Mini-Project 2

Entong Li, Zijia Cao, Xindi Lu, Jingyu Xie

11/20/2019

Introduction:

Research Question:

To what extent do attitudes toward immigration explain the switching of votes of 2012 Obama supporters who became 2016 Trump supporters?

Exploration steps:

1. Clean the data, only include the people who took both tests and vote to Obama in 2012.
2. As all demographic variables are meaningful and should be included in the model, find out the which demographic variables will interact with immigration attitude, then add the interaction term in the model.
3. Fit two models: one include immigration attitude, another doesn't, find out that whether the probability of switching votes for selected demographic variables will be affected by the immigration attitude.

Data Description:

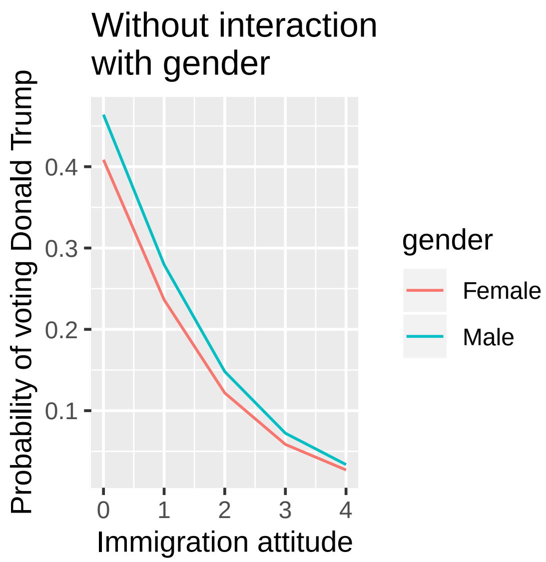
We'll use the 2016 Cooperative Congressional Election Study, which is a national survey that colllects the responses of 64,600 investigators of the same set of questions before and after the election, but not all investigtor who already done the pre-investigation will do the post one. The dataset could be accessed from: http://cces.gov.harvard.edu/data.

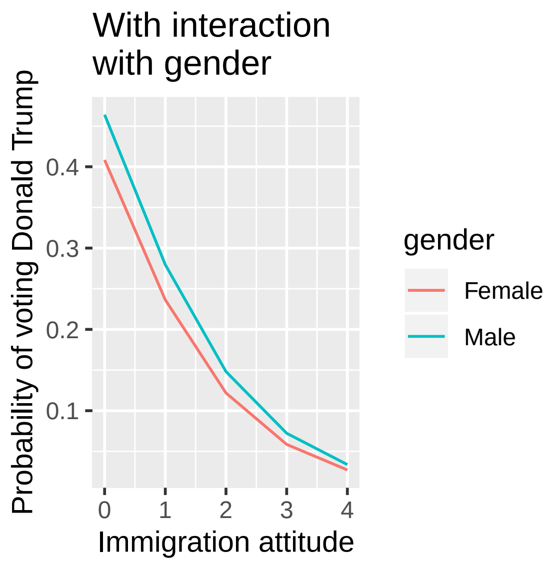
Data cleaning:

As we want to know about the switching of the votes from Obama to Trump, we will only include the investigators who took the post-election survey, and who voted to Obama in 2012. Overall, the dataset we used to explore will only have 23,395 observations.

Step 1 & Step 2:

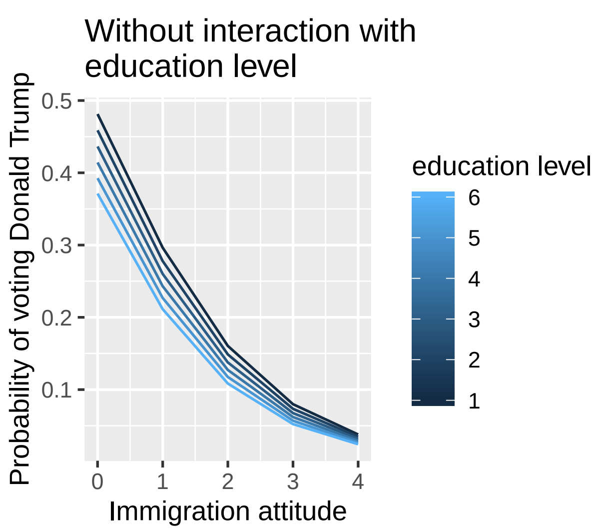
Immigration attitude with gender:

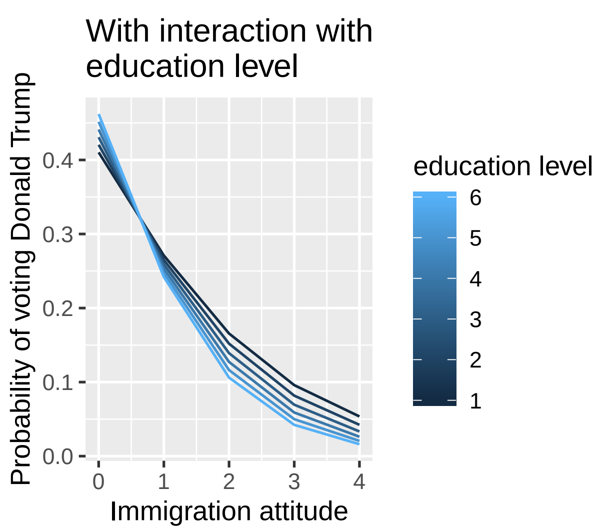




The immigration attitude doesn't interact with gender. Because the slopes for male and female from the model without interaction with gender do not have difference with the slopes of male and female from the model with interaction with gender.

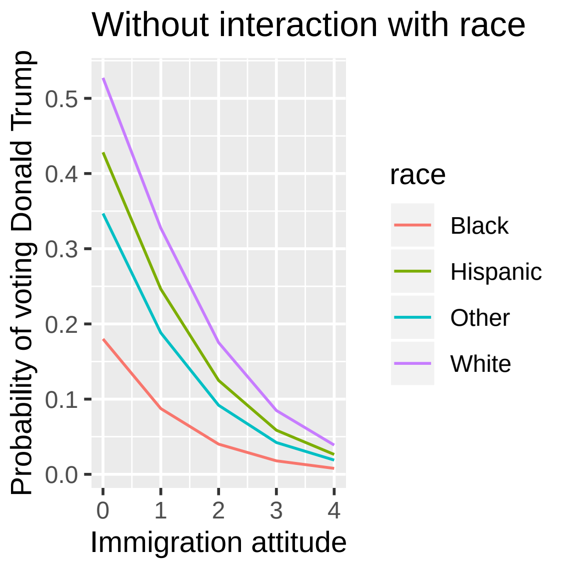
Immigration attitude with education level:

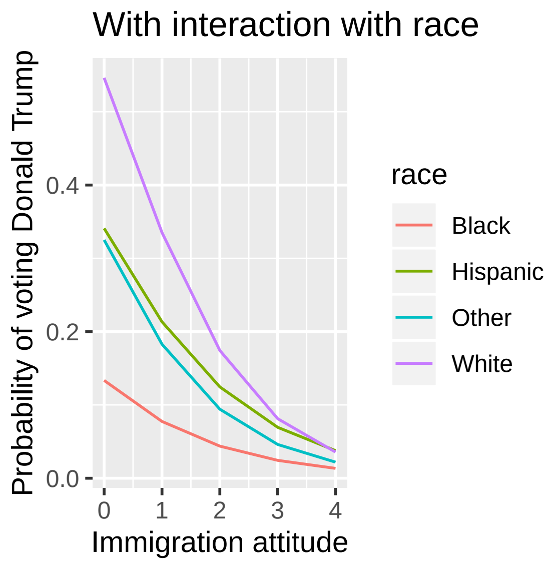




The immigration attitudes interact with educational level. Because for the model without interaction with education level, each level of education level seems to have similar, almost the same, slope, and when the immigration attitude is the same, higher level of education will have higher probability of voting Donald Trump. In contrast, for the model with interaction with education level, different education level will have different slope, and all of them will intersect when immigration attitude is around 0.7. As the education level will affect the probability of voting Donald Trump differently when there is an interaction between education level and immigration attitude, the interaction with education level should be included in model.

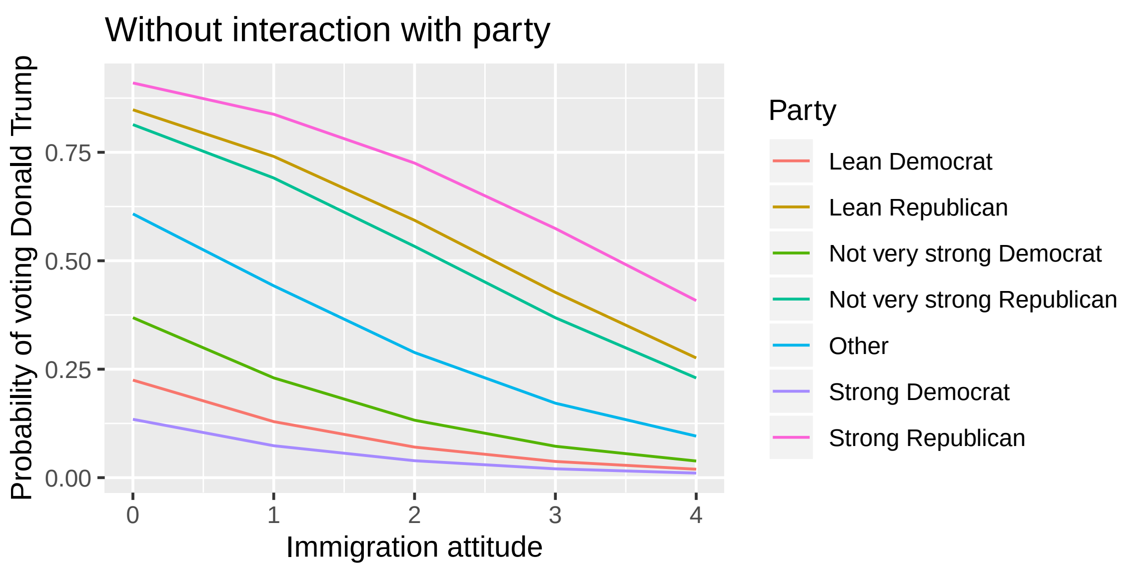
Immigration attitude with race:

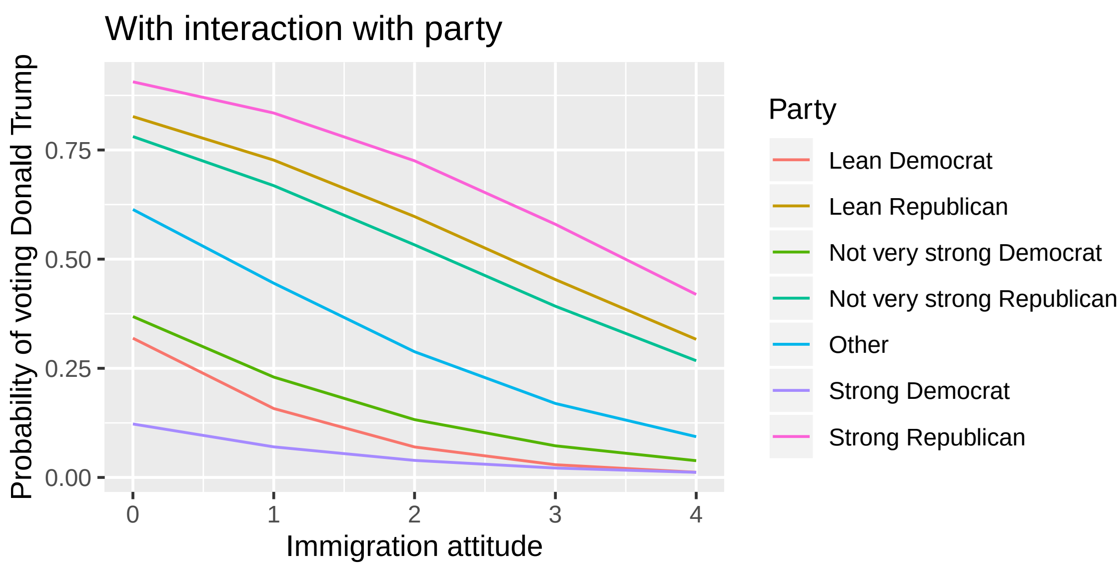




The immigration attitude interacts with the race. Because in model includes the interaction with race, Hispanic, Other, and White seem to have different slope from in the model does not include the interaction with race. Thus the interaction between immigration attitude and race should be included.

Immigration attitude with party:





The immigration attitude interacts with party. In the model with interaction with party, Lean Democrat, Lean Republican, Not very strong Republican, and Strong Democrat have different slope from in the model without interaction with party. Thus the interaction between immigration attitude and party should be included.

Step 3:

The coefficients from the model without immigration attitude:

```
## glm(formula = vote.T ~ race + pid7 + educ, family = "binomial",
##     data = obama, weights = commonweight_vv_post)
##                                  coef.est coef.se
## (Intercept)                      -2.44     0.08
## raceBlack                        -1.01     0.09
## raceHispanic                     -0.54     0.10
## raceOther                        -0.72     0.11
## pid7Not very strong Democrat      1.48     0.08
## pid7Lean Democrat                 0.45     0.10
## pid7Strong Republican             4.55     0.15
## pid7Not very strong Republican    3.49     0.10
## pid7Lean Republican               3.98     0.13
## pid7Other                         2.50     0.08
## educ                             -0.23     0.02
## ---
##   n = 22208, k = 11
##   residual deviance = 10618.5, null deviance = 14425.0 (difference = 3806.5)
```

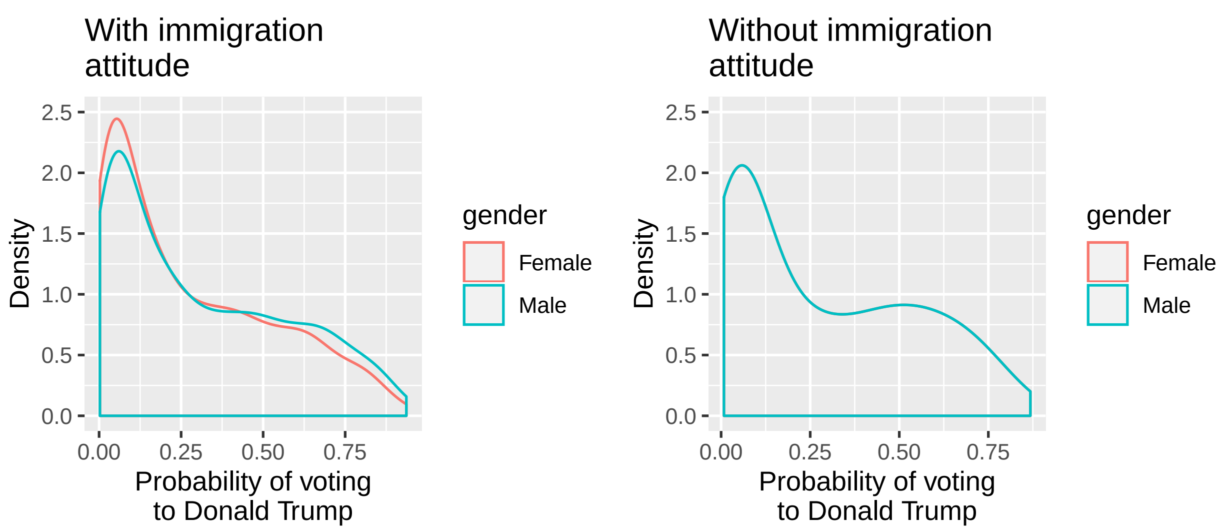
The coefficients from the model with immigration attitude:

```
## glm(formula = vote.T ~ immig.at + race + pid7 + educ + gender +
##     immig.at:race + immig.at:educ + immig.at:pid7, family = "binomial",
##     data = obama, weights = commonweight_vv_post)
##                                           coef.est coef.se
## (Intercept)                               -1.30     0.15
## immig.at                                  -0.46     0.07
## raceBlack                                 -1.33     0.16
```

```
## raceHispanic                          -0.81     0.21
## raceOther                             -1.06     0.19
## pid7Not very strong Democrat           1.27     0.14
## pid7Lean Democrat                      0.94     0.19
## pid7Strong Republican                  4.02     0.30
## pid7Not very strong Republican         3.02     0.19
## pid7Lean Republican                    3.24     0.24
## pid7Other                              2.18     0.15
## educ                                  -0.03     0.03
## genderFemale                          -0.22     0.05
## immig.at:raceBlack                     0.12     0.07
## immig.at:raceHispanic                  0.19     0.08
## immig.at:raceOther                     0.10     0.09
## immig.at:educ                         -0.07     0.01
## immig.at:pid7Not very strong Democrat -0.02     0.06
## immig.at:pid7Lean Democrat            -0.24     0.08
## immig.at:pid7Strong Republican         0.04     0.13
## immig.at:pid7Not very strong Republican 0.06    0.08
## immig.at:pid7Lean Republican           0.11     0.11
## immig.at:pid7Other                    -0.01     0.06
## ---
##   n = 22208, k = 23
##   residual deviance = 9487.2, null deviance = 14425.0 (difference = 4937.8)
```

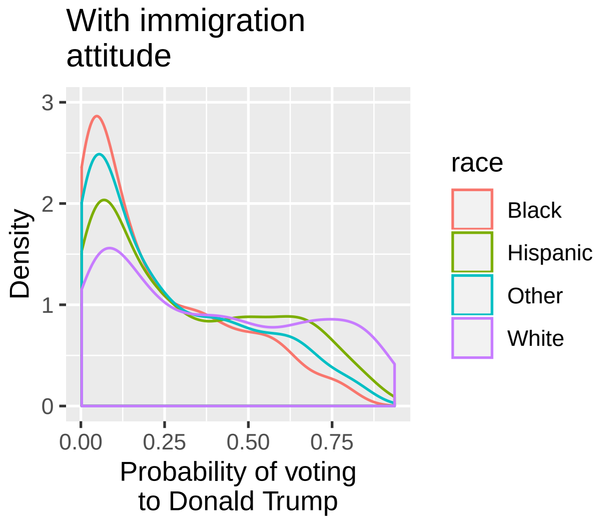
Graphic comparison:

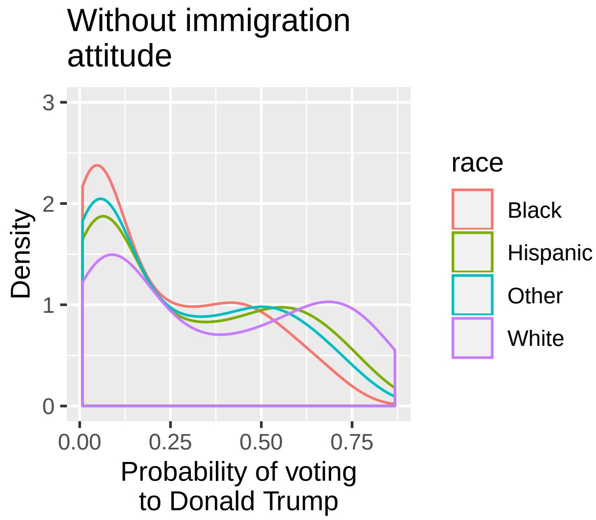
Gender:



The immigration attitude does make some difference when talking about gender. In the density plot without immigration attitude, lines for male and female seem to be overlapped which means that they tend to have the same probability of voting to Donald Trump. However, from the plot with immigration attitude, the line for male deviates from the line for women. In other words, taking immigration attitude into consideration results in some changes in probability of voting to Donald Trump between different gender.

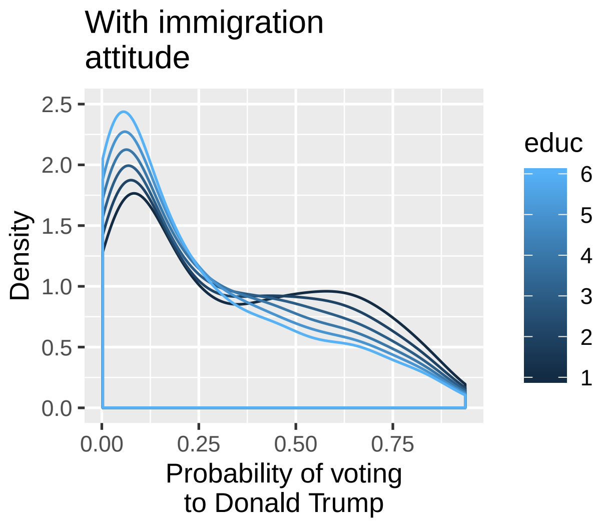
Race:

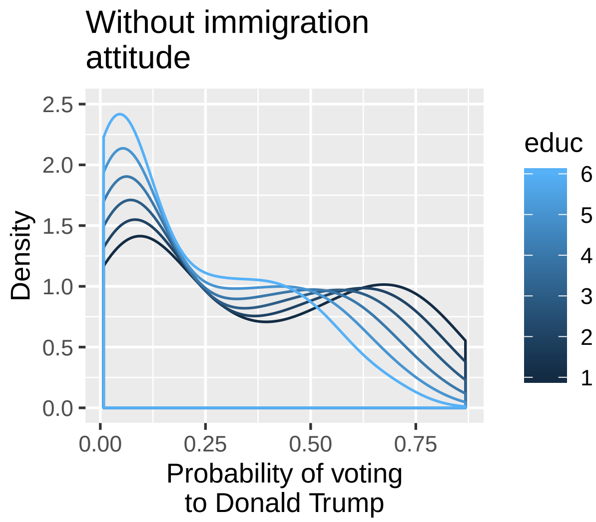




The immigration attitude does not make a substantive difference when talking about race. From the density graph, it seems like the overall shape for probability of voting to Donald trump for all four races doesn't change a lot when we add immigration attitude in the model.

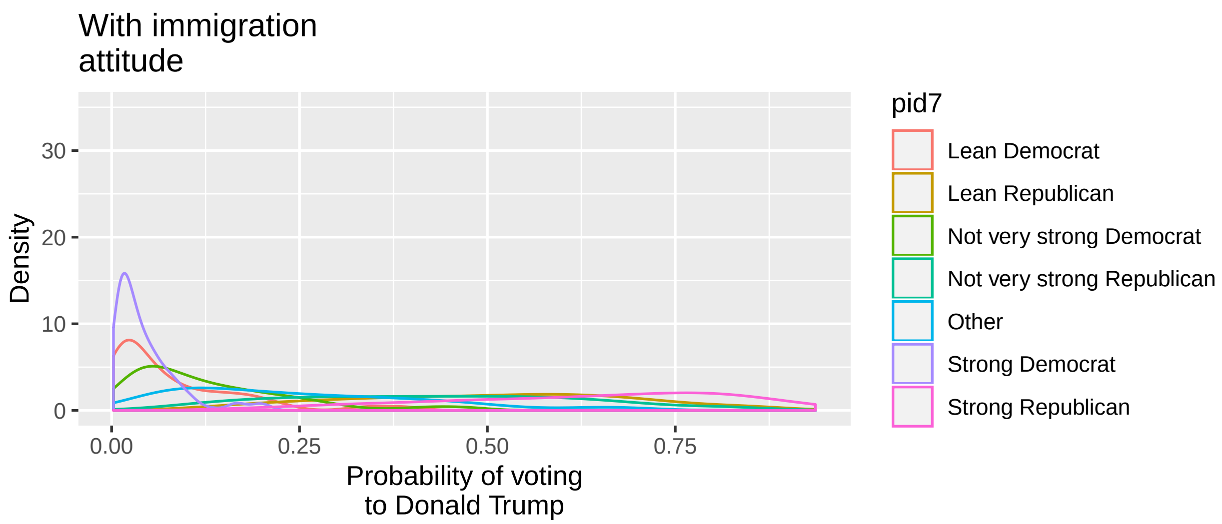
Education level:

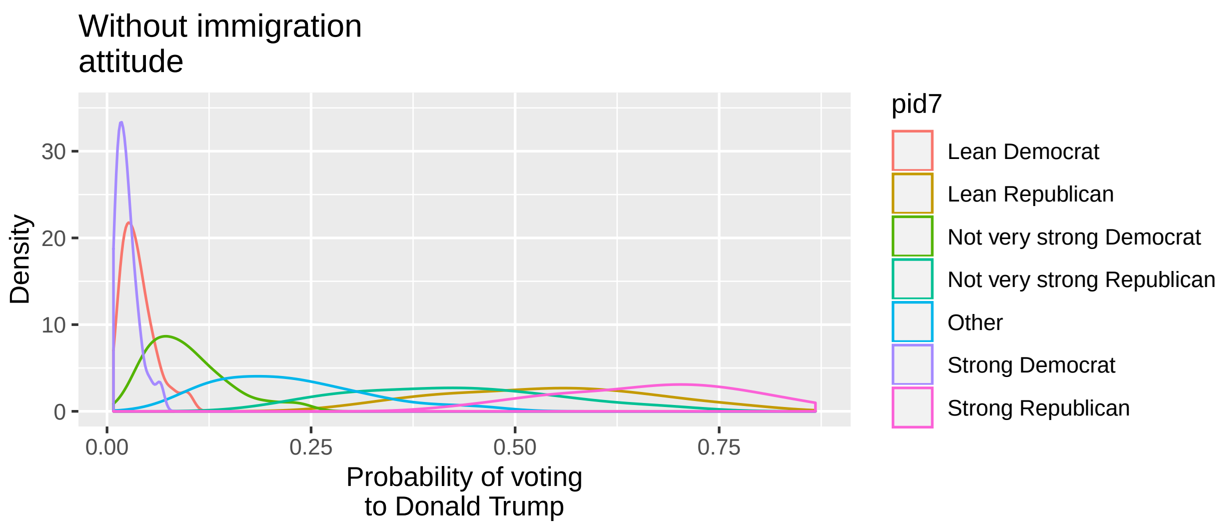




For model with variable education, adding immigration attitude does make a substantive difference. The plot shows that without immigration attitude, there exist relatively larger gaps in probabilities between each educational levels. After including immigration attitude in the model, the differences of probabilities between various educational levels become smaller than before.

Party:





From the plot, we could see that the immigration attitude does make a substantive difference when talking about party. It's obvious to find that for parties like Strong Democrat and Lean Democrat ,their probabilities of voting to Donald Trump shrink a lot when including immigration attitude in the model.

Overall, including immigration attitudes make a substantive difference in demographic groups: gender, education level, and party. For these three demographic groups, immigration attitude seems to have the same level of effect on them.

Conclusion:

The attitudes toward immigration explain the probability of switching of votes of 2012 Obama supporters who became 2016 Trump supporters within the demographic groups: gender, education level, and party.