OBAMA TO TRUMP

Ajinkya Pawale, Shivani Vogiral, Sharanbasav A Sumbad, Anuhya Sankranti

Introduction

To what extent do attitudes toward immigration explain why 2012 Obama supporters switched to Trump supporters in 2016? We try to answer this question using data from the 2016 Cooperative Congressional Election Study. All people who voted for Barack Obama in the 2012 Presidential general election and were eligible to vote in the 2016 Presidential general election are included in the demographic of interest. In this project, we'll look at how individual demographic characteristics like gender, education, race, and political party identification can explain why people are voting for Donald Trump. In addition, we also try to determine whether immigration attribute makes a substantive difference for estimating the probability of switching. Also, are there some demographic variables that are more important than the others?

Data Preprocessing

Before we begin with any analysis it is important to perform some necessary data preprocessing steps. We created a data frame called **Obama** than contains the following attributes:

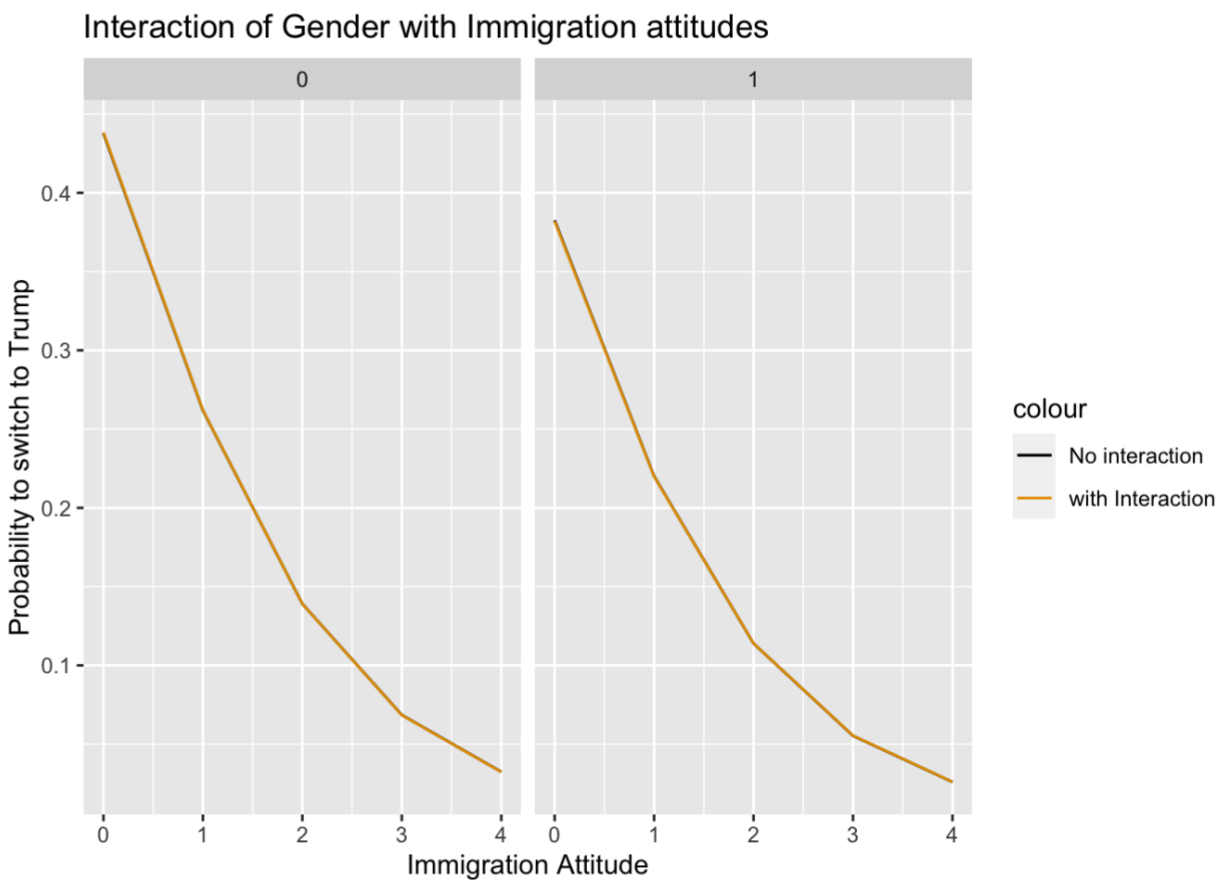
- Commonweight_vv_post: The survey weights for people who took the post-election survey.
- Gender: Male or Female
- Educ: An ordered factor with six levels of education
- Race: A factor with with 4 main races - White, Black, Hispanic, Others
- Pid7: An ordered factor with seven levels from "Strong Democrat" to "Strong Republican."
- Pro.img: a quantitative variable that measures the respondent's attitude toward immigration using the four immigration variables. The variable has a range of values between 0 to 4 with 4 being the most pro-immigration and 0 the least.
- isTrump - A binary variable that has values "Yes" indicating that the respondent voted for Trump in 2016 and "No" meaning that the respondent's vote is for Obama (his/her vote being the same as that in 2012)

Interactions between immigration attitudes and demographic variables

In this section, we analyze the effects of the Immigration Attitude variable on all the demographic variables: Gender, Race, Education, Party Identification

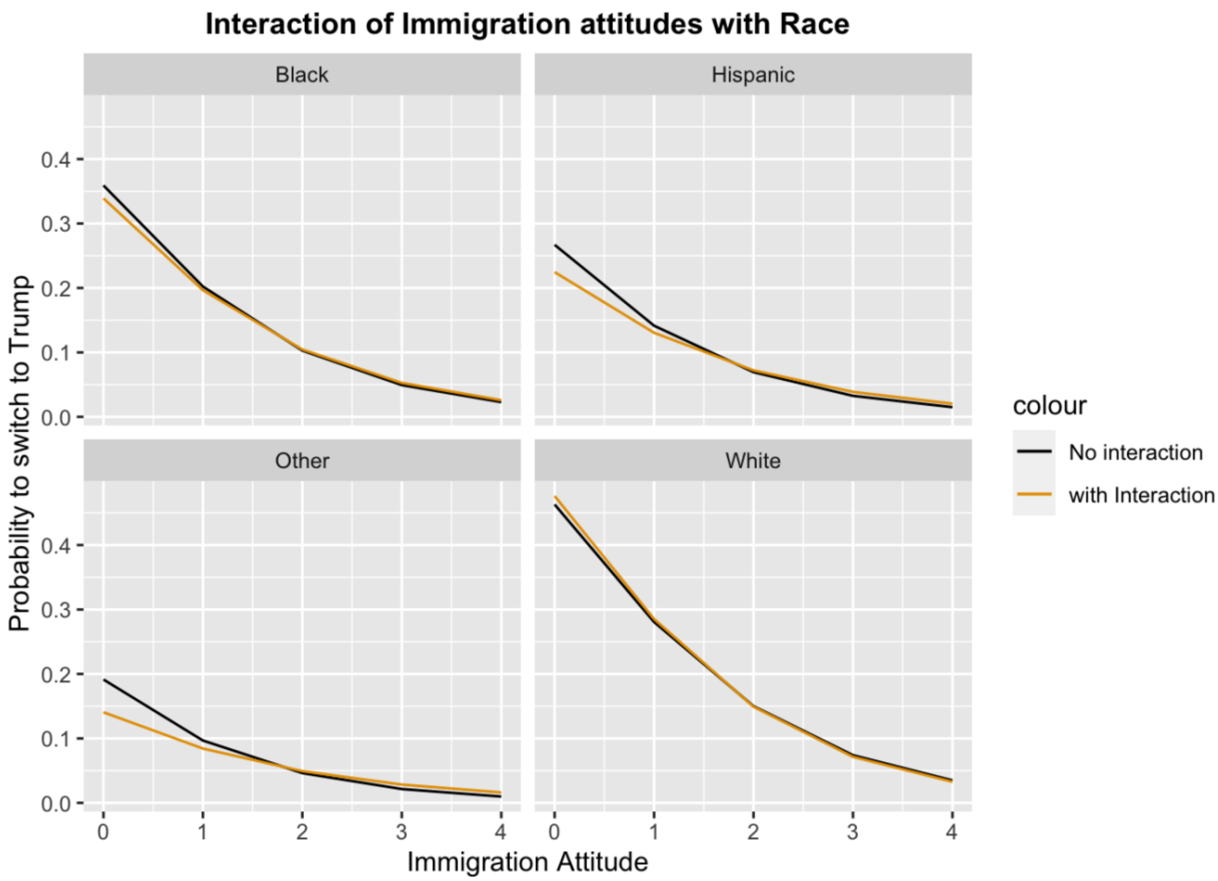
Interaction of Gender and Immigration Attitude v/s Probability of Switching to Trump:

For the below purpose, **With Interaction** stands for a multiplicative model between the two groups and **No Interaction** stands for addictive model between the two groups.



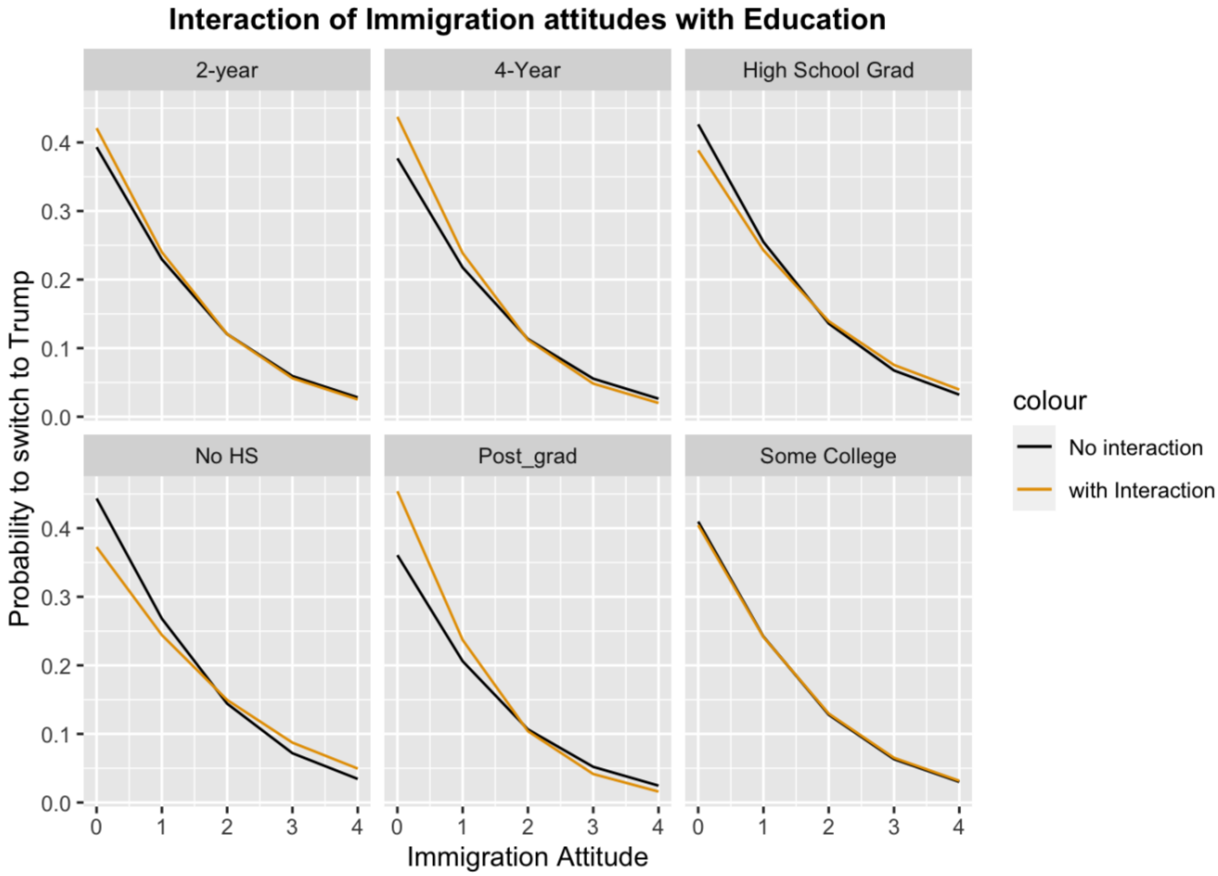
The probabilities of switching to Trump predicted by two models (with interaction and without interaction) are the same in the above plots on gender and immigration opinions interaction. As a result, gender does not explain any differences in male and female attitudes toward immigration.

Interaction of Race and Immigration Attitude v/s Probability of Switching to Trump:



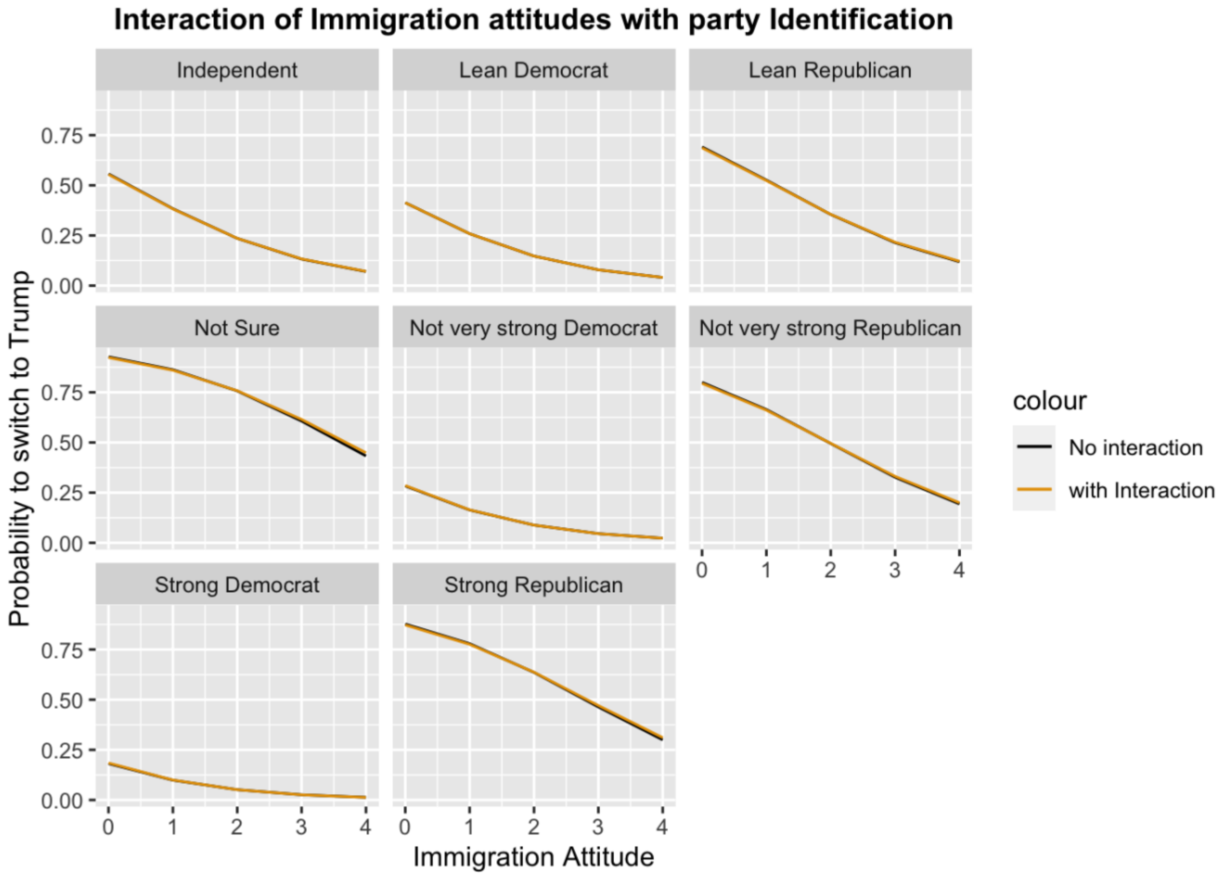
We can observe that there are interactions between race and immigration attitude variables in the plot above. In the Hispanic, black, and other levels of race, there is some significant difference between the interaction and no interaction curve which suggests that Race along with Immigration can be considered as an interaction for creating the model. The chance of switching to Trump is slightly lower for those attributes who have the most negative attitude about immigration when using the interaction model than when using the no interaction model.

Interaction of Education and Immigration Attitude v/s Probability of Switching to Trump:



We can observe that there are some relationships with immigration attitudes at six levels of schooling. The No high school and High School Grad levels of education demonstrate similar relationships. In both, when compared to the no interaction model, we receive somewhat reduced probabilities for the most unfavorable immigration attitude response with interaction. Post-Grad, 2-year, 4-year: For the most unfavorable response to immigration attitudes, we get somewhat higher probabilities with interaction than with no interaction. Thus, education can be considered to have an interaction with the Immigration attitude.

Interaction of Party Identification and Immigration Attitude v/s Probability of Switching to Trump:



We can see from the above plots that party identification has little bearing on immigration attitudes when estimating the likelihood of a Trump switch. Because there is no interaction between the two variables, there is no benefit to considering the demographic variable of party identification along with the immigration attitude.

Logistic Regression Model without Immigration Attitude:

The logistic regression model described here predicts the probability of a 2012 Obama voter switching to Trump in 2016 without using the Immigration Attitude variable.

```
```{r}
model1.logit=glm(isTrump~ pid7_num+ educ_num + race_num + gender_num
,family="binomial",data=obama,weights = obama$commonweight_vv_post)
summary(model1.logit)
```
```

```
Call:
glm(formula = isTrump ~ pid7_num + educ_num + race_num + gender_num,
    family = "binomial", data = obama, weights = obama$commonweight_vv_post)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.5957  -0.3797  -0.2446  -0.1653   9.5497

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.93662    0.07613 -38.574  < 2e-16 ***
pid7_num     0.64952    0.01392  46.651  < 2e-16 ***
educ_num    -0.19474    0.01645 -11.835  < 2e-16 ***
race_num    -0.31425    0.03114 -10.091  < 2e-16 ***
gender_num  -0.19025    0.04855  -3.919 8.89e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 14691  on 23376  degrees of freedom
Residual deviance: 11944  on 23372  degrees of freedom
  (18 observations deleted due to missingness)
AIC: 12224
```

This is an additive model consisting of Party Identification, Education, Race and Gender. It is a weighted model, so the weights are also specified.

Here we can see that the coefficient for Part Identification i.e., pid7_num is the highest (0.65).
All the other variables have a negative impact with race having the most negative impact (-0.31), race and gender have similar negative impact (-0.19).

Logistic Regression Model with Immigration Attitude:

The logistic regression model described here predicts the probability of a 2012 Obama voter switching to Trump in 2016 that makes use of the Immigration Attitude variable.

```
```{r}
model2.logit=glm(isTrump~ pro.img:educ_num
+pid7_num+pro.img:race_num+gender_num,family="binomial",data=obama,weights = obama$commonweight_vv_post)
summary(model2.logit)
```
```

```
glm(formula = isTrump ~ pro.img:educ_num + pid7_num + pro.img:race_num +
    gender_num, family = "binomial", data = obama, weights = obama$commonweight_vv_post)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-6.1887  -0.3437  -0.1843  -0.0928   9.6830

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      -2.358996   0.064039 -36.837  < 2e-16 ***
pid7_num          0.592081   0.014449  40.979  < 2e-16 ***
gender_num       -0.175029   0.050517  -3.465 0.000531 ***
pro.img:educ_num -0.147876   0.005204 -28.416  < 2e-16 ***
pro.img:race_num -0.142057   0.014797  -9.600  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

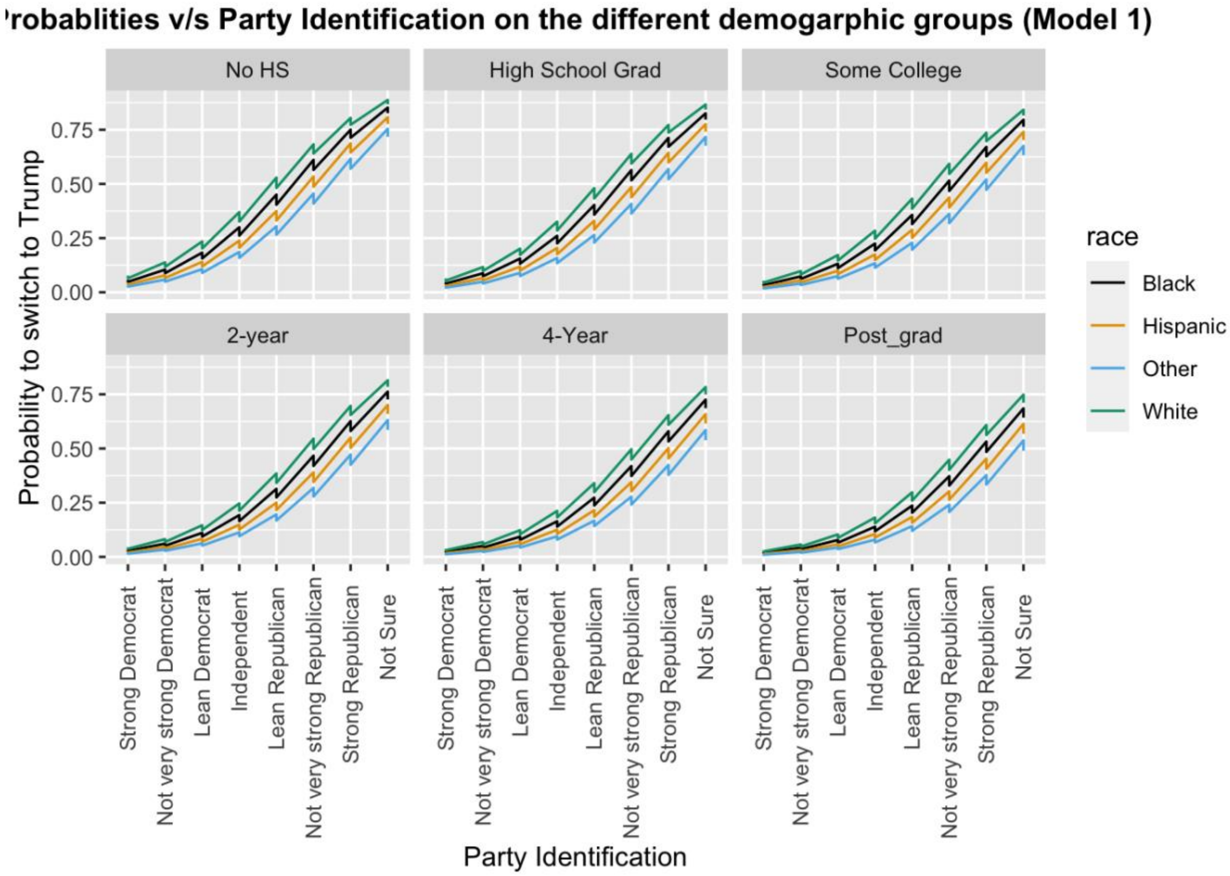
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 14691  on 23376  degrees of freedom
Residual deviance: 10874  on 23372  degrees of freedom
  (18 observations deleted due to missingness)
AIC: 11100
```

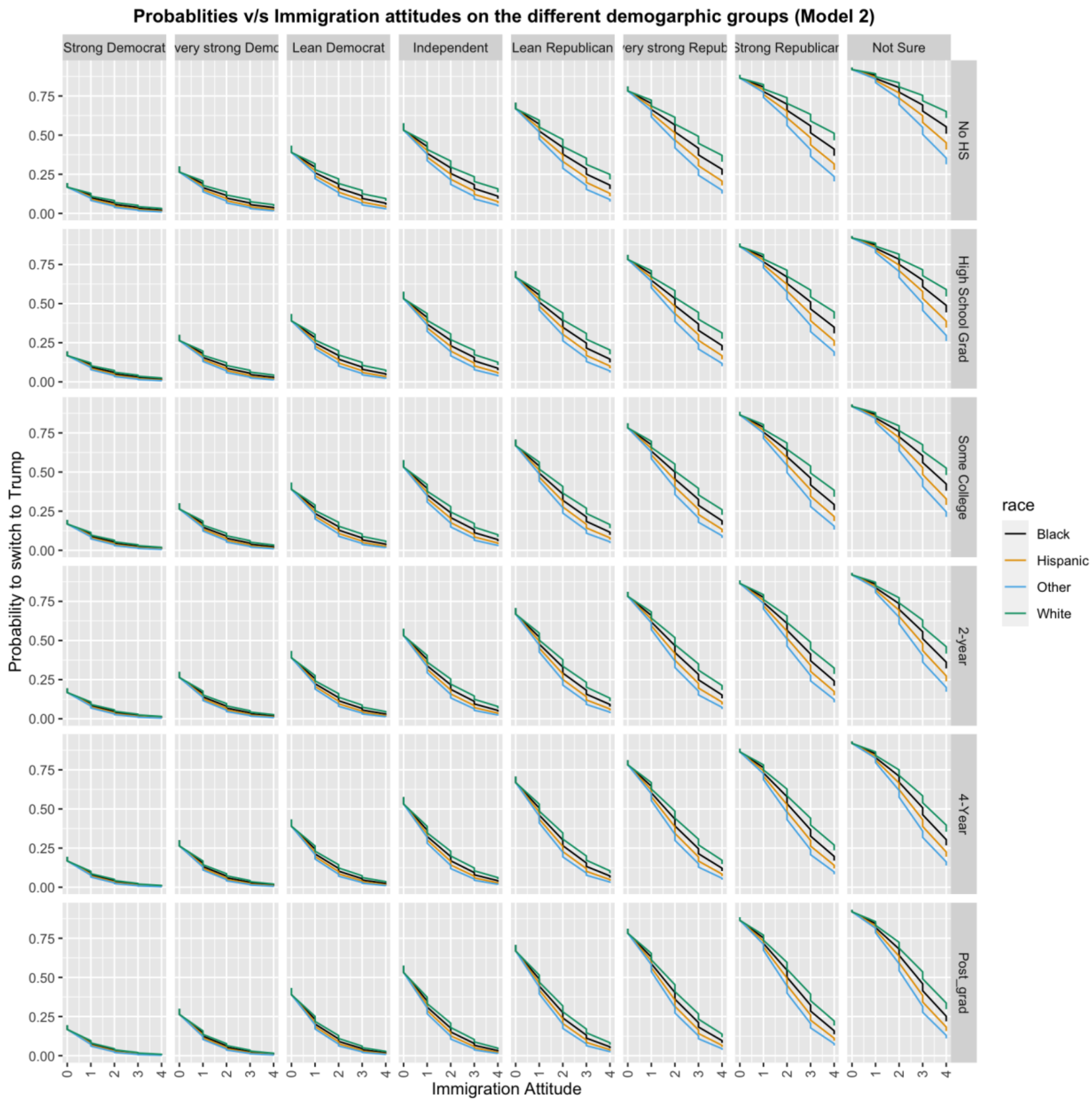
In this we have considered all the demographic variables along with Pro Immigration. In Part 2 we saw that Immigration attitude with education and with race as an additive model and as an interaction model shows separate curves which suggests that Immigration attitude makes a difference when taken as an interaction with gender and race, thus we have taken their interaction over here. We did the same for Immigration attitude with gender and party identification but did not get different curves with or without interaction thus we took gender and party identification without the Immigration attitude interaction.

Here we can see that the coefficient for Part Identification i.e. pid7_num is the highest (0.59). The other variables have a negative impact with gender having the most negative impact (-0.17), pro_immigration:education and pro_immigration:race have similar negative impact (-0.14).

Influence of Immigration Attitude of computing probability of switching to Trump in 2020:



Here we can see that according to the party identifier, if the person falls under the category of Democrat, the probability of switching to Trump is less but if the person falls under the category of Republican the probability of switching to trump is more. Also, based on race we can see that the curves follow the following order of switching to Trump: White > Black > Hispanics > Other. Education doesn't offer much information as the trends are similar for all the education groups.



Here we have added the Immigration attitude variable along with the other demographic variables. We can see that according to the party identifier, if the person falls under the category of Democrat the probability of switching to Trump is less but if the person falls under the category of Republican the probability of switching to trump is more. Also, Immigration attitude provides additional information here, if the person is strongly against immigration i.e., pro-immigration = 0 then the probability of switching is more as compared to the ones with higher pro-immigration i.e., pro-immigration = 1,2,3,4. This immigration trend is visible in all the facets and groups. The trend for race is similar as above i.e., the curves follow the following order of switching to Trump: White > Black > Hispanic > Other. Education as above doesn't provide any additional information; the trends are similar for all the education groups.

Does including the Immigration Attitude variable matter more for some demographic groups than others?

The logistic regression models with and without interaction are quite different in general, demonstrating that immigration opinions have a lot of predictive value even after demographic variables are considered.

For instance, while race is well-known to play a significant role in voting, the models show that minorities who were strongly anti-immigration were more likely to switch than whites who were very pro-immigration (keeping gender and party constant).

When we compare both the models and plot the probabilities, we find substantive differences with and without immigration attitude, for model 1 where immigration attitude is not considered, we see increasing probabilities as we move towards strong republicans closing towards 0.75. With White race being the one with higher probability. But when we introduce immigration attitude, we see decreasing trend even with strong republicans i.e., as people are more supportive towards immigration, the probability for switching to trump decreases, even for the white race, it's the same trend. Comparatively No High school demographic shows higher probability when compared with other education levels, the same can be observed with white race, the probabilities switching to trump for white is a bit high when compared with other races, but overall, it is decreasing trend in probabilities when immigration attitude is considered.

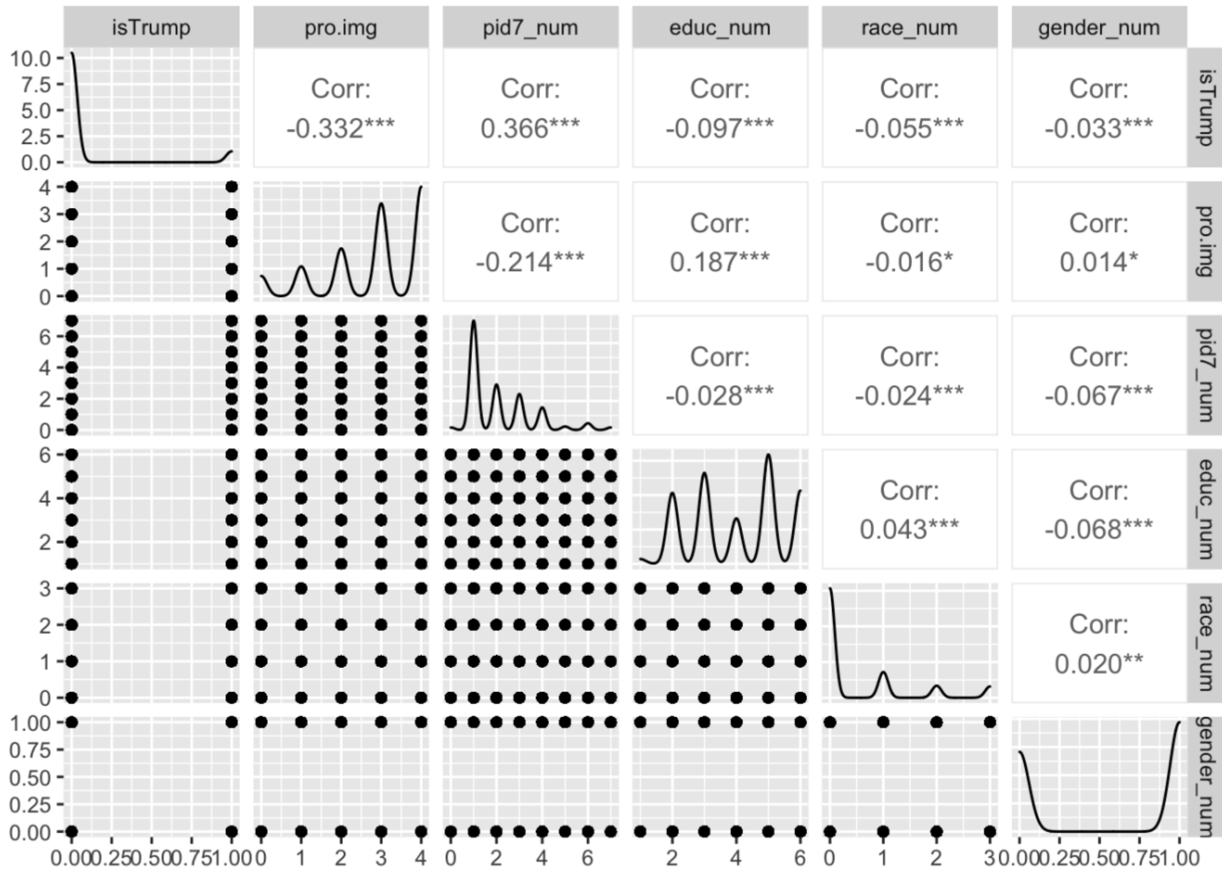
To summarize, we can say that immigration variable has an interaction with the 2 demographic variables i.e., race and education. It does not have an interaction with gender and party identification.

Conclusion:

In conclusion, the voting attitude of Republicans is not very surprising and as expected these voters have a high probability of switching their vote to Trump in the 2016 elections. An important observation from the data is that around 25% of independent Obama voters seemed to have switched their vote to Trump in the 2016 elections. This trend looks very unusual and with more information from the previous elections we can analyze the switching behavior more intricately. Perhaps even more importantly, attitudes toward immigration were a strong predictor of switching to Trump. A significant portion of the Obama voters who were least favorable for immigration switched their vote to Trump in the 2016 elections. Immigration opinions continue to be a bigger factor of vote switching among white Obama voters and Obama voters with at least a bachelor's degree. Speculatively, the Democratic Party might be able to reclaim these people without changing policy by making immigration less of a priority, though this would likely result in the loss of other voters and is, in any event, easier said than done. The Republican Party, on the other hand, may benefit from making immigration a significant issue for the foreseeable future.

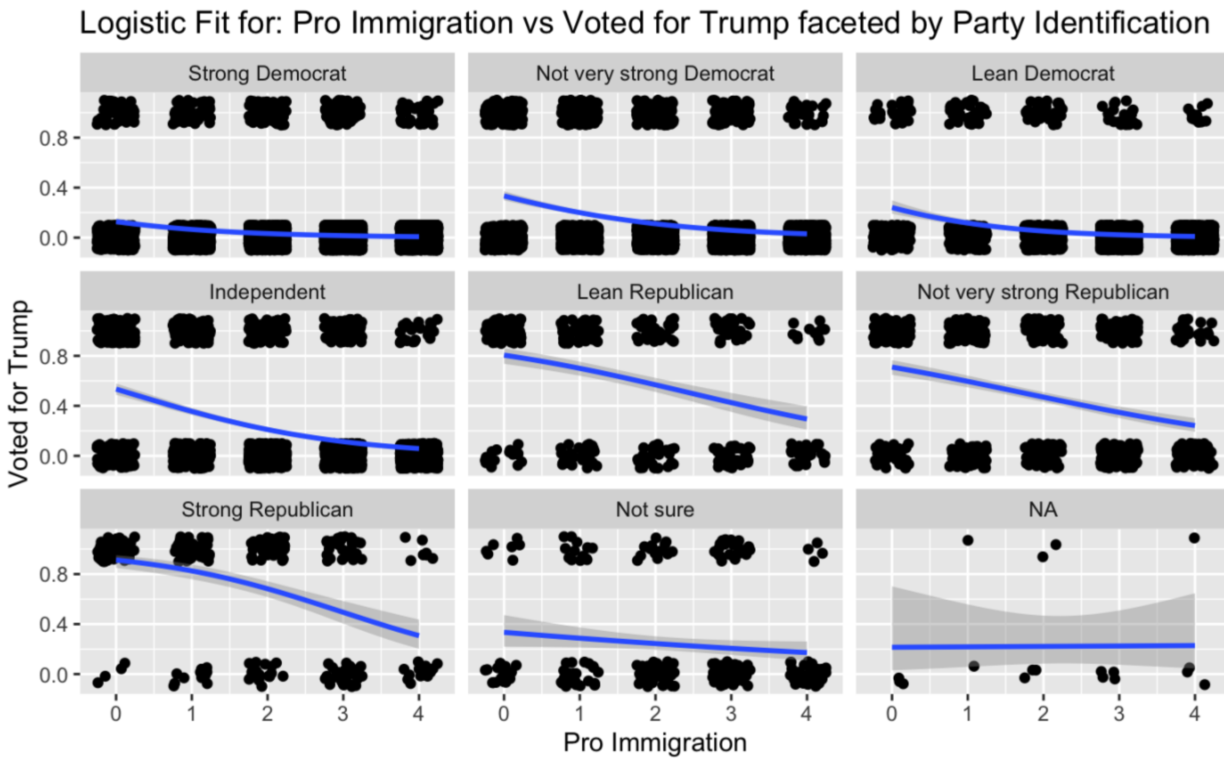
Appendix:

Analysis of the Correlation Plot:



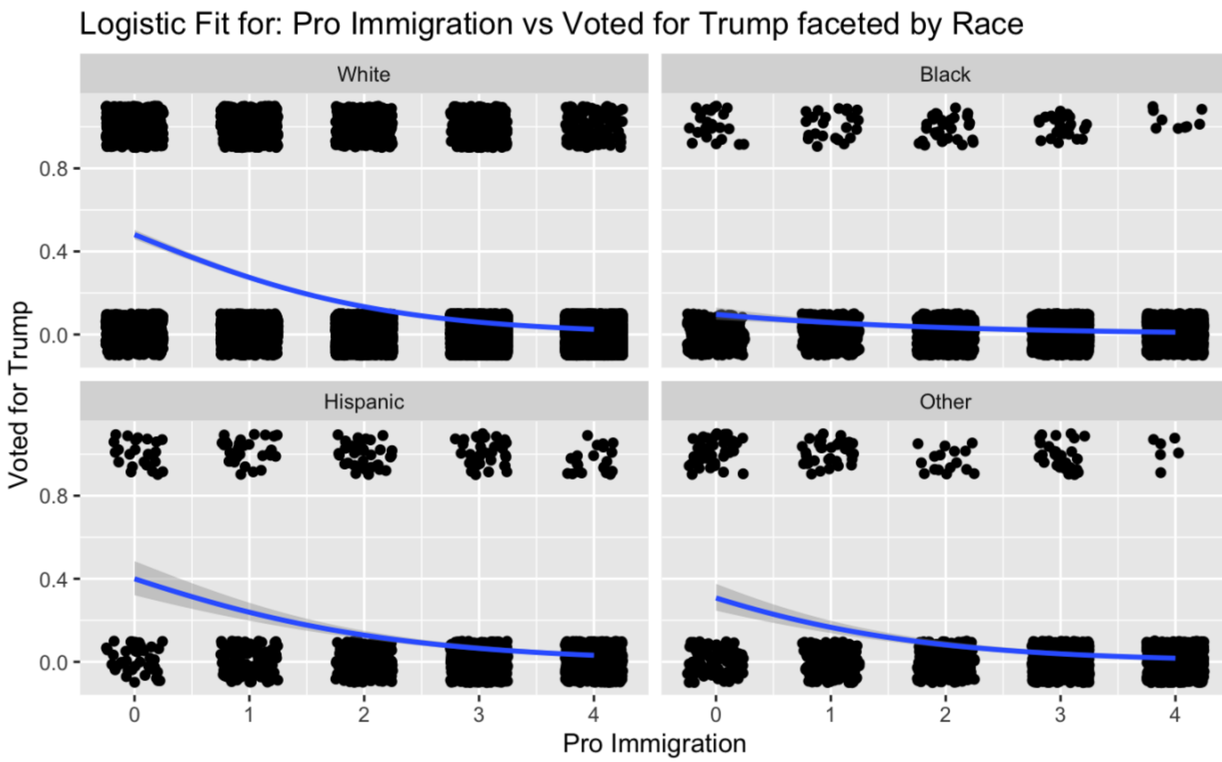
If we consider the first row, we can see the correlation of all the variables with our target variable isTrump(Probability of switching to Trump from Obama). Pro-Immigration i.e. pro.img has a good negative correlation with the target variable whereas education i.e. educ_num and Race i.e. race_num have a weak negative correlation with the target. Similarly, Party Identification i.e., pid7_num has a good positive correlation with the target variable.

Relationship between the Pro Immigration and Vote variables grouped by Party Identification:



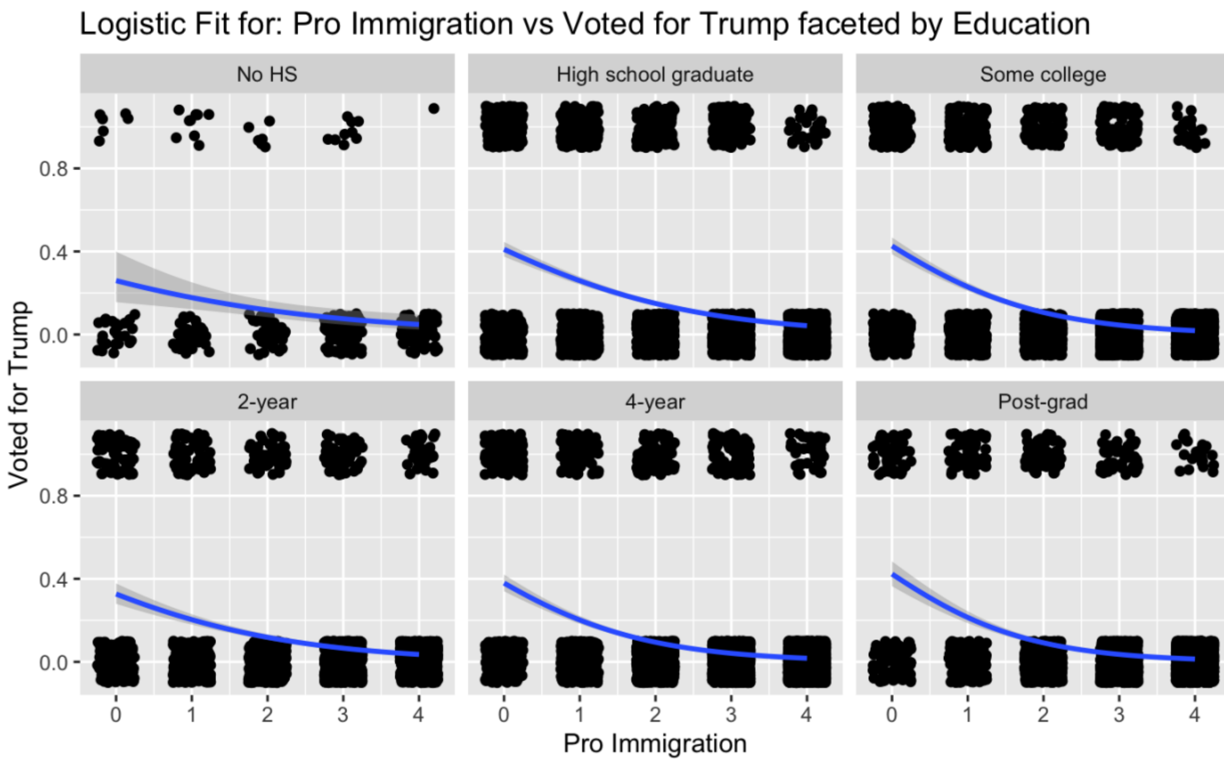
Here we can say that, for the party identifiers such as independent, lean republican, not very strong republican and strong republican the probability of voting for trump goes down as the person is more in support of immigration. The Pro Immigration value 0 suggests being against immigration and 4 suggests being in favor of immigration. In the party identifiers such as strong democrat, not very strong democrat, lean democrat and not sure we can't see much change with respect to Pro Immigration for the probability of switching to Trump. The main reason is because most of the categories belong to democrats and since Trump was a republican it does make sense.

Relationship between the Pro Immigration and Vote variables grouped by Race:



Here we can see that the Race groups White, Hispanic and Other have a good initial probability of switching to Trump i.e., when the Pro Immigration is 0, but as the Pro Immigration shifts towards 4, we can see a dip in the probability of switching to Trump in those groups. For the Race group Black, the probability of switching to Trump is very low irrespective of the Pro Immigration variable.

Relationship between the Pro Immigration and Vote variables grouped by Education:



Here we can see that all the education groups follow a general trend i.e., all of them have good initial probabilities of switching to Trump when the Pro Immigration is 0 but as the Pro Immigration shifts to 4 the switching probabilities go low.

Note: We see the same trend for the Gender group, for both Male and Female have good initial probabilities of switching to Trump when the Pro Immigration is 0 but as the Pro Immigration shifts to 4 the switching probabilities go low.