. Models with lower mean log-loss values indicate better predictive potential than models with higher mean log-losses.

GLM Models

Given that the response variable is binary, we began by constructing a basic binary GLM. Four predictors had missing values – age, percent of the county that is religious, percent of the county that is evangelical, and education status. While the first three variables only had a handful of missing values (19, 2, and 5 respectively), education had 943 missing values, equivalent to nearly 10% of the entire dataset. We originally handled this missing data using the na.convert.mean function provided by Professor Glickman. This method imputes missing values in quantitative predictors with the mean of the remaining data in that column, and adds an additional column of 1's and 0's that identifies whether or not an observation is missing a value for that specific predictor. The function handles categorical predictors by adding an extra level to the predictor that represents NA values.

At first, we fit our model on all 47 main effects in the dataset using a logit link function, ensuring that missing values were accounted for via na.convert.mean. Although this framework outperformed both benchmarks on Kaggle, with a score of 0.58432 on the public leaderboard, nearly half of the predictors yielded insignificant Wald tests, indicating that many could likely be dropped without affecting our model's predictive potential. Indeed, a higher ratio of the number of predictors to the number of data points increases the chance of overfitting the data. Thus, we continued by dropping certain predictors from our model.

Given the many predictors in the dataset, running an in-depth analysis of deviance through R's anova command was difficult, particularly since we were unsure which variables would be most influential. Anova tests are best used when trying to determine the effects of one or two specific variables or when one has a general idea of what predictors to include. With 47 variables, there was a seemingly endless number of possible combinations of predictors to try and compare to one another. Therefore, we opted for the stepwise AIC method for variable selection (using the stepAIC command), which automates the process of comparing models with different sets of predictors and penalizes those with a higher number of predictors. Unlike anova tests, in which deviances of different models are compared, this method compares the Akaike Information Criterion (AIC) of models to one another and chooses the one that yields a lower AIC. After running the stepAIC command, a set of 29 predictors was outputted, including variables we would expect to be related to one's support of Democrats, such as whether an individual is a liberal donor and their education level. As a result of this variable selection, the discrepancy measure on the Kaggle leaderboard dropped from 0.58432 to 0.58383.

Moving forward, we were curious as to whether the choice of link function affects the predictive power of our model. Hence, in addition to constructing the model with 29 predictors using a logit link function, we also tested its success with a probit and complementary log-log link function. The probit function tends to predict more extreme probabilities than logit, while the complementary log-log function predicts extremely low probabilities and conservative high probabilities. Interestingly, although the complementary log-log model performed significantly worse than the logit model, the probit model had a log-loss measure that was nearly as low as the logit model. Given the way the discrepancy evaluation measure is calculated, we expect that, had our model more accurately predicted the "suppdem" value for each observation, the probit function would have yielded a lower discrepancy measure. However, when the model is not too reliable, the probit function often produces extreme, yet incorrect probabilities, which implies larger errors. Accordingly, we decided to continue using the logit function, since a probit function is best used when the model is known to be a strong fit already.

Finally, to ensure we could detect and remedy any collinearity in our model (and future models), we computed the generalized variance inflation factor for each of the predictors in our 29-predictor model (*Appendix C, Figure 7)*. Although most variance inflation measures (obtained by squaring the final column of the table) were under 10, two predictors – single and married – were understandably highly collinear.

We decided to handle this problem by collapsing the single and married columns into one predictor variable, labeled relationship status. This process led to a new insight – perhaps single and married were not the only set of redundant predictors. There might exist other predictors that were intuitively related and could be generalized into multi-factor categorical variables. This would reduce the dimensionality of our dataset, which would both control for any collinearity between predictors and simplify our own quantitative analyses. Thus, in the same way that we handled the predictors single and married, we collapsed a number of indicator variables that described different levels of the same entity into multi-factor categorical variables. For example, out of the the rural, suburban, and urban indicator variables, we constructed one categorical variable that took on the values 1 for rural, 2 for suburban, and 3 for urban. Analogously, we replaced the two variables, renters and homeowners, with a categorical predictor consisting of four levels: whether someone rented a home, owned a home, neither, or both.

We not only decreased the number of categorical variables, however. We also transformed and combined quantitative variables into variables that had greater explanatory power. Intuitively, household income (ppi) and median income individually have less explanatory power than a variable that can measure how someone earns relative to his community. After all, whether a person earns above, below, or the same relative to his peers might be an important indication of voter sentiment. We created a new variable called relative income (one's household income divided by median income of the community) to capture this effect, and removed the individual household income and median income predictors from the dataset.

GAM Models

After working extensively with GLM models, we shifted our attention to GAM models. Using GAMs allowed us to move past GLM's restrictive assumption that the relationship between the predictors and the response variable is linear. As seen by the data exploratory plots (*Appendix A*), the predictors each have different distributions, indicating unique relationships to the response variable. Through the flexible GAM framework, in which smooth, nonlinear functions can be specified for certain quantitative variables, we could individualize these various relationships between the predictor and response variable.

Indeed, we immediately saw an improvement in our score (both through our internal cross-validation function and the Kaggle scoreboard) when we applied a GAM model to our cleaned dataset. Simply by smoothing all of the quantitative variables, such as age and relative income, the log-loss score dropped to 0.58340. The success of our basic GAM model eventually led us to use this framework for our final and best-performing model. This will be discussed further in following sections.

Random Forests

Before moving onto a description of our final model, we should note that we constructed a number of random forest models, yet all performed worse than both GAM and GLM models. This was surprising to us, given that random forests are known to be one of the most accurate learning algorithms and can aptly handle datasets with large amounts of variables.

The simplest random forest model that we built involved all the variables without any tuned parameters. This yielded a score of 0.59715. We rationalized that this might be because random forest is overfitting the data and needs to be tuned. Thus, we chose three parameters to tune: number of trees (ntrees), number of variables randomly selected at each split (mtry), and the prior weights of each class (classwt).

We began by tuning ntrees and mtry in a similar manner to our cross-validation methodology. We ran 100 iterations for every combination of ntrees and mtry, made a 90%-10% split in the data, and found the mean log-loss score at the end of the 100 simulations. The combination with the lowest mean log-loss score – 1500 trees and 8 variables – were considered the optimal tuning parameters. However, even after tuning these variables, we found that our random forest was greatly overestimating the number of voters who would not support Democrats, leading to a high misclassification error. Noticing that approximately 60% of the “suppdem” column consisted of No’s, and only 40% consisted of Yes’s, we decided to tune the parameter, classwt, which ultimately returned prior class weights of 67% and 33%. After all these modifications to the random forest model, the discrepancy measure on Kaggle only dropped to 0.59632, indicating to us that we should stop pursuing this overly complex model and opt for a GAM instead.

Nevertheless, one helpful insight we gained from running random forest models was that education was ranked one of the most important features in the dataset. This was consistent with what we found with our GLM and GAM models; education was always considered statistically significant and selected during stepwise AIC selection. We realized that perhaps one way to improve our predictions, apart from testing different types of models, was to improve how we handled the over 900 missing values for education. Thus, although we continued to use the na.convert.mean function for variables with few missing values, we decided against this approach for the education column. Instead, we regressed the available education data on the other predictors. We then used this regression to predict and impute values for the missing data using the “mice” package. Alternative options for handling missing data in categorical variables are limited – including filling missing data with the mode of the predictor – so this appeared to be the best option. The regression method for imputing missing data – common in econometric analysis – proved fruitful, as the models we re-ran using this data outperformed those that came before it.

III. Final Model

Given the success of GAM in our initial modeling attempts, and the lackluster performance of random forests, we pursued GAM models further. Our final model, with a discrepancy score of 0.58096, is a binary GAM with a reduced set of predictors, smoothed variables, interaction terms, and tuned parameters. We arrived at this model through rigorous cross-validation testing, stepwise AIC selection, and a number of likelihood ratio tests to assess the significance of specific variable interactions. Our model was trained off of the adjusted dataset with 43 predictors, in which certain indicator and quantitative variables were combined into one column and education’s missing values were imputed through a regression technique. The attributes of our model can be seen in *Appendix D, Figure 8*.

In order to reduce the number of predictors in our model and determine which quantitative variables should be smoothed, we performed stepwise AIC variable selection. As previously mentioned, this method is more efficient than analysis of deviance because it automates the process of comparing hundreds of models with different combinations of predictors. After running the stepAIC function, we reduced the number of predictors to 24. Furthermore, only two of the four quantitative predictors – age and the percent of the county that is religious – were necessary to smooth, according to stepwise AIC. The smoothed variable plots, in which variables are plotted against their smooth functions, are shown in *Appendix D, Figure 9*.

In these graphs, we see that for the age variable, younger people and people about to retire are more likely to support Democrats, whereas middle-aged people and those post retirement are less likely to do so. This falls in line with our intuition about the bases of support for the Democratic party. Similarly, we see that counties with higher percentages of religious citizens are less likely to support Democrats.

Once we had our final list of main effects, we began testing different interaction terms for significance through analysis of deviance. Running stepwise AIC selection on all possible combinations of interactions had the risk of adding too many interaction terms and then causing the model to overfit the data. Thus, we only ran pairwise anova tests on interaction terms that we thought made intuitive sense to add to our model or that our earlier data visualizations suggested were relevant. For example, we tested combined_ethnicity_4way*catholic because of our knowledge that Catholic voters often vote Republican yet minority voters often support Democrats. An interaction that surprisingly yielded no interesting results was guns_1*sex, until we realized that sex perhaps was not the best interaction term for gauging gun owners' support for Democrats. Based off of our knowledge of the makeup of the NRA and the likelihood of a member of the NRA supporting a Democrat candidate, we instead interacted the guns_1 term with relative_income, which yielded significant results and was ultimately included in the final model. All in all, we incorporated four interaction terms, in addition to the 24 main effects, into our model.

We also thought it would be beneficial to tune certain parameters in the GAM model to reduce overfitting and improve the fit of our model. Although the tuning parameter, λ, which balances the trade-off between model fit and model wiggliness, is automatically optimized through Generalized Cross Validation, we also decided to tune two additional parameters in the model – epsilon and BF epsilon. These parameters are useful in controlling the convergence and numerical stability of the model fit. Specifically, epsilon controls the convergence threshold for local scoring, and BF epsilon controls the threshold for the number of back-fitting iterations. In order to tune both parameters, we made a 90%-10% split in the data and then tested different combinations of epsilon and BF epsilon over multiple iterations. After taking the mean log-loss score for every combination of these two parameters, we settled on the values of epsilon = .001 and bf.epsilon = .001, which we then plugged into our final model.

Diagnostics of the Final Model

We performed diagnostics on the model to determine if there were any clear outliers in the data or problematic patterns in the spread of residuals. After plotting the Average Binned Fitted Values against Average Binned Residuals, we found a slightly positive linear relationship between fitted values and residuals. Nevertheless, this trend likely holds limited significance, as our main goal is to avoid curvature, not linear patterns. Moreover, the plot of jackknifed residuals displayed points primarily falling between the 2 and -2 range, implying that most residuals were not too extreme in magnitude. Thus, both residual plots (*Appendix E, Figure 10 and 12)* assured us that there were no glaring flaws in the model. Additionally, the Cook's distances plot, which can identify influential points in the data, produced satisfying results (*Appendix E, Figure 11)*. All of the Cook's distances were significantly below 1, indicating that there were no influential points affecting our model.

Analysis of Model Coefficients

Appendix F, Figure 14 shows the predictor coefficients for the final model. The coefficients for both smoothed predictors – age and the percent of the population who reported religiosity in the subject's country – are positive. However, because we have applied smoothing functions to these predictors, we cannot interpret these coefficients easily; we should instead rely on graphical representations of these predictors to glean the relationships between them and the response. Looking at the plots in *Appendix D, Figure 9*, then, we see polynomial functions – closely resembling cubic functions – that show an overall negative relationship between increased age and religiosity in a county and the functions used to smooth them. Because these functions trend negative, we can say that when age and and county religiosity increase, the probability of a subject supporting a Democratic candidate usually decreases, which matches our intuition on these relationships.

Other predictors had coefficient estimates that were more interpretable and mostly unsurprising. For example, liberal_donor has a large positive coefficient while guns_1 and conservative_donor had large negative ones. Beyond individual predictors that were either smoothed or left parameterized, some of the interaction terms are worth nothing as well. Research shows that the Catholic vote is generally split across party lines, but that minority Catholics generally are left-leaning. This might explain why ethnicity interacted with Catholic in our model has a large, positive coefficient. Following similar rationale couched in political theory and existing research, we explored interactions amongst a number of other predictors, though ethnicity*catholic is the only that appears significant at the 95% level. In future iterations of the model, we might wish to remove these other interactions and explore others further.

IV. Conclusion

Our dataset, provided by BlueLabs, consisted of 47 predictor variables and a binary response variable, "suppdem." After performing stepwise AIC and likelihood ratio tests, as well as condensing redundant information in the dataset into single variables, our final binary GAM model consisted of 24 single variables and 4 interaction terms.

There are numerous aspects of the model that surprised us, as well as things to improve on in the future. Our original hypothesis was that a random forest model would perform better than a GAM model considering that many of the predictors in the dataset were binary variables that therefore could not be smoothed. Yet, perhaps this greater control over smoothing was an important reason why our GAM model succeeded over random forests. A common issue that machine learning methods like random forest, trees, and support vector machines face is the bias-variance tradeoff. That is, the fewer iterations the machine does, the more susceptible it is to returning results that are biased by a few, relevant portions of the data, i.e., underfitting. On the other hand, the more iterations it does, the more likely it is to overfit the data to accommodate outliers, thus increasing the variance of the predictions. By fitting specific smoothing functions to only certain predictors, we can capture nonlinearities in the data that linear models cannot without overfitting like random forests are prone to. We thus avoid the bias-variance tradeoff.

That said, we still ran into problems of overfitting, which became clear when we dropped from 9th on the public leaderboard to 25th in the final rankings. By continuously revising our model until it was highly ranked on the public leaderboard, we accidentally overfit our data to that small subset of the test data.

Given our relatively high log-loss score, there are certain improvements that must be made in order to yield a better fitting model. One approach for future exploration is a more critical evaluation of the variables that appeared in the final model – cat_1, for example, consistently appeared in models outputted by stepwise AIC. This variable encoded whether a person in question was likely a cat owner or not. Our intuition led us to believe that this variable captured information that was not directly relevant to the prediction of the response variable, yet we kept this variable in due to the results of the automated process. This may have contributed to the overfitting issue we experienced in the final scoring. Furthermore, the addition of interaction terms, while significant in our likelihood ratio tests, did not improve the final scoring as much as we had hoped.

The greatest gains in this model were realized following the reduction of the number of predictors in the dataset as well as the use of a revised missing data imputation approach. These steps greatly reduced the dimensionality and inaccuracies of the data that we fit our model on, and therefore it makes sense that vast improvements in the log-loss score were subsequently seen.

Ultimately, this project helped us understand that there is no one right way to fit a dataset – different models have relative strengths and weaknesses that become evident through an iterative process of testing and cross-validating various approaches. While we eventually decided on a GAM, a GLM approach could have also provided similar results. Our decisions throughout this process relied on how we understood the dataset, our intuitions, and our priorities as data scientists.

Appendix A: Data Visualization

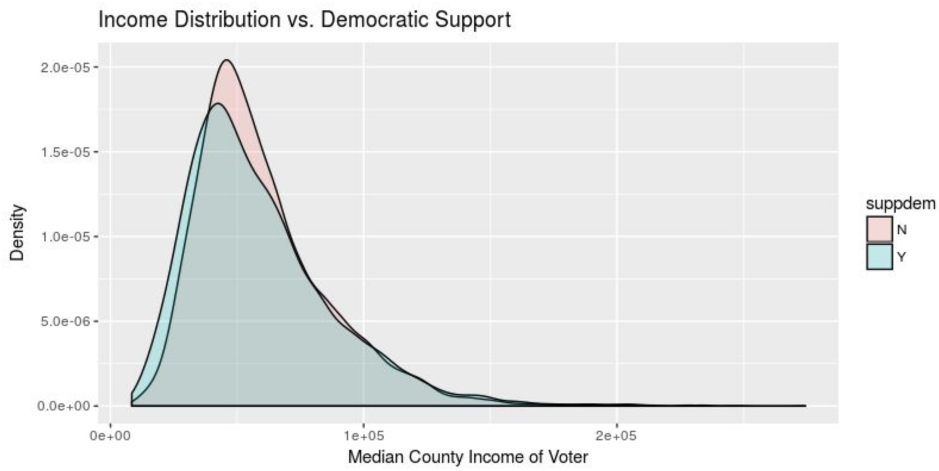

Figure 1: Distribution of Income for Democratic Support

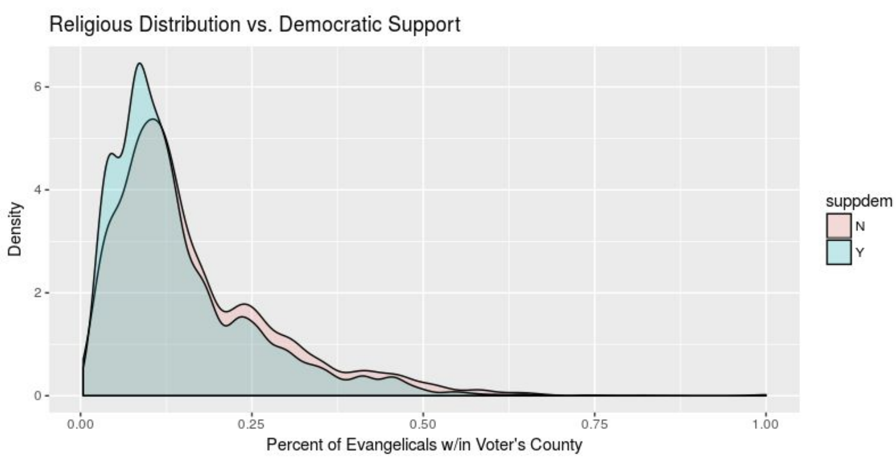

Figure 2: Distribution of Age for Democratic Support

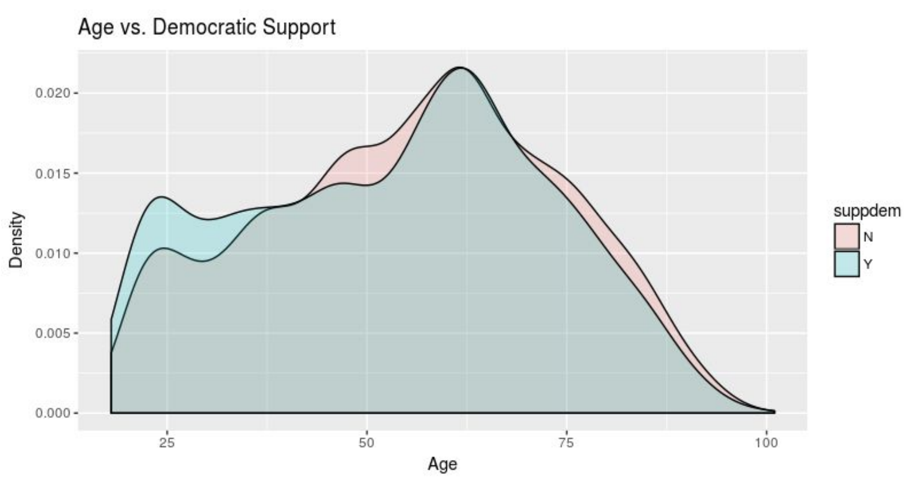

Figure 3: Religious Distribution by County for Voter Preference

Appendix B: Selection of Indicator Interaction Jitter Plots

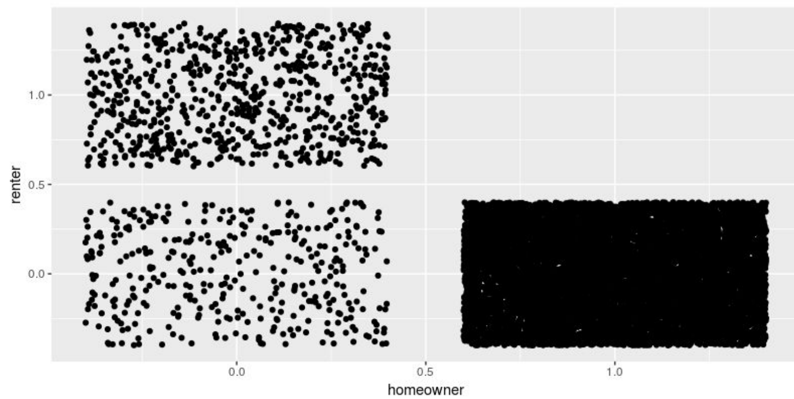

Figure 4: Voters that are Homeowners vs Renters

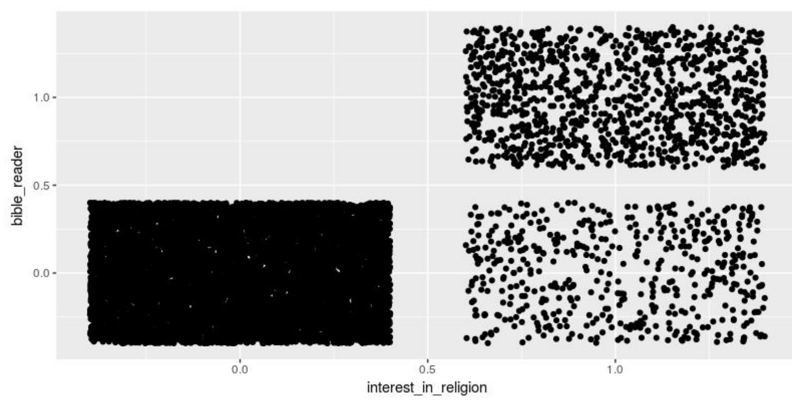

Figure 5: Voters with Interests in Religion vs. Bible Readers

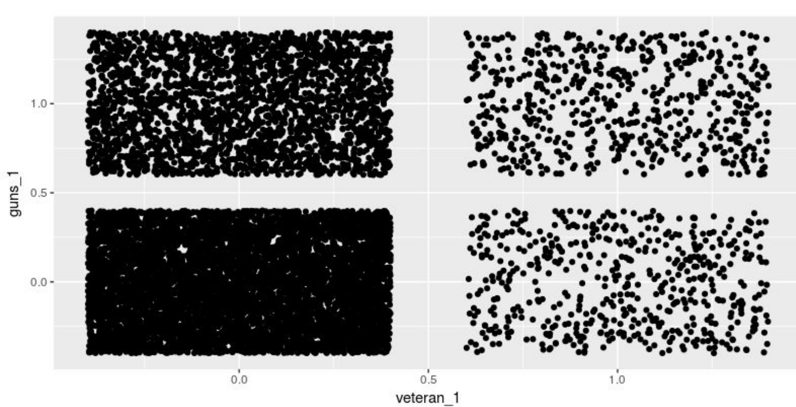

Figure 6:Voters with Interests in Guns vs. Veterans

Appendix C: Collinearity Analysis

	GVIF	Df	GVIF^(1/(2*Df))
density_rural	2.194973	1	1.481544
density_suburban	1.925995	1	1.387802
sex	1.126487	2	1.030224
combined_ethnicity_4way	1.471220	3	1.066464
ppi	1.107287	1	1.052277
single	215.989323	1	14.696575
married	216.844718	1	14.725648
num_children	1.068194	1	1.033535
renter	1.192623	1	1.092073
education	1.464963	5	1.038921
hasreligion	2.575981	1	1.604986
catholic	2.057136	1	1.434272
christian	1.284460	1	1.133340
interest_in_religion	1.438814	1	1.199506
donrever_1	1.705123	1	1.305804
liberal_donor	1.150449	1	1.072590
conservative_donor	1.027234	1	1.013525
contbrel_1	1.054768	1	1.027019
blue_collar	1.149844	1	1.072308
retired	1.208937	1	1.099517
apparel_1	1.317704	1	1.147913
boatownr_1	1.168435	1	1.080942
cat_1	1.282918	1	1.132660
outdgrdn_1	1.827102	1	1.351704
outdoor_1	2.140605	1	1.463081
guns_1	1.462500	1	1.209339
cnty_pct_religious	1.135144	1	1.065431
cnty_pct_evangelical	1.291101	1	1.136266
cnty_pct_religious.na	1.002847	1	1.001423

Figure 7: VIF Chart of Selected Predictors.
**Note high VIFs for single and married status*

Appendix D: Final Model Summary

```
model_final_gam_tuned = gam(suppdem ~ s(age) + sex + combined_ethnicity_4way +
                             relationship_status + num_children + hasreligion + catholic +
                             christian + interest_in_religion + donrever_1 + liberal_donor +
                             conservative_donor + contbrel_1 + apparel_1 + boatownr_1 +
                             cat_1 + environm_1 + outdgrdn_1 + guns_1 + s(cnty_pct_religious) +
                             cnty_pct_evangelical + district_status + collapsed_educ +
                             relative_income + liberal_donor*collapsed_educ +
                             combined_ethnicity_4way*catholic + sex*combined_ethnicity_4way +
                             guns_1*relative_income,epsilon = .001, bf.epsilon = .001,
                           family = binomial, data = train.na, trace = FALSE)
```

Figure 8: Final Model

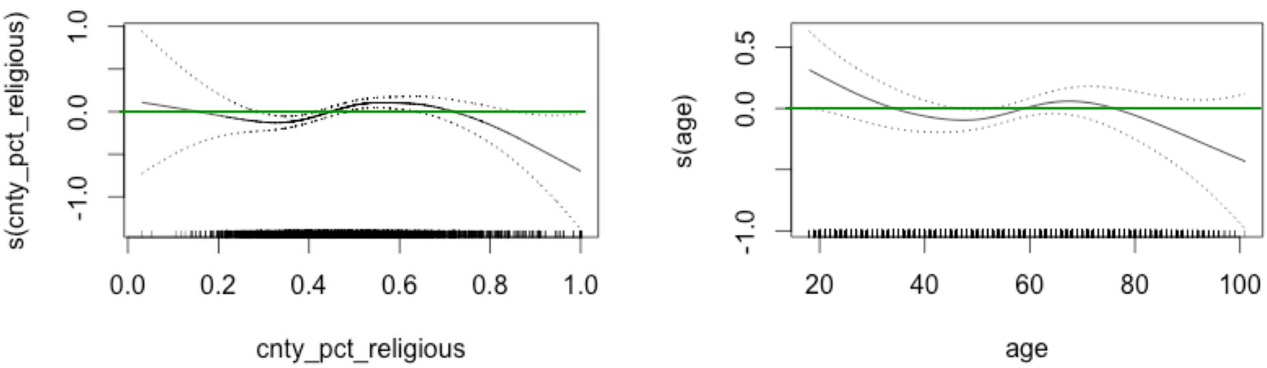

Figure 9: Plot of Smoothed Variables

Appendix E: Diagnostic Plots

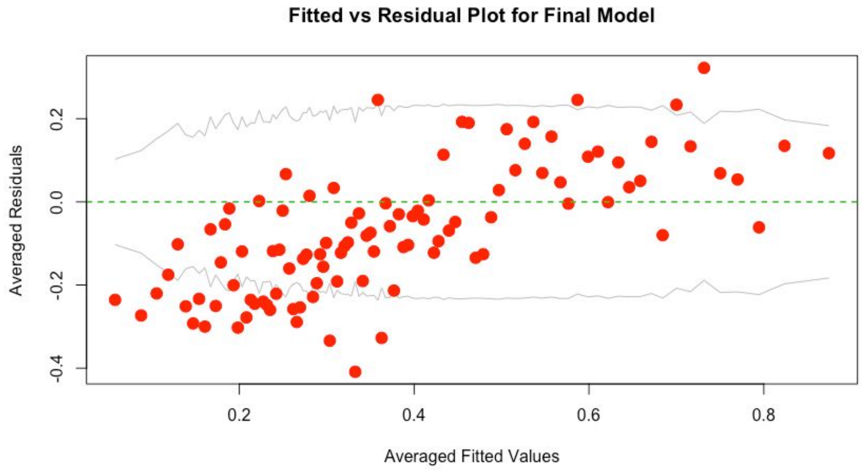

Figure 10: Average Binned Fitted Values vs. Average Binned Residuals Plots

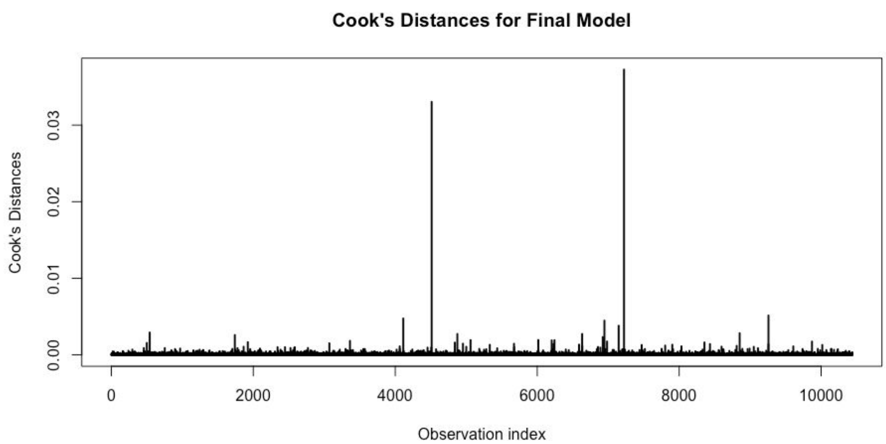

Figure 11: Cook's Distances

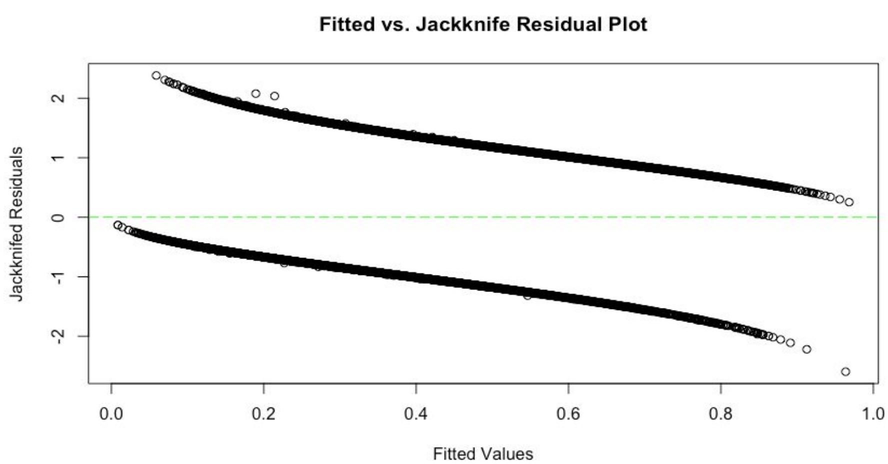

Figure 12: Fitted Values vs Jackknifed Residuals

Appendix F: Analysis of Coefficients

```
Deviance Residuals:
    Min      1Q  Median      3Q     Max 
-2.3819 -0.9002 -0.6562  1.0925  2.3313 

(Dispersion Parameter for binomial family taken to be 1)

    Null Deviance: 13873.28 on 10438 degrees of freedom
Residual Deviance: 12312.17 on 10395 degrees of freedom
AIC: 12400.17 

Number of Local Scoring Iterations: 3 

Anova for Parametric Effects
                                  Df  Sum Sq Mean Sq  F value    Pr(>F)    
s(age)                             1    25.3  25.333  25.3647 4.825e-07 ***
sex                                2    74.3  37.128  37.1748 < 2.2e-16 ***
combined_ethnicity_4way            3   587.2 195.725 195.9722 < 2.2e-16 ***
relationship_status                1    42.0  41.994  42.0472 9.314e-11 ***
num_children                       1     8.3   8.324   8.3341 0.0038987 ** 
hasreligion                        1    10.2  10.180  10.1929 0.0014140 ** 
catholic                           1    30.4  30.400  30.4386 3.529e-08 ***
christian                          1    11.9  11.943  11.9580 0.0005463 ***
interest_in_religion               1    16.2  16.159  16.1792 5.803e-05 ***
donrever_1                         1     8.8   8.800   8.8109 0.0030012 ** 
liberal_donor                      1    99.5  99.546  99.6719 < 2.2e-16 ***
conservative_donor                 1    25.8  25.781  25.8137 3.825e-07 ***
contbrel_1                         1    15.6  15.633  15.6526 7.661e-05 ***
apparel_1                          1     4.3   4.330   4.3354 0.0373515 *  
cat_1                              1    18.8  18.761  18.7845 1.477e-05 ***
environm_1                         1     3.0   3.025   3.0283 0.0818522 .  
outdgrdn_1                         1     6.8   6.772   6.7801 0.0092312 ** 
guns_1                             1    30.1  30.087  30.1246 4.147e-08 ***
s(cnty_pct_religious)              1     1.1   1.133   1.1341 0.2869226    
cnty_pct_evangelical               1    77.3  77.307  77.4045 < 2.2e-16 ***
district_status                    1    82.1  82.133  82.2371 < 2.2e-16 ***
collapsed_educ                     1    21.3  21.308  21.3346 3.903e-06 ***
relative_income                    1     7.0   6.989   6.9981 0.0081717 ** 
liberal_donor:collapsed_educ       1     0.0   0.003   0.0035 0.9528429    
combined_ethnicity_4way:catholic   3     8.9   2.969   2.9728 0.0304414 *  
sex:combined_ethnicity_4way        6    10.1   1.687   1.6891 0.1192229    
guns_1:relative_income             1     0.6   0.579   0.5802 0.4462708    
Residuals                      10395 10381.9   0.999                      
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figure 13: Final Model Output

```
> coef(model_final_gam_tuned)
                       (Intercept)                             s(age)                               sexM
                     -6.163421e-02                       1.540155e-05                      -7.927642e-01
                              sexU           combined_ethnicity_4wayB           combined_ethnicity_4wayH
                     -8.751433e-01                       3.507446e-01                      -4.659954e-01
          combined_ethnicity_4wayW                relationship_status                       num_children
                     -1.231229e+00                      -2.294742e-01                      -7.182679e-02
                       hasreligion                           catholic                          christian
                     -2.664253e-01                      -1.345873e+00                       2.717569e-01
              interest_in_religion                         donrever_1                      liberal_donor
                     -4.392662e-01                       1.444007e-01                       8.500071e-01
                conservative_donor                         contbrel_1                          apparel_1
                     -1.502132e+00                      -4.965182e-01                       1.746000e-01
                             cat_1                         environm_1                         outdgrdn_1
                     -1.658849e-01                       8.749347e-02                      -1.141585e-01
                            guns_1              s(cnty_pct_religious)                cnty_pct_evangelical
                     -2.255557e-01                       3.679553e-01                      -1.233138e+00
                   district_status                     collapsed_educ                     relative_income
                      3.255264e-01                       1.182617e-01                       6.838919e-02
      liberal_donor:collapsed_educ combined_ethnicity_4wayB:catholic combined_ethnicity_4wayH:catholic
                      4.079995e-03                       1.441497e+00                       1.723933e+00
combined_ethnicity_4wayW:catholic      sexM:combined_ethnicity_4wayB      sexU:combined_ethnicity_4wayB
                      1.599794e+00                       4.449936e-01                       5.599570e-01
     sexM:combined_ethnicity_4wayH      sexU:combined_ethnicity_4wayH      sexM:combined_ethnicity_4wayW
                      1.104713e-01                       8.918339e-02                       4.405302e-01
     sexU:combined_ethnicity_4wayW              guns_1:relative_income
                      9.887772e-01                      -3.313326e-02
```

Figure 14: Coefficients of Final Model