Obama to Trump: The American shift

Authors: Choi Jane, Han Da Hee, Mestre Luis Miguel, Watson, Katharine, & Xu, Xiao

Team Hungary-Sri Lanka

Some contacts emails: lmestre@iu.edu, dh1@iu.edu , xx1@iu.edu

Introduction

Barack Obama won the Presidency of the United States (US) in 2008, becoming the first African-American president of the nation's history winning the popular vote and the Electoral College vote. Obama represented a change in the US politics being a central-left politician that explicitly supports increment of the minimum wage, human rights to the LGBTIQ+ community, more accessibility to healthcare, and more support to immigrants. He got reelected in 2012 with both the popular vote and the Electoral College vote. During the second term, it seems that Americans were moving to the left, becoming a more liberal country, mainly respect human rights and immigration. Policies as DACA and other refugees program were created to those who were children of an illegal immigrant or those affected by civil wars can have legal status in the US. However, during the 2016 election a non-expected shift occurs, Donald J. Trump won both the nominee of the Republican Party (GOP) and won the general election only through the college electoral vote.

The polls have shown that people from certain demographic groups that supported Obama during his terms, were more likely to switch their votes than others. However, it is not very clear what might explain why some people within a group switched, while others did not. One of the theories is that attitudes toward immigration became especially salient during the 2016 campaign. ***Our main objective*** is to be exploring the data about the 2016 elections and then fit a "complicated" logistic regression model to get a sense of whether attitudes had explanatory power over and above demographic shifts.

Objective

To accomplish the main objective we used the 2016 Cooperative Congressional Election Study dataset, a very large survey of a nationally representative sample of 64,000 adults. The investigators asked questions prior and post elections, however not all respondents prior elections replied post-elections. The dataset contains *demographic variables* (i.e., gender, education, race, and party identification), *technical variables* (i.e., survey weights and if the respondent took the post-election survey), *voting variable* (i.e. respondent voted in 2012 elections for Barack Obama, respondent voting decision in 2016 election) and *immigration variables* focus on what the US government should do about immigration (i.e., granting legal status to illegal immigrants that pay taxes, no felony crimes, etc., decision about border patrols in US-Mexican border, grant legal status to people who were brought to the US illegally, and deport illegal immigrants).

All variables are important to understand the switching from Obama to Trump. With such information, we will have two main aims. Aim 1 to explain if there is any interaction between immigration attitude and demographic variables. Aim 2 to determine if immigration attitudes make a substantive difference and if it is a matter for some demographic groups than others.

Methods

The approach of this exploratory analysis will be through logistic regressions. Logistic regression is a statistical model that explains a binary dependent (only two outcomes) variable in a logistic form. It is used to predict, or measure risks and odds (the ratio of the probability of success respect to the probability of not). This statistical model often has the mathematical form $\log(odds) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_k X_k$, where β is the linear coefficients that represent the odds ratio of the kth covariate X (predictor variable) respect to the log transformation of the odds of success. The odds ratio is respect to the odds of a group of interest respect to the odds of a reference group (the group that is used to compare outcomes). Normally, they are interpreted respect the group of interest having risk more or less amount of times than the reference group.

As other statistical model, the logistic regression have some assumptions in its residuals that should be verified. Residuals are the difference error between the ith observed outcome and the ith predicted outcome. It is assumed: (1) they follow a normal distribution with mean zero and constant variance ($\epsilon \sim N(0, \sigma_\epsilon^2)$), (2) homoscedasticity (equal variances) and (3) they are independent (we do not need an error to calculate other).

The methods applied are based on the main Aims. For Aim 1, we fitted logistic models respect to the Obama-Trump switching (a dichotomous variable where is 1 if the shift happens and zero otherwise) using immigration attitudes (how pro-immigrant someone can be on a scale 0 to 4, where 0 is no pro-immigrant and 4 is being highly pro-immigrant) as a predictor along with the education, gender, race, and party identification. The logistic models have some level of complexity because the interactions between immigration attitude respects to demographic variables were considered. The residuals of the logistic regression were verified to check the assumptions of this kind of model (e.g., verified to which extent the results of the model can be trusted). For Aim 2, we fitted two weighted logistic regressions for 2012 Obama voters that switch to Trump 2016 one without immigration attitude and other with immigration attitude. In statistics, the weighting is the emphasis of an outcome respect to a particular interest (e.g., probability of something occurring, population size, survey size). For this case, we weighted respect to the respondents who took the post-election survey (e.g., survey weights).

Results

The results will be given by the main aims as describe in the objectives.

Aim 1 - to explain if there is any interaction between immigration attitude and demographic variables

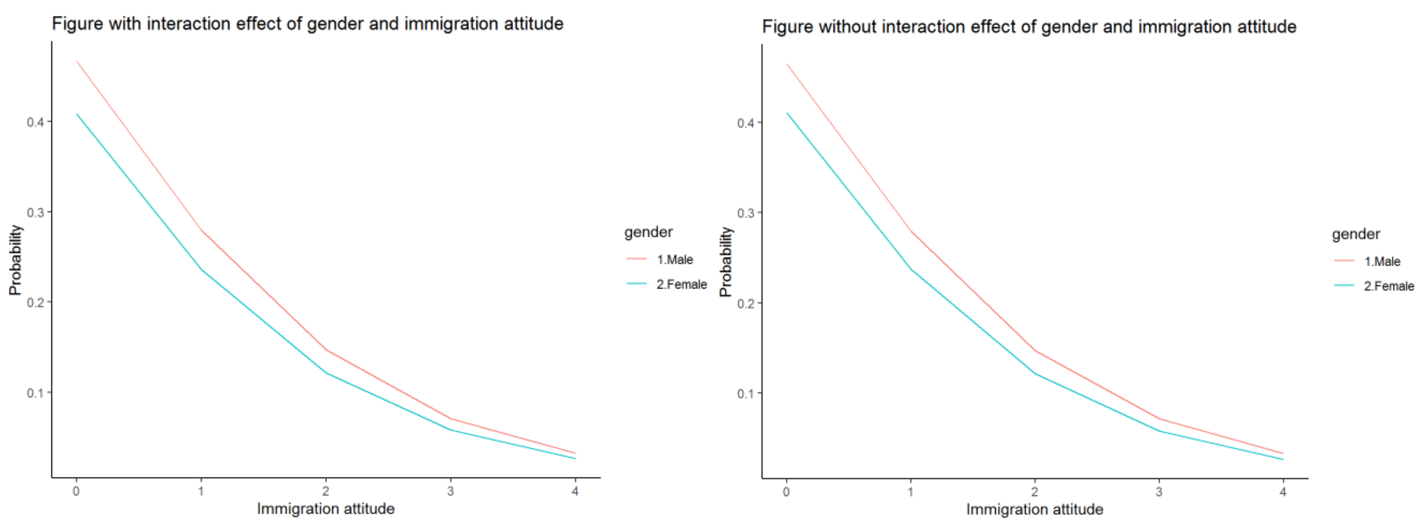

Figure 1 - Attitude and gender interaction effect plot

Figure 1 shows the interaction of the attitude about immigration respect to gender. Both plots shows that higher the attitude about immigration, lower the probability of voting for Trump. However, there is no significant difference between the model with and without interaction term of gender and immigration attitude.

Figure 2 shows that doesn't matter the education level, it seems that those who have a worst pro-immigration attitude have a higher probability of voting for Trump. Therefore, the more pro-immigrant, lower the probability of voting for Trump, regardless the education level. In this case, there is interaction effect regarding education levels and immigration attitude. Probability without interaction effect suggests nearly paralleling decreased whereas curves of probability with interaction effect crossed. Thus, we cannot ignore the effect between education levels and immigration status.

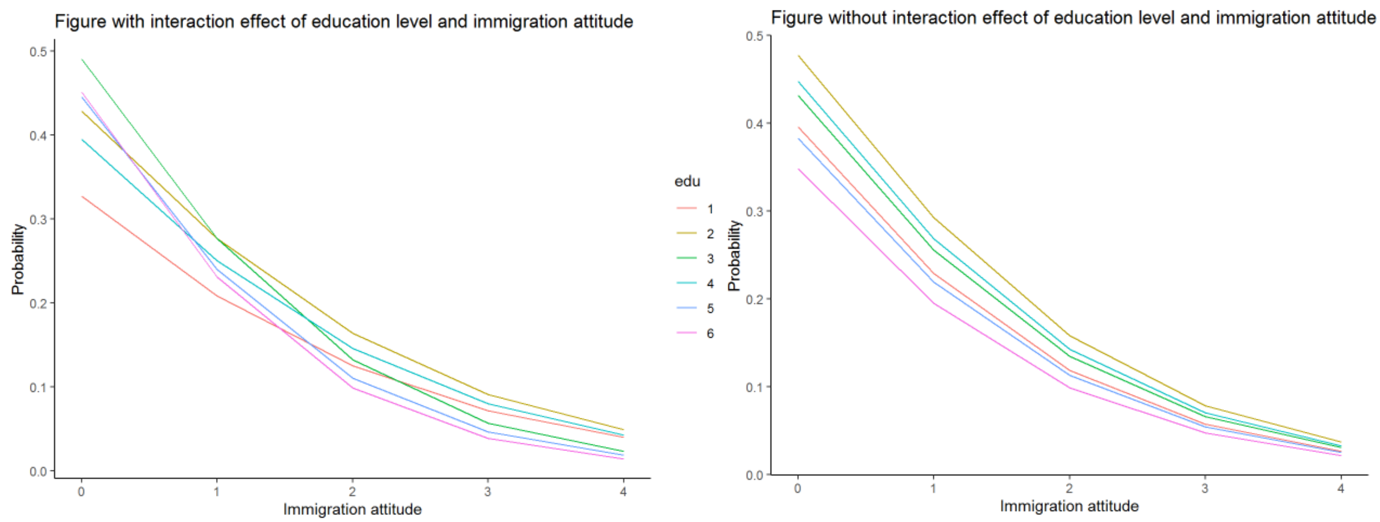

Figure 2 - Attitude and Education level interaction effect plot

Figure 3 shows the effect by race and immigration attitude. Based on the plot, there is some differences by race. All races have the same trend that higher the score of pro-immigration attitude lower the probability of voting for Trump. However, the slope changes. The Whites and Others categories have similar slopes, the probability for voting for Trump is very high is there is no pro-immigration attitude and very low if there is the highest score for pro-immigration attitude. Blacks, in the other hand have the lowest probability of voting for Trump even though not having any pro-immigration attitude score. Interestingly, they are also the lowest probability of voting for Trump if the score pro-immigration attitude is four (the highest). There is interaction effect regarding race and immigration attitude. Probability without interaction effect suggests nearly paralleling decreased whereas curves of probability with interaction effect has significant difference. Thus, we cannot ignore the effect between race and immigration status.

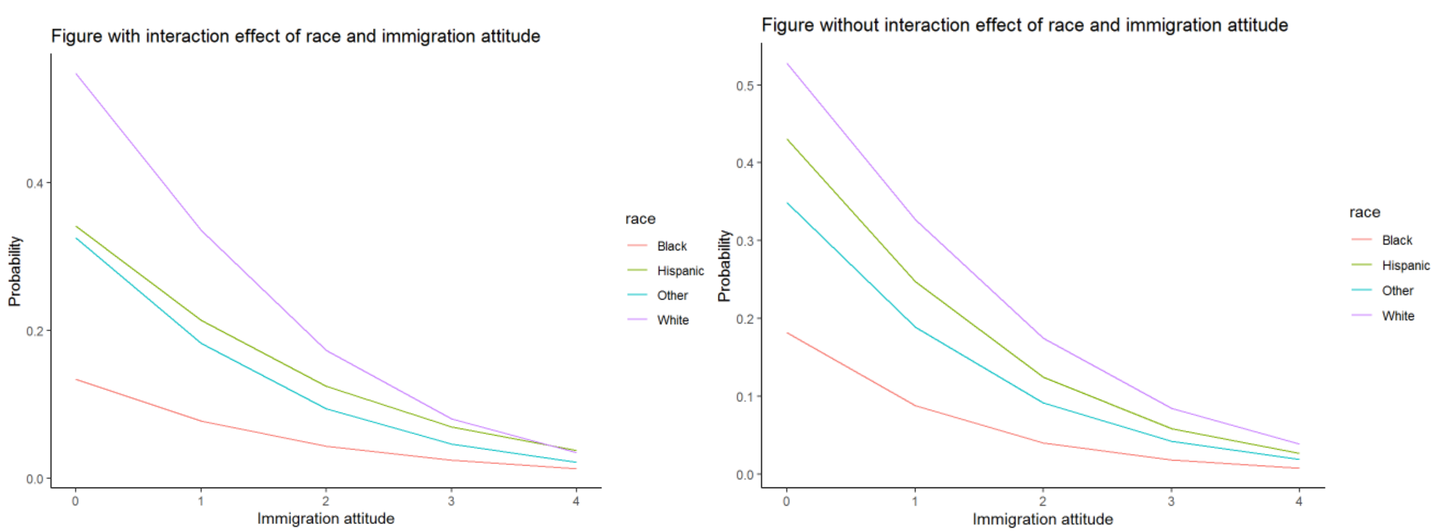

Figure 3 - Attitude and race interaction effect plot

Figure 4, shows the pro-immigration attitude score by party level (higher the value, the more republican is the party level, and lower the score, the more democrat the respondent is). The higher the party level, higher the probability of voting for Trump, and lower the party level, lower the probability of voting for Trump, regardless the pro-immigration attitude score. However, it is still observed the pattern that lower the pro-immigration attitude score, higher the probability of voting for Trump, however it is very clear that such probability will decrease the less republican is the person, regardless the pro-immigration attitude score. There is interaction effect regarding party and immigration attitude. Probability without interaction effect suggests nearly paralleling decreased whereas curves of probability with interaction effect crossed. We cannot ignore the effect between party and immigration status.

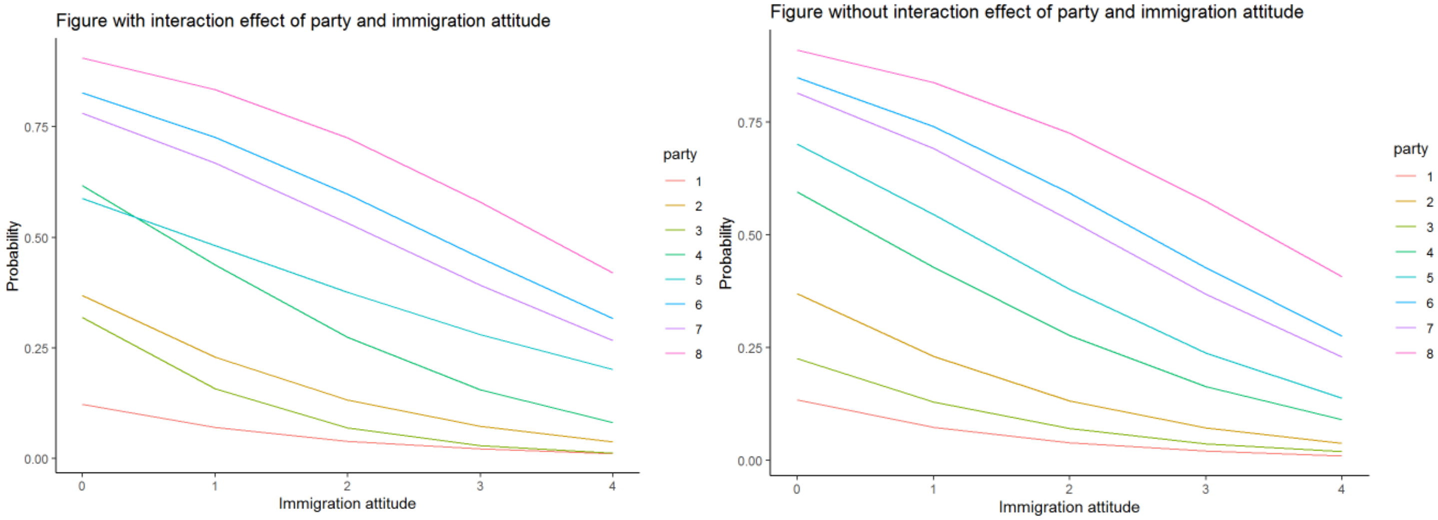

Figure 4 - Attitude and Party interaction effect plot

Aim 2 - to determine if immigration attitudes makes a substantive difference and if it is matter for some demographics groups than others (For final submission).

For selected demographic groups, the probability of switching from Obama to Trump (based on Model 1) is shown in Table 3. The biggest difference between the two models is apparent in the male and female, least educated, white, most anti-immigration, strong democrat group. Based on Model 1, the probability of males of this group switching increases from 2.3 to 19.3% (an 8.4-fold increase) when the immigration attitude changes from being pro- to anti-immigration. The probability that a female of this groups would make this change increases from 2.1 to 14.3% (a 6.8-fold increase). Based on Model 2, the probability of switching for the males of this group is 8.6%. The probability of the females switching is 6.8%.

Table 1. Probability of switching from voting for Obama in 2012 to Trump in 2016 for selected demographic groups, using Model 1.

Immigration attitude	Sex	Race	Education level	Political party	Probability of switching
4 (Most'pro')	Female	White	Highest	Strong dem	0.74%
4 (Most'pro')	Male	White	Highest	Strong dem	0.81%
1 (Most'anti')	Female	White	Highest	Strong dem	5.5%
1 (Most'anti')	Male	White	Highest	Strong dem	7.6%
4 (Most'pro')	Female	Black	Highest	Strong dem	0.52%
4 (Most'pro')	Male	Black	Highest	Strong dem	0.57%
1 (Most'anti')	Female	Black	Highest	Strong dem	2.3%
1 (Most'anti')	Male	Black	Highest	Strong dem	3.3%
4 (Most'pro')	Female	White	Lowest	Strong dem	2.1%
4 (Most'pro')	Male	White	Lowest	Strong dem	2.3%
1 (Most'anti')	Female	White	Lowest	Strong dem	14.3%
1 (Most'anti')	Male	White	Lowest	Strong dem	19.3%
4 (Most'pro')	Female	Black	Lowest	Strong dem	0.68%
4 (Most'pro')	Male	Black	Lowest	Strong dem	0.74%
1 (Most'anti')	Female	Black	Lowest	Strong dem	3.0%
1 (Most'anti')	Male	Black	Lowest	Strong dem	4.3%

Probabilities were also calculated for Model 2 (the simplified model without immigration attitude or any interaction terms involving immigration). Model 2 does not fit these data as well. These values are shown in Table 2.

Table 2. Probability of switching from voting for Obama in 2012 to Trump in 2016 for selected demographic groups, using Model 2.

Sex	Race	Education level	Political party	Probability of switching
Female	White	Highest	Strong dem	1.7%
Male	White	Highest	Strong dem	2.2%
Female	Black	Highest	Strong dem	1.0%
Male	Black	Highest	Strong dem	1.3%
Female	White	Lowest	Strong dem	6.8%
Male	White	Lowest	Strong dem	8.6%
Female	Black	Lowest	Strong dem	1.7%
Male	Black	Lowest	Strong dem	2.1%

Table 3 - Model considering immigration terms

Coefficient	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.98148	0.37028	-2.651	0.008039
proinmi	-0.61287	0.08804	-6.961	3.47E-12
genderMale	0.44651	0.09651	4.627	3.74E-06
Lean Democrat	-1.82278	0.10242	-17.797	< 2e-16
Lean Republican	1.31513	0.14212	9.253	< 2e-16
NA	1.51954	0.5069	2.998	0.002723
Not sure	0.5731	0.16104	3.559	0.000374
Not very strong democrat	-1.05232	0.0766	-13.738	< 2e-16
Not very strong republican	0.92378	0.10306	8.964	< 2e-16
Strong Democrat	-2.32065	0.08328	-27.865	< 2e-16
Strong Republican	1.89507	0.16222	11.682	< 2e-16
Race Black	0.40182	0.42735	0.94	0.347103
Race Hispanic	0.63533	0.45746	1.389	0.164906
Race White	2.40977	0.37085	6.498	8.33E-11
Education Level	0.14943	0.0741	2.017	0.043741
Proinmi*gender Male	-0.08831	0.04138	-2.134	0.032833

Table 4 - Model without immigration terms

Coefficient	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-2.13005	0.3397	-6.27	3.67E-10
Lean Democrat	-2.16465	0.10008	-21.629	< 2e-16
Lean Republican	1.50507	0.13359	11.267	< 2e-16
NA	0.85886	0.49994	1.718	0.08582
Not sure	0.46001	0.15484	2.971	0.00297
Not very strong Democrat	-1.1311	0.07284	-15.528	< 2e-16
Not very strong Republican	0.96261	0.09784	9.839	< 2e-16
Strong Democrat	-2.62447	0.08076	-32.499	< 2e-16
Strong Republican	2.02438	0.15243	13.281	< 2e-16
Gender Male	0.25149	0.05156	4.877	1.08E-06

Race Black	0.76555	0.39123	1.957	0.05039
Race Hispanic	0.86048	0.40848	2.107	0.03517
Race White	2.43267	0.3409	7.136	9.93E-13
Education Level	0.14327	0.07347	1.95	0.0512
Race Black * education Level	-0.24147	0.09577	-2.521	0.0117
Race Hispanic * Education Level	-0.11179	0.10196	-1.096	0.27295
Race White * education Level	-0.4344	0.07615	-5.705	1.18E-08

An analysis of variance with Chi-square was used to compare the GLM models. There was a significant difference between the two models ($p < 0.001$). Therefore, the more complex model, which includes immigration and its interactions, needs to be kept (Figure 1). The coefficients for each model are shown below (Tables 3 and 4).

Discussion

Demographics factors such as education level, race identity and political affiliation have an interaction that has an effect in the prediction of the probability of voting for Trump in 2016, given a vote for Obama in 2012 respect to the pro-immigration attitude. The three demographics significant factors shown a similar pattern for all its categories: better the pro-immigration attitude lower the risk of voting for Trump in 2016. All education levels shows this pattern but in different trends, where higher the level of the education of a person and its pro-immigration attitude, lower the probability of voting for Trump in 2016. Similarly, by race identity and pro-immigration attitude is shown that those in ethnic minorities (i.e. non-whites) have a lower probability of voting for Trump in 2016.

However, by political affiliation is not a surprise the pattern shown. High values in this ordered variable represents those who strong support to the Republican Party and those with a lower value represents a strong support to the democrat party. Higher the value of the political affiliation by pro-immigration attitude higher the probability of voting for Trump.

For selected demographic groups, the probability of switching from Obama to Trump (based on Model 1) is shown in Table 1. The biggest difference between the two models is apparent in the male and female, least educated, white, most anti-immigration, strong democrat group. Based on Model 1, the probability of males of this group switching increases from 2.3 to 19.3% (an 8.4-fold increase) when the immigration attitude changes from being pro- to anti-immigration. The probability that a female of this groups would make this change increases from 2.1 to 14.3% (a 6.8-fold increase). Based on Model 2, the probability of switching for the males of this group is 8.6%. The probability of the females switching is 6.8%, indicating that the interaction of some demographics factors play a role in understanding the switch between of Obama 2012 to Trump 2016.

Limitations

The immigration variables were merge into a new ordered variable or simplification. Our rank about immigration (0 no pro-immigrant and 4 high pro-immigrant) is a scale that measure the “big picture” about how those who voted for Obama in 2012 and then for Donald Trump in 2016 feels about immigration. However, we did lose the information of what specific topic of immigration they do or do not support (for values 1, 2 and 3), having some information bias. Only the survey’s respondent post-election were consider in the analysis, having some selection bias due to the elimination of those who participated prior 2016 elections but not after election, having some possibility of selection bias.