Voter Stance on Immigration as a Predictor of Voting for Donald Trump in 2016.

April 2018

Anthony Colombo, Vatsal Jatakia, Deidra Miniard

Indiana University

Introduction

The 2016 election was unique in three ways: first, Donald Trump was a new type of candidate who utilized populist and extreme rhetoric throughout his campaign (Oliver and Rahn, 2016); second, he won the presidential election even though most polls suggested that the Democratic candidate Hillary Clinton would win (Cohn, 2017a); finally, Donald Trump won many of the key swing states that Clinton was expected to win, largely due to voters who had supported President Barack Obama in 2012 and switched to vote for Donald Trump in 2016 (Cohn, 2017b). A primary reason explaining Donald Trump's success was his victories in key swing states, with voters who supported President Barack Obama in 2012 but switched to support Donald Trump in the 2016 election. An important question to consider is what factors may have caused those who voted for the Democratic candidate Obama in 2012 to vote for Republican Trump in 2016. Donald Trump was a vehement critic of the Obama presidency, both before and throughout Trump's campaign. Thus, what policies or self-identifiers may have caused some previous Barack Obama supporters to vote for Donald Trump in the 2016 election?

Immigration policy was rallying point Donald Trump utilized throughout his campaign, examples of which include his proposed border wall and call for a Muslim ban. Donald Trump differed from previous president Barack Obama on many issues including climate change, economy, trade, foreign policy, and immigration, healthcare, and so on. Of these policy domains, immigration is interesting and valuable to examine for two reasons. First, according to PEW Research Center (2016), polling showed that immigration was important to voter decision making; seventy percent of Americans indicated that immigration was "very important" in determining their vote. Of Trump supporters surveyed, 79% indicated that immigration was an important factor in determining their vote. In addition, when looking at the issues that Trump supporters listed as "very important" to their vote, the top four issues were the economy, foreign policy, terrorism, and immigration. Finally, immigration was a highly salient issue during the 2016 campaign. Thus, immigration is important to examine because issues that are highly salient and important to voters are likely to be the issues on which people decide which party and candidate to vote for (Belanger and Meguide, 2008; Krosnick, 1988).

While President Obama deported high numbers of illegal immigrants, he was also responsible for DACA, or Deferred Action for Childhood Arrivals which allowed undocumented children who have graduated from high school in the U.S. to remain in the country under certain stipulations. In contrast, President Trump referred to Obama as weak on immigration throughout his 2016 campaign, and his campaign included anti-immigration rhetoric. This brings us to our two hypotheses. First, individuals who support path-to-citizenship policies are less likely to switch from voting for Democrat Barack Obama in 2012 and vote for Republican Donald Trump in 2016. Second, individuals who support deportation-based policies are more likely to switch from voting for Obama and vote for Trump. Regression analysis is the best analytical tool because explaining vote choice of individuals is highly complex Many variables determine the candidate a person will vote for, and it is important to control for these variables.

The Model

The data for the model comes from the Cooperative Congressional Election Study. The study is supported by a National Science Foundation grant, and the principle investigators are individuals at Harvard University, the University of Massachusetts Amherst, and YouGov. For the research, sixty separate teams were tasked with collecting a nationally representative sample of approximately 1,000 respondents using matched random sampling. Matched random sampling is a method used to collect opt-in survey respondents that match the demographics of a random sample. The data from the CCES consisted of five parts: sample identifiers to provide locations of respondents; profile questions which reported demographic characteristics; pre-election questions collected before the presidential election; post-election questions collected immediately after the presidential election; and finally, contextual data such as congressional representatives of the respondent.

There were two anticipated issues with the data set. The first was the presence of missing values. Second, while the data set included over 500 variables some questions were only asked to a subset of the sample. Thus, some variables that could otherwise be included as controls could not be added because it would drastically reduce the sample size of our study. Similarly, since our analyzed subset consists of only people who voted for Obama in 2012 we can expect a disproportionate number of self-identified Democrats relative to Republicans; however, this should not bias the results because the low number of republicans in the sample is likely to match the actual percentage of Republicans in the population that voted for Obama in 2012.

We use a logit regression model because the dependent variable is dichotomous, meaning the dependent variable is separated into two categories where a person either voted for Trump or they voted for any other candidate. Our base model includes variables of interest on Immigration, as well as several control variables selected based on previous models of vote choice predictors, factors related to immigration attitudes, and policy positions of Trump voters. We also investigated an interaction model that examined the effect political identity has on the difference in likelihood of switching depending on gender. To allow for a visual representation of this interaction, political party was coded as an ordinal scale from strongly Democrat to strongly Republican. All control variables from the base model are included in the interaction model.

Four variables were included as proxy measures of voter stance on immigration. The first two variables measured support for path-to-citizenship policies. The first variable was a dichotomous variable assessing support for granting legal status to children who were brought to the U.S. illegally as children but have graduated from high school. The second path-to-citizenship variable was also dichotomous, measuring support for granting legal status to illegal immigrants who have worked in the U.S. for at least three years and paid taxes without any criminal convictions. In contrast, two variables were used to assess support for deportation-based policies. The first was a dichotomous measure of whether an individual supported identifying and deporting illegal immigrants, and the second was a dichotomous measure of support for increasing border patrol.

Control variables were selected based on previous models of vote choice predictors, factors related to immigration attitudes, and aggregate policy positions of Trump voters. Immigration attitudes have been shown to have a relationship to employment status (Boomgaarden & Vliegenthart, 2007) so we included a dichotomous variable for unemployment as a control. In addition, regression models have been used to evaluate how media on immigration relates to the likelihood an individual voted for an anti-immigrant candidate (Boomgaarden & Vliegenthart, 2007; Burscher & van Spanje, 2015). To control for media, we included an indexed variable which is a measure of how many sources of media an observation uses (i.e. television, social media, newspapers, etc.). Religion has been linked to immigration policy attitudes as well (Knoll, 2009) so a variable was included to control for religion This variable was an ordinal measure that gauged how important religion was for respondents, ranging from "not at all important" to "very important. In addition, we included education, family income, gender, political identity, and race as demographic control variables. Finally, we controlled for additional policies that might be important to Trump voters and thus might affect the likelihood of voters switching. Pew reported that the issues most Trump voters cared about were economy, terrorism, foreign policy, and immigration. To control for economy, we use a categorical measure of voter stance on economic conditions. In addition, we use a dichotomous variable to assess stance on terror that asks participants whether they support military action to eliminate terrorist camps. Voter stance on foreign policy could not be controlled for because the CCES data did not provide a good proxy.

Results

Variable	N	Percent
Gender		
Female	8,326	55.6%
Male	6,656	44.4%
Race		
White	10,240	68.3%
Black	2603	17.4%
Hispanic	1200	8.0%
Other	939	6.3%
Political Identity		
Strong Republican	178	1.2%
Lean Republican	239	1.6%
Not very strong Republican	466	3.1%
Independent	1310	8.7%
Not very strong Democrat	2955	19.7%
Lean Democrat	2213	14.8%
Strong Democrat	7621	50.9%
Education		
No high school	2,604	17.4%
Graduated High School	2,828	18.9%
Some college	3,489	23.2%
Two-year degree	1,784	11.9%
Four-year degree	4,108	27.4%
Post-grad	169	1.1%

Table 1. Distribution of Demographic Characteristics

The distributions of demographic characteristics are presented in Table 1. The sample had a greater percentage of females than males. Additionally, most of the sample consisted of Democrats, which is expected because the sample is a selection of those who voted for a democratic candidate in 2012. Summary information on additional variables in the model can be found in the technical appendix. To evaluate for multicollinearity between variables in our model, we ran a correlation matrix to examine correlations between our independent variables. All correlations between our independent variables were below 0.5, so we conclude that multicollinearity is not an issue in our model and no corrective action was taken. A correlation table can be found in the technical appendix.

Results of the logit models are presented in Table 2. Statistical significance is indicated in the table. Referent categories are Independent, white, some college, and believe the economy has stayed about the same. Coefficients, signs, and significance values were consistent across the base model and interaction model. The only variable that differed was gender. Additionally, political identity will be interpreted as a linear relationship for the interaction model.

Table 2. Regression Results for Base Model and Interaction model

	Dependent Variable: Vote for Trump	
	Base Logit	Interaction
Immigration		
Grant Legal Status: workers	-0.340*** (0.077)	-0.331*** (0.077)
Grant Legal Status: children	-0.563*** (0.075)	-0.563*** (0.075)
Deport Illegal Immigrants	0.860*** (0.077)	0.864*** (0.077)
Increase Border Patrols	0.554*** (0.073)	0.551*** (0.073)
Political Identity		
Strong Republican	1.848*** (0.216)	
Lean Republican	1.323*** (0.173)	
Not very strong Republican	1.098*** (0.132)	
Not very strong Democrat	-0.743*** (0.098)	
Lean Democrat	-1.311*** (0.128)	
Strong Democrat	-1.926*** (0.108)	
Race		
Black	-1.373*** (0.132)	-1.369*** (0.132)
Hispanic	-0.224*** (0.126)	-0.241*** (0.126)
Other	-0.619*** (0.147)	-0.622*** (0.146)
Gender: Male	-.413*** (0.072)	0.235*** (0.081)
Family Income	0.025*** (0.012)	0.028*** (0.012)
Media Index	0.109*** (0.032)	0.107*** (0.032)
Terrorism	0.155*** (0.074)	0.153*** (0.074)
Religion	0.307*** (0.033)	0.310*** (0.033)
Unemployed	0.071 (0.169)	0.064 (0.170)
Political Identity Recode		0.784*** (0.033)
Gender: Male*Political Identity		-0.207*** (0.045)

Individuals who supported granting legal status to children and granting legal status to workers were significantly less likely to switch and vote for Trump This is indicated by the negative coefficient in Table 2. Conversely, Individuals who supported finding and deporting illegal immigrants and increasing border patrols were significantly more likely to switch and vote for Donald Trump, as indicated by the positive coefficient. Political identity variables are as expected; individuals who indicate they are Democrats are less likely to switch and vote for Trump as compared to Independents, while individuals who indicated they were Republicans were more likely to switch as compared to Independents.

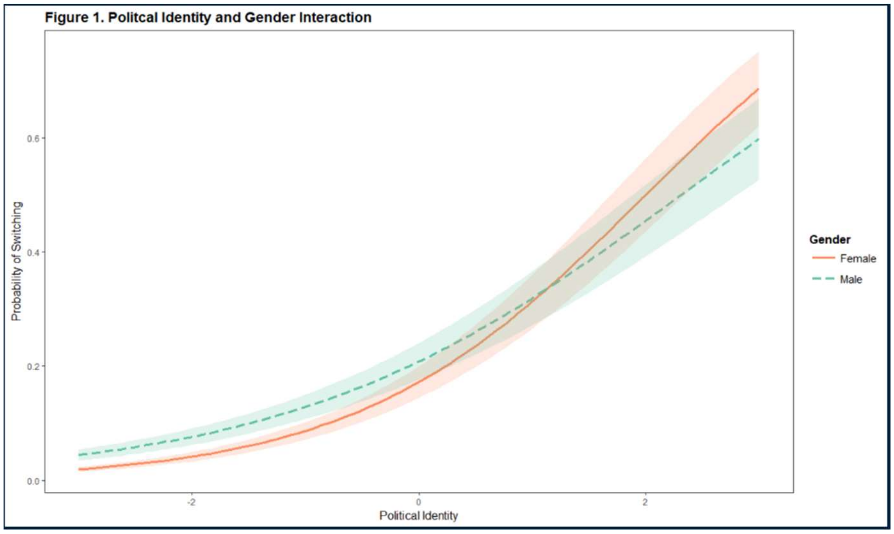
In the base model males were more likely than females to switch and vote for Trump. As

Education		
No High School	0.051	0.082
	(0.308)	(0.307)
High school graduate	0.294***	0.291***
	(0.098)	(0.098)
Two-year degree	0.097	0.101
	(0.120)	(0.120)
Four-year degree	-0.051	-0.066
	(0.104)	(0.104)
Post-grad	-0.182	-0.184
	(0.127)	(0.126)
Economy		
Gotten much worse	1.427***	1.420***
	(0.148)	(0.149)
Gotten worse	0.871***	0.863***
	(0.090)	(0.091)
Gotten Better	-0.775***	-0.766***
	(0.091)	(0.090)
Gotten much better	0.052	0.045
	(0.143)	(0.140)
Constant	-2,561***	-2.443***
	(0.178)	(0.169)
Observations	14,970	14,970
Akaike Inf. Crit.	6118.5	6106.799
Note: *p<0.1,	**p<0.05,	***p<0.01

the number of media sources increased, the likelihood of switching also increased. In addition, Individuals with a high school degree and no college were more likely to switch. Also, as family income increased, the likelihood of voting for Trump also increased and as individuals reported religion was more important, the likelihood they switched and voted for Trump also increased. Similarly, those that supported the military action against terrorism were more likely to switch. Finally, respondents who reported the economy had gotten worse or much worse were more likely to switch as compared to those who believed the economy had stayed about the same.

Some control variables indicated that individuals would be less likely to switch and vote for Trump. Black, Hispanic, and individuals who indicated “other” for race were less likely than white voters to switch. Additionally, individuals who indicated that the economy had gotten better were less likely to switch as well. Finally, some variables were not statistically significant, including some education variables, and the control for unemployment.

A significant interaction was present between gender and political identity. The relationship is presented in Figure 1. For voters with Democratic political identities males who voted for Obama in 2012 were slightly more likely to vote for Trump in 2016 in comparison to females with self-reported Democratic political identities. This relationship appears to change as voters move into



the Republican end of the political identity spectrum. The graph indicates that female voters who voted for Obama in 2012 and are Republican were more likely to vote for Trump in 2016 than male Republicans. However, the sample size of observations for Republicans is small so the confidence intervals widen, and we cannot be certain of the relationships between gender and political identity on the Republican end of the spectrum. The change in odds indicates that those in favor of granting legal status to immigrants who have worked full time for three years, payed taxes, and have no criminal record have more than a 28% decrease in the odds of voting for Trump in comparison to those who are against this policy. The positive percentage in the table indicate that those

Table 3. Change in Odds for Base Model

	Dependent Variable: Vote for Trump
Immigration	
Grant Legal Status: workers	-28.80%
Grant Legal Status: children	-43.07%
Deport Illegal Immigrants	136.30%
Increase Border Patrols	73.99%
Political Identity	
Strong Republican	534.86%
Lean Republican	275.47%
Not very strong Republican	199.89%
Not very strong Democrat	-52.43%
Lean Democrat	-73.04%
Strong Democrat	-85.43%
Race	
Black	-74.66%
Hispanic	-20.04%
Other	-46.16%

Unemployed	7.32%
Economy	
Gotten much worse	316.72%
Gotten worse	139.02%
Gotten Better	-53.95%
Gender: Male	51.09%
Media Index	11.55%
Terrorism	16.76%
Religion	35.94%

who support the deportation of illegal immigrants and increased border patrol have an increase in odds of voting for Trump.

Unsurprisingly voters with Democrat political identities show a decrease in odds while Republicans show an increase in reference to Independents. Additionally, the output for race indicates that all other categories show a decrease in odds of switching in 2016 in comparison to white voters. Furthermore, voters who indicated they felt the economy has gotten worse show a significant increase in the odds of switching.

Conclusion

Our hypotheses were supported by the logistical regression. Individuals who supported the path-to-citizenship based policies were less likely to vote for Trump while individuals who supported the deportation-based policies were more likely to vote for Trump. In addition, control variables in the model were also found to have a relationship to whether an individual switched to voting for Trump, suggesting that several variables beyond salient immigration policies were important factors in determining vote choice.

The current analysis can be extended in four ways. First, additional interactions can be explored, such as interactions between political identity and race. Second, further investigation can be conducted into variables with interesting results in our model. For example, our model showed that as the number of media sources increased (i.e. tv, social media, newspaper), the probability for switching increased. A next step in analysis should consider why media sources had a meaningful impact, and how this might relate to how many different channels or outlets an individual consumes. Additional analysis might also consider how this media variable interacts with political identity and immigration attitudes.

Another extension of the research might consider adding more variables. The presented study includes policy areas that most Trump supporters indicate are important to them. A future model extension can include variables that were indicated as important, but less so than immigration, such as healthcare or gun rights. Similarly, because our sample was based on previous Obama voters, variables that are important to Democrats and Independents should be included as well, such as education or the environment.

Finally, an interesting and useful extension might consider how individuals who previously voted for Obama in 2012 and switched to voting for Donald Trump in 2016 differ from those voters who voted for the Republican candidate in 2012 and stayed Republican in the 2016 election. This comparison would be interesting because the issues that caused individuals to switch and vote for Trump may be different from the issues that caused voters to stay within the Republican party and vote for Trump.

The results from our study are important in three major ways. First, we show that immigration, a salient issue during the 2016 Trump campaign, was an important factor that contributed to previous Obama supporters switching to vote for Donald Trump. Second, we show that while immigration is an important factor, other factors were important as well, showing that vote choice behavior is highly complex and dependent upon many different issues. Lastly, recent work has shown that while the general consensus was that economic issues drove individuals to vote for Trump, it is more likely that individuals voted for Trump because they felt that their status and dominance in the country were threatened (Cox, Lienesch, & Jones, 2017).

Our study further supports this research, showing that support for more strict immigration policies is associated with an increased likelihood of voting for Donald Trump in 2016 for those who previously voted for Barack Obama in 2012. Perhaps when evaluating votes that switch across party lines, an important consideration is not only party identity, but also the opinions individuals have on salient issues like immigration, that stem from deeply held fears and beliefs.

References

Belanger, E. and Meguid, B.M. (2008). Issue salience, issue ownership, and issue-based vote choice. *Electoral Studies, 27,* 477-491.

Burscher, B., van Spanje, J., and de Vreese, C.H. (2015). Owning the issues of crime and immigration: The relation between immigration and crime news and anti-immigrant voting in 11 countries. *Electoral Studies, 38*, 59-69.

Cohn, N. (2017a). A 2016 review: Why key state polls were wrong about Trump. *The New York Times.* https://www.nytimes.com/2017/05/31/upshot/a-2016-review-why-key-state-polls-were-wrong-about-trump.html.

Cohn, N. (2017b). The Obama-Trump voters are real. Here's what they think. *The New York Times.* https://www.nytimes.com/2017/08/15/upshot/the-obama-trump-voters-are-real-heres-what-they-think.html

Cox, D., Lienesch, R., and Jones, R.P. (2017). Beyond Economics: Fears of cultural displacement pushed the white working class to Trump. *PRRI/The Atlantic Report.*

Knoll, B.R. (2009). "And who is my neighbor?" Religion and Immigration Policy Attitudes. *Journal for the Scientific Study of Religion, 48*(2), 313-331.

Krosnick, J.A. (1988). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. *Journal of Personality and Social Psychology, 55*(2), 196-210.

Oliver, J.E. and Rahn, W.M. (2016). Rise of the Trumpenvolk: Populism in the 2016 election. *The ANNALS of the American Academy of Political and Social Science, 667,* 189 – 206.

Pew Research Center. U.S. Politics and Policy. (2016). 2016 campaign: Strong interest, widespread Dissatisfaction. Section 4. http://assets.pewresearch.org/wp-content/uploads/sites/5/2016/07/07-07-16-Voter-attitudes-release.pdf.

Technical Appendix

Summary of Variables included in the Model

Frequency Table for Independent Variable of Interest

Variable	N	Percent
Grant Legal Status: workers		
Support	11,130	74.3%
No Support	3,852	25.7%
Grant Legal Status: children		
Support	9,679	64.6%
No Support	5,303	35.4%
Increase Border Patrols		
Support	5,671	37.8%
No Support	9,311	62.2%
Deport Illegal Immigrants		
Support	3,598	24%
No Support	11,384	76%

Frequency Table for Dependent Variable

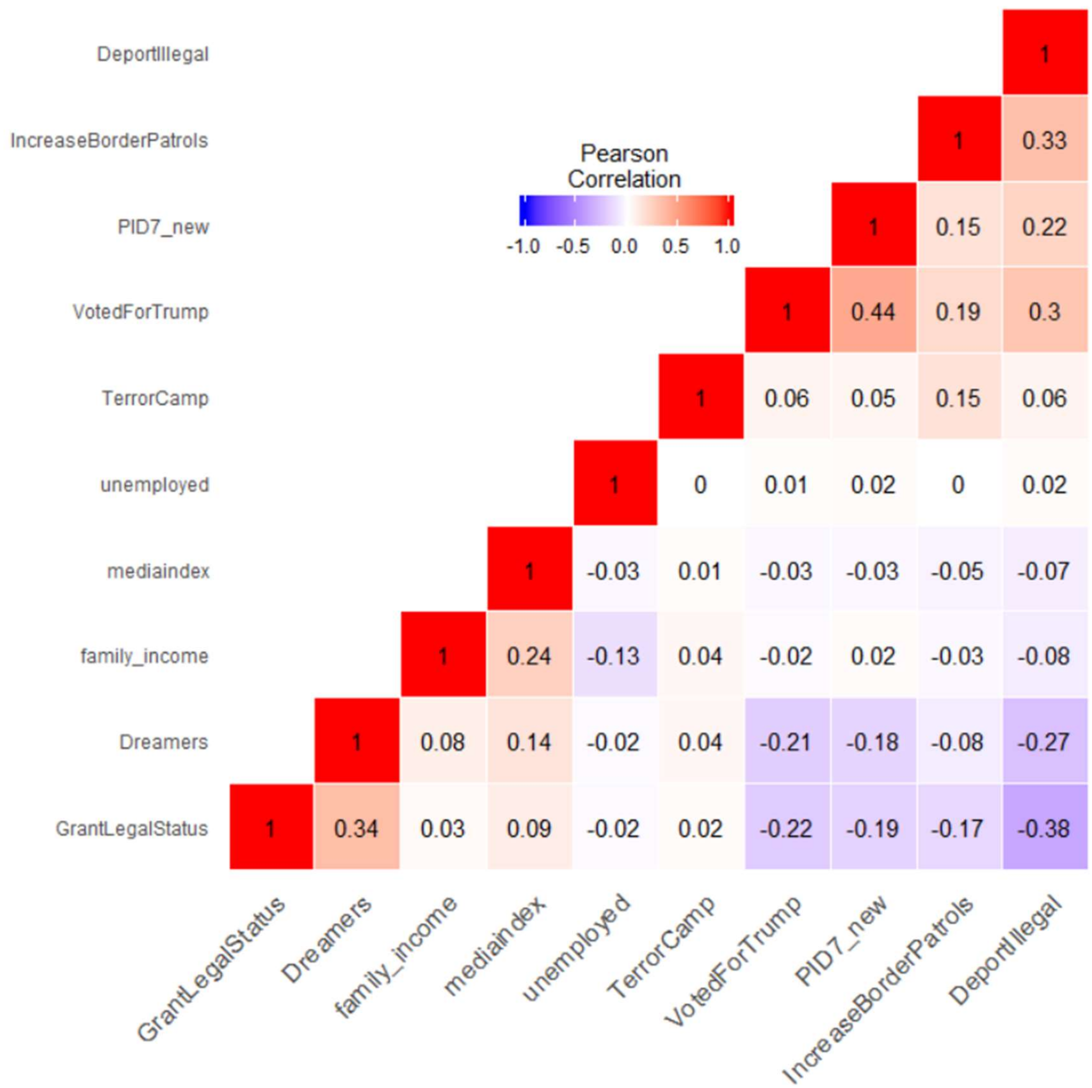
Variable	N	Percent
2016 Vote		
Trump	1,546	10.3%
Other	13,436	89.7%

Frequency Table for Independent Variables in the model

Variable	N
Unemployed	
Unemployed	580
Other	14,402
Family Income	
0	10,240
Black	2603
Hispanic	1200
Other	939
Media Index	

1	1561
2	4230
3	4763
4	3089
5	1339
Religion	
0: Not at all important	3901
1: Not too important	2567
2: Somewhat important	3903
3: Very important	4599
Education	
No high school	2,604
Graduated High School	2,828
Some college	3,489
Two-year degree	1,784
Four-year degree	4,108
Post-grad	169
Economy	
Gotten Much Worse	384
Gotten Worse	1468
Stayed the same	4640
Gotten Better	7112
Gotten Much Better	1378
Terrorism	
Support	9315
No indication of Support	5667
Family Income	
0: Less than $10,000	369
1: $10,000 - $19,999	952
2: $20,000 - $29,999	1489
3: $30,000 - $39,999	1589
4: $40,000 - $49,999	1431
5: $50,000 - $59,999	1501
6: $60,000 - $69,999	1272
7: $70,000 - $79,999	1317

8: $80,000 - $99,999	1643
9: $100,000 - $119,999	1183
10: $120,000 - $149,999	1106
11: $150,000 - $199,999	608
12: $200,000 - $249,999	274
13: $250,000 - $349,999	135
14: $350,000 - $499,999	71
15: $500,000 or more	42



Initial Cleaning

The original data has 563 variables and 64,600 rows. Our sample includes only those who voted for Obama in 2012 and then filtered the candidates who took this survey both before and after the 2016 elections. We also filtered out the candidates who did not answer who they voted for during the 2016 election. We created our response variable(VotedForTrump) as a binary which indicates 1 if the person voted for Trump and 0 otherwise.

Immigration Variables: These are the predictors for the response variable and represent the stance of a person on immigration. These are also binary variables depicting whether the person supports the stance or not.

1. Granting legal status to a person who came to the US as an illegal immigrant but has worked and paid taxes for 3 years and not been convicted for any felonies (1 = Agree, 0 = Disagree).

2. Granting legal status to immigrants who were brought to the United States illegally as children but have graduated from a U.S. High School(1 = Agree, 0 = Disagree).

3. Increasing the Border Patrols on the U.S-Mexico Border(1 = Agree, 0 = Disagree).

4. Identifying and Deporting the Illegal immigrants(1 = Agree, 0 = Disagree).

Control variables:

After doing an extensive research survey, we selected the following control variables:

1. Political Identification: (Ranging from -3 which corresponds to a Strong Democratic to +3 which indicates a Strong Republican)

2. Race : This indicates the race of the respondent and due to the bias in the data, we have recoded the Race variable to be “White”,“Black”, “Hispanic” or “Other” as the categories.

3. Education: The Education variable is a factor ranging from No High School to Post-Grad.

4. Media Index: Sum of the no of media a person uses including newspaper, social media, radio, and/or TV

5. Family Income: Ordinal variable suggesting the range of family income of a respondent (0-“Less than $10000” to 15- “$500,000 or more”)

6. Employment Status of the respondent. (1 = Unemployed, 0 = otherwise)

7. Economy: Whether or not the economy has improved since Obama? (-2: Gotten Way Worse to 2: Gotten Way Better)

8. Religion: Whether or not Religion is an important factor (0: Not important to 3: Very important)

9. TerrorCamp: Whether or not American soldiers must be used to fight against terrorists in other countries (0 =No, 1 = yes)

```r
setwd("C:/Users/Vatsal Jatakia/OneDrive/Study Material/Data Analysis and Modeling/Project")
load("CCES16_Common_OUTPUT_Feb2018_VV.RData")

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.4.2

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

#Selecting the above mentioned variables and filtering the NA's while selecting
obama = x[x$CC16_326 == 'Barack Obama' & is.na(x$CC16_326)== FALSE & x$tookpost == 'Yes' ,c('commonweight_vv_post', 'tookpost', 'gender', 'educ','race', 'pid7', 'CC16_326', 'CC16_410a', 'CC16_331_1','CC16_331_2', 'CC16_331_3', 'CC16_331_7')]

cols = c("SurveyWeight", "TookPost", "Gender", "Education", "Race", "PID7", "Vote_2012", "Vote_2016","GrantLegalStatus", "IncreaseBorderPatrols", "Dreamers", "DeportIllegal")

names(obama) = cols

#Creating a VotedForTrump variable to get the DV (1 = Voted for Trump, 0 = Didn't vote for Trump)
obama$VotedForTrump = ifelse(obama$Vote_2016 == 'Donald Trump (Republican)', 1, 0)

#Recoding the race variable using dplyr
obama$Race = recode(obama$Race, "White"="White", "Black"="Black", "Hispanic"="Hispanic", .default = "Other" )

#Recoding the Immigration Attitude variables
obama$GrantLegalStatus = as.numeric(as.character(recode(obama$GrantLegalStatus, "Yes"= 1, "No"= 0)))
obama$IncreaseBorderPatrols = as.numeric(as.character(recode(obama$IncreaseBorderPatrols, "Yes"= 1, "No"= 0)))
obama$Dreamers = as.numeric(as.character(recode(obama$Dreamers, "Yes"= 1, "No"= 0)))
```

```
obama$DeportIllegal = as.numeric(as.character(recode(obama$DeportIllegal,
"Yes"= 1, "No"= 0)))

#Creating the new immigration Attitude score variable
obama$ImmigrationAttitudeScore =
obama$DeportIllegal+obama$IncreaseBorderPatrols+obama$Dreamers+obama$GrantLeg
alStatus

#Removing the missing values from the variables
obama = subset(obama, !is.na(obama$Vote_2016))
obama = subset(obama, PID7!= "Not sure")

write.csv(obama, "Obama_Trump1.csv") #Saving the file is important to get the
row numbers for adding the new variables. Do save before running!
```

Adding new variables like media sources, family income and employment status and recoding them.

```
obama = read.csv("Obama_Trump1.csv")

#Selecting the media sources variable
Obama_Trump = data.frame(obama, loc_nat = x$CC16_300b[obama$X], newspaper =
x$CC16_300_3[obama$X], blog = x$CC16_300_1[obama$X],tv =
x$CC16_300_2[obama$X],radio = x$CC16_300_4[obama$X], social_media =
x$CC16_300_5[obama$X], none = x$CC16_300_6[obama$X] )

#Remove missing values
Obama_Trump = subset(Obama_Trump, !is.na(loc_nat))

#Recoding them to numeric form for getting index
Obama_Trump$newspaper<-recode(Obama_Trump$newspaper,'Yes'=1 , 'No'=0)
Obama_Trump$blog<-recode(Obama_Trump$blog,'Yes'=1,'No'=0)
Obama_Trump$tv<-recode(Obama_Trump$tv,'Yes'=1,'No'=0)
Obama_Trump$radio<-recode(Obama_Trump$radio,'Yes'=1,'No'=0)
Obama_Trump$social_media<-recode(Obama_Trump$social_media,'Yes'=1,'No'=0)

#Converting media variables to numeric form
Obama_Trump$tv<-as.numeric(as.character(Obama_Trump$tv))
Obama_Trump$newspaper<-as.numeric(as.character(Obama_Trump$newspaper))
Obama_Trump$blog<-as.numeric(as.character(Obama_Trump$blog))
Obama_Trump$social_media<-as.numeric(as.character(Obama_Trump$social_media))
Obama_Trump$radio<-as.numeric(as.character(Obama_Trump$radio))

#Creating the media index variable
media_index = rowSums(Obama_Trump[17:21])
Obama_Trump$mediaindex<-media_index

#Adding the family income and employment status as variables
Obama_Trump$family_income = x$faminc[Obama_Trump$X]
```

```
Obama_Trump$employ_status = x$employ[Obama_Trump$X]

library(car)

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##     recode

#Creating the unemployed status variable as a binary variable
Obama_Trump$unemployed<-
recode(Obama_Trump$employ_status,"c('Unemployed','Temporarily laid
off')=1;c('Full-time','Homemaker','Other','Part-time','Permanently
disabled','Retired','Student')=0")

#Converting teh family income to character form
Obama_Trump$family_income = as.character(Obama_Trump$family_income)

#Remove rows with family income = Prefer not to say
Obama2 = subset(Obama_Trump, Obama_Trump$family_income!= "Prefer not to say")

#Remove rows with family income = $150000 or more
Obama3 = subset(Obama2, Obama2$family_income != "$150,000 or more")



#Recoding the family income to make it numeric
Obama3$family_income[which(Obama3$family_income == "Less than $10,000" )] = 0
Obama3$family_income[which(Obama3$family_income == "$10,000 - $19,999" )] = 1
Obama3$family_income[which(Obama3$family_income == "$20,000 - $29,999" )] = 2
Obama3$family_income[which(Obama3$family_income == "$30,000 - $39,999" )] = 3
Obama3$family_income[which(Obama3$family_income == "$40,000 - $49,999" )] = 4
Obama3$family_income[which(Obama3$family_income == "$50,000 - $59,999" )] = 5
Obama3$family_income[which(Obama3$family_income == "$60,000 - $69,999" )] = 6
Obama3$family_income[which(Obama3$family_income == "$70,000 - $79,999" )] = 7
Obama3$family_income[which(Obama3$family_income == "$80,000 - $99,999" )] = 8
Obama3$family_income[which(Obama3$family_income ==  "$100,000 - $119,999" )]
= 9
Obama3$family_income[which(Obama3$family_income ==  "$120,000 - $149,999" )]
= 10
Obama3$family_income[which(Obama3$family_income ==  "$150,000 - $199,999" )]
= 11
Obama3$family_income[which(Obama3$family_income ==  "$200,000 - $249,999" )]
= 12
Obama3$family_income[which(Obama3$family_income ==  "$250,000 - $349,999" )]
= 13
Obama3$family_income[which(Obama3$family_income ==  "$350,000 - $499,999" )]
= 14
Obama3$family_income[which(Obama3$family_income ==  "$500,000 or more" )] =
```

```
15


#Saving the file as a CSV
write.csv(Obama3,'Obama_Trump2.csv') #This saving is optional. But do it just
incase you decide to play around with more variables.
```

Adding the Religion, Foreign Jobs, TerrorCamp and Economy

```
Obama_Trump = read.csv("Obama_Trump2.csv")


### ADDING THE VARIABLES
Obama_Trump$Religion = x$pew_religimp[Obama_Trump$X]
Obama_Trump$Economy = x$CC16_302[Obama_Trump$X]
Obama_Trump$TerrorCamp = x$CC16_414_2[Obama_Trump$X]

#Removing the NA's
Obama_Trump = subset(Obama_Trump, Economy!= "Not sure")
Obama_Trump$Economy = factor(Obama_Trump$Economy, ordered = F)

## Recoding the Terror Camp values
Obama_Trump$TerrorCamp = ifelse(Obama_Trump$TerrorCamp == "Yes", 1,0)

#Recoding the Religion
Obama_Trump$Religion = as.character(Obama_Trump$Religion)
Obama_Trump$Religion[which(Obama_Trump$Religion == "Not at all important")] =
0
Obama_Trump$Religion[which(Obama_Trump$Religion == "Not too important")] = 1
Obama_Trump$Religion[which(Obama_Trump$Religion == "Somewhat important")] = 2
Obama_Trump$Religion[which(Obama_Trump$Religion == "Very important")] = 3
Obama_Trump$Religion = as.numeric(Obama_Trump$Religion)

#Recoding PID7

Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Strong Democrat")] = -3
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Strong Republican")] = 3
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Lean Democrat")] = -2
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Lean Republican")] = 2
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Not very strong Republican")]
= 1
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Not very strong Democrat")] =
-1
Obama_Trump$PID7_new[which(Obama_Trump$PID7 == "Independent")] = 0

write.csv(Obama_Trump, "Obama_Trump_FINALLY_FINAL.csv")
```

Identifying correlations among the variables in the base model

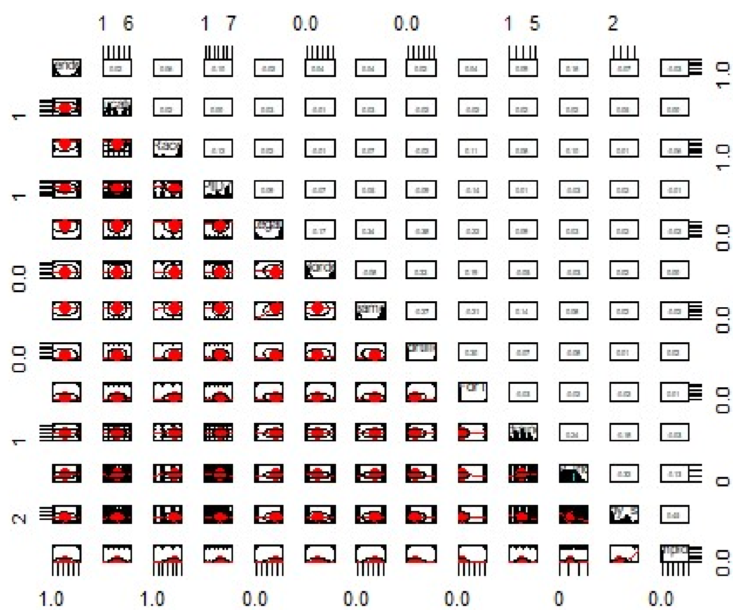
```
library(psych)

## Warning: package 'psych' was built under R version 3.4.3

##
## Attaching package: 'psych'

## The following object is masked from 'package:car':
##
##     logit

#Analyzing the correlation among the variables in the model
Obama_subset = Obama_Trump[names(Obama_Trump) %in% c("VotedForTrump", "PID7",
"Gender", "Education","Race","unemployed","employ_status",
"mediaindex","family_income","Dreamers","GrantLegalStatus","IncreaseBorderPat
rols","DeportIllegal")]
pairs.panels(Obama_subset)
```



Base Logit

```
Obama_Trump$Economy<-relevel(Obama_Trump$Economy,ref="Stayed about the same")
Obama_Trump$Race<-relevel(Obama_Trump$Race,ref="White")
Obama_Trump$Education<-relevel(Obama_Trump$Education,ref="Some college")
summary(baselogit<-glm(VotedForTrump~ as.factor(PID7) + GrantLegalStatus +
IncreaseBorderPatrols + Dreamers + DeportIllegal + Gender + as.factor(Race) +
as.factor(Education) + unemployed + family_income + mediaindex + TerrorCamp +
Religion + Economy, data=Obama_Trump,family=binomial(link="logit")))
```

```
##
## Call:
## glm(formula = VotedForTrump ~ as.factor(PID7) + GrantLegalStatus +
##     IncreaseBorderPatrols + Dreamers + DeportIllegal + Gender +
##     as.factor(Race) + as.factor(Education) + unemployed + family_income +
##     mediaindex + TerrorCamp + Religion + Economy, family = binomial(link =
"logit"),
##     data = Obama_Trump)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5001  -0.3254  -0.1875  -0.1168   3.3611
##
## Coefficients:
##                                             Estimate Std. Error z value
## (Intercept)                                 -2.56124    0.17757 -14.424
## as.factor(PID7)Lean Democrat                -1.31081    0.12789 -10.250
## as.factor(PID7)Lean Republican               1.32302    0.17315   7.641
## as.factor(PID7)Not very strong Democrat     -0.74287    0.09792  -7.587
## as.factor(PID7)Not very strong Republican    1.09825    0.13164   8.343
## as.factor(PID7)Strong Democrat              -1.92621    0.10794 -17.846
## as.factor(PID7)Strong Republican             1.84823    0.21559   8.573
## GrantLegalStatus                            -0.33968    0.07731  -4.394
## IncreaseBorderPatrols                        0.55385    0.07323   7.563
## Dreamers                                    -0.56336    0.07522  -7.489
## DeportIllegal                                0.85992    0.07745  11.103
## GenderMale                                   0.41268    0.07186   5.743
## as.factor(Race)Black                        -1.37295    0.13199 -10.402
## as.factor(Race)Hispanic                     -0.22367    0.12626  -1.772
## as.factor(Race)Other                        -0.61916    0.14663  -4.223
## as.factor(Education)2-year                   0.09746    0.11988   0.813
## as.factor(Education)4-year                  -0.05118    0.10426  -0.491
## as.factor(Education)High school graduate     0.29444    0.09838   2.993
## as.factor(Education)No HS                    0.05116    0.30784   0.166
## as.factor(Education)Post-grad               -0.18163    0.12715  -1.428
## unemployed                                   0.07062    0.16901   0.418
## family_income                                0.02468    0.01248   1.977
## mediaindex                                   0.10927    0.03236   3.376
## TerrorCamp                                   0.15496    0.07421   2.088
## Religion                                     0.30705    0.03317   9.258
## EconomyGotten much better                    0.05193    0.14252   0.364
## EconomyGotten better                        -0.77542    0.09070  -8.550
## EconomyGotten worse                          0.87138    0.09050   9.629
## EconomyGotten much worse                     1.42725    0.14845   9.614
##                                             Pr(>|z|)
## (Intercept)                                  < 2e-16 ***
## as.factor(PID7)Lean Democrat                 < 2e-16 ***
## as.factor(PID7)Lean Republican              2.16e-14 ***
## as.factor(PID7)Not very strong Democrat     3.29e-14 ***
## as.factor(PID7)Not very strong Republican    < 2e-16 ***
```

```
## as.factor(PID7)Strong Democrat             < 2e-16 ***
## as.factor(PID7)Strong Republican           < 2e-16 ***
## GrantLegalStatus                          1.11e-05 ***
## IncreaseBorderPatrols                     3.95e-14 ***
## Dreamers                                  6.92e-14 ***
## DeportIllegal                              < 2e-16 ***
## GenderMale                                9.30e-09 ***
## as.factor(Race)Black                       < 2e-16 ***
## as.factor(Race)Hispanic                   0.076459 .
## as.factor(Race)Other                      2.41e-05 ***
## as.factor(Education)2-year                0.416253
## as.factor(Education)4-year                0.623496
## as.factor(Education)High school graduate  0.002765 **
## as.factor(Education)No HS                 0.868014
## as.factor(Education)Post-grad             0.153179
## unemployed                                0.676062
## family_income                             0.048018 *
## mediaindex                                0.000735 ***
## TerrorCamp                                0.036792 *
## Religion                                   < 2e-16 ***
## EconomyGotten much better                 0.715589
## EconomyGotten better                       < 2e-16 ***
## EconomyGotten worse                        < 2e-16 ***
## EconomyGotten much worse                   < 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: 9933.6  on 14969  degrees of freedom
## Residual deviance: 6060.5  on 14941  degrees of freedom
##   (12 observations deleted due to missingness)
## AIC: 6118.5
##
## Number of Fisher Scoring iterations: 6
```

PID*Gender interactions

```
summary(PIDGenderlogit<-glm(VotedForTrump~ Gender*PID7_new +
as.factor(Education) + GrantLegalStatus + IncreaseBorderPatrols + Dreamers +
DeportIllegal + as.factor(Race) + unemployed + family_income + mediaindex +
TerrorCamp + Religion + Economy,
data=Obama_Trump,family=binomial(link="logit")))
```

```
##
## Call:
## glm(formula = VotedForTrump ~ Gender * PID7_new + as.factor(Education) +
##     GrantLegalStatus + IncreaseBorderPatrols + Dreamers + DeportIllegal +
##     as.factor(Race) + unemployed + family_income + mediaindex +
##     TerrorCamp + Religion + Economy, family = binomial(link = "logit"),
```

```
##     data = Obama_Trump)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.7312  -0.3283  -0.1859  -0.1095   3.3451
##
## Coefficients:
##                                          Estimate Std. Error z value
## (Intercept)                              -2.44337    0.16932 -14.431
## GenderMale                                0.23463    0.08081   2.904
## PID7_new                                  0.78432    0.03300  23.767
## as.factor(Education)2-year                0.10056    0.11977   0.840
## as.factor(Education)4-year               -0.06592    0.10403  -0.634
## as.factor(Education)High school graduate  0.29106    0.09838   2.959
## as.factor(Education)No HS                 0.08234    0.30740   0.268
## as.factor(Education)Post-grad            -0.18373    0.12637  -1.454
## GrantLegalStatus                         -0.33142    0.07711  -4.298
## IncreaseBorderPatrols                     0.55087    0.07305   7.541
## Dreamers                                 -0.56334    0.07493  -7.518
## DeportIllegal                             0.86376    0.07721  11.187
## as.factor(Race)Black                     -1.36916    0.13178 -10.390
## as.factor(Race)Hispanic                  -0.24059    0.12607  -1.908
## as.factor(Race)Other                     -0.62209    0.14584  -4.266
## unemployed                                0.06417    0.16991   0.378
## family_income                             0.02753    0.01243   2.214
## mediaindex                                0.10667    0.03226   3.307
## TerrorCamp                                0.15335    0.07411   2.069
## Religion                                  0.31045    0.03283   9.458
## EconomyGotten much better                 0.04460    0.14028   0.318
## EconomyGotten better                     -0.76575    0.09000  -8.509
## EconomyGotten worse                       0.86300    0.09058   9.527
## EconomyGotten much worse                  1.42030    0.14939   9.508
## GenderMale:PID7_new                      -0.20690    0.04483  -4.615
##                                          Pr(>|z|)
## (Intercept)                               < 2e-16 ***
## GenderMale                               0.003690 **
## PID7_new                                  < 2e-16 ***
## as.factor(Education)2-year               0.401100
## as.factor(Education)4-year               0.526310
## as.factor(Education)High school graduate 0.003090 **
## as.factor(Education)No HS                0.788799
## as.factor(Education)Post-grad            0.145954
## GrantLegalStatus                         1.72e-05 ***
## IncreaseBorderPatrols                    4.66e-14 ***
## Dreamers                                 5.56e-14 ***
## DeportIllegal                             < 2e-16 ***
## as.factor(Race)Black                      < 2e-16 ***
## as.factor(Race)Hispanic                  0.056339 .
## as.factor(Race)Other                     1.99e-05 ***
## unemployed                               0.705691
```

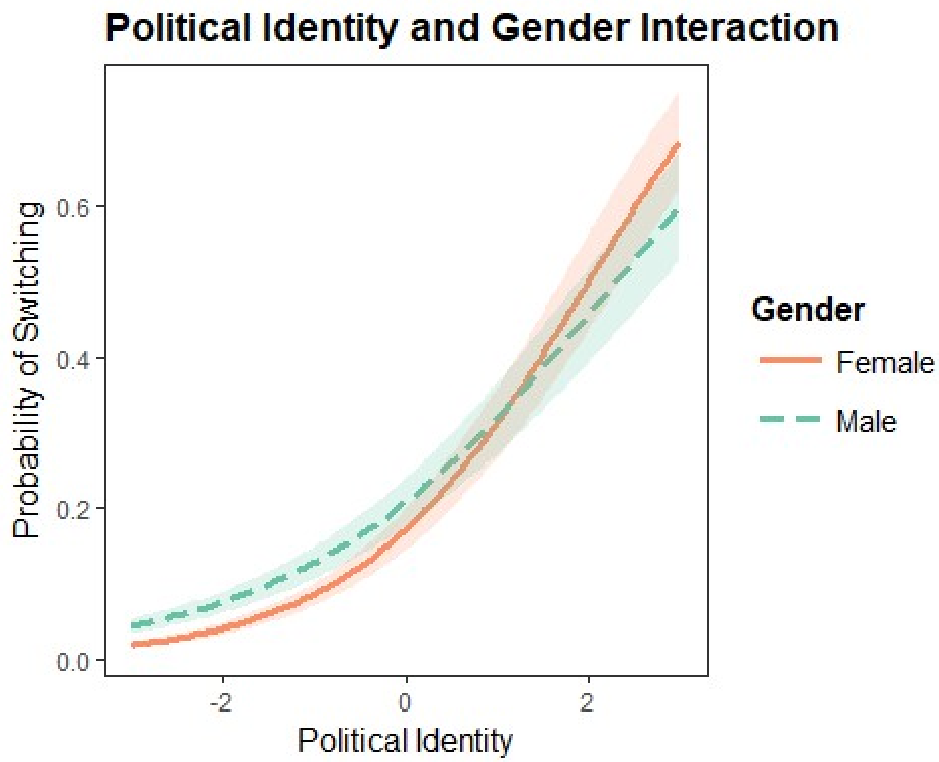
```
## family_income                                 0.026810 *
## mediaindex                                    0.000944 ***
## TerrorCamp                                    0.038523 *
## Religion                                       < 2e-16 ***
## EconomyGotten much better                     0.750525
## EconomyGotten better                           < 2e-16 ***
## EconomyGotten worse                            < 2e-16 ***
## EconomyGotten much worse                       < 2e-16 ***
## GenderMale:PID7_new                           3.94e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 9933.6  on 14969  degrees of freedom
## Residual deviance: 6056.8  on 14945  degrees of freedom
##   (12 observations deleted due to missingness)
## AIC: 6106.8
##
## Number of Fisher Scoring iterations: 6
```

Interact plot

```
library(jtools)

## Warning: package 'jtools' was built under R version 3.4.3

interact_plot(PIDGenderlogit, pred=PID7_new, modx=Gender, interval = TRUE,
x.label = "Political Identity", y.label = "Probability of
Switching",main.title = "Political Identity and Gender Interaction" )
```



Comparing the 2 models using Stargazer

```
stargazer::stargazer(baselogit, PIDGenderlogit, type = "text")
```

```
##
## ======================================================================
##                                               Dependent variable:
##                                           ----------------------------
##                                                  VotedForTrump
##                                              (1)             (2)
## ----------------------------------------------------------------------
## as.factor(PID7)Lean Democrat              -1.311***
##                                            (0.128)
##
## as.factor(PID7)Lean Republican             1.323***
##                                            (0.173)
##
## as.factor(PID7)Not very strong Democrat   -0.743***
##                                            (0.098)
##
## as.factor(PID7)Not very strong Republican  1.098***
##                                            (0.132)
##
## as.factor(PID7)Strong Democrat            -1.926***
##                                            (0.108)
```

```
##
## as.factor(PID7)Strong Republican           1.848***
##                                            (0.216)
##
## GrantLegalStatus                          -0.340***      -0.331***
##                                            (0.077)        (0.077)
##
## IncreaseBorderPatrols                      0.554***       0.551***
##                                            (0.073)        (0.073)
##
## Dreamers                                  -0.563***      -0.563***
##                                            (0.075)        (0.075)
##
## DeportIllegal                              0.860***       0.864***
##                                            (0.077)        (0.077)
##
## GenderMale                                 0.413***       0.235***
##                                            (0.072)        (0.081)
##
## PID7_new                                                  0.784***
##                                                           (0.033)
##
## as.factor(Race)Black                      -1.373***      -1.369***
##                                            (0.132)        (0.132)
##
## as.factor(Race)Hispanic                    -0.224*        -0.241*
##                                            (0.126)        (0.126)
##
## as.factor(Race)Other                      -0.619***      -0.622***
##                                            (0.147)        (0.146)
##
## as.factor(Education)2-year                  0.097          0.101
##                                            (0.120)        (0.120)
##
## as.factor(Education)4-year                  -0.051         -0.066
##                                            (0.104)        (0.104)
##
## as.factor(Education)High school graduate   0.294***       0.291***
##                                            (0.098)        (0.098)
##
## as.factor(Education)No HS                   0.051          0.082
##                                            (0.308)        (0.307)
##
## as.factor(Education)Post-grad               -0.182         -0.184
##                                            (0.127)        (0.126)
##
```

```
## unemployed                                 0.071          0.064
##                                            (0.169)        (0.170)
##
## family_income                              0.025**        0.028**
##                                            (0.012)        (0.012)
##
## mediaindex                                 0.109***       0.107***
##                                            (0.032)        (0.032)
##
## TerrorCamp                                 0.155**        0.153**
##                                            (0.074)        (0.074)
##
## Religion                                   0.307***       0.310***
##                                            (0.033)        (0.033)
##
## EconomyGotten much better                  0.052          0.045
##                                            (0.143)        (0.140)
##
## EconomyGotten better                      -0.775***      -0.766***
##                                            (0.091)        (0.090)
##
## EconomyGotten worse                        0.871***       0.863***
##                                            (0.090)        (0.091)
##
## EconomyGotten much worse                   1.427***       1.420***
##                                            (0.148)        (0.149)
##
## GenderMale:PID7_new                                      -0.207***
##                                                           (0.045)
##
## Constant                                  -2.561***      -2.443***
##                                            (0.178)        (0.169)
##
## -----------------------------------------------------------------------
## Observations                               14,970         14,970
## Log Likelihood                            -3,030.271     -3,028.400
## Akaike Inf. Crit.                          6,118.543      6,106.799
## =======================================================================
## Note:                                 *p<0.1; **p<0.05; ***p<0.01
```