

POLS 306

Dr. Brichoux/Bronson Herrera

3 May, 2017

Theory and Hypothesis

The 2016 presidential election proved to be a contentious race that resulted in a way many pundits did not anticipate. Both of the major party candidates, Donald Trump and Hillary Clinton, proved to be incredibly divisive figures in American politics, so perhaps that had something to do with the radically different than expected outcome. In the current political sphere there is a myriad of issues that could be considered somewhat controversial or divisive, much like the political candidates, with one of the most major ones being feminism and anything related to it. Often being associated with the social justice movement, there has appeared to be a sharp divide between the two major political parties with the Democratic party being more in support of it and the Republican party being more against it. In this fierce election, did a person's feelings toward the issue of feminism make them approve of Hillary Clinton more or less? Analyzing the factors that went into liking a candidate ultimately can reveal the reasons behind people voting for certain candidates, which can paint us a more informative picture of what happened on election day and what made Hillary Clinton fail in spite of the pundits' predictions.

I hypothesize that the more positive a person's feelings are toward feminism, the more likely they were to view Hillary Clinton in a more favorable light. People who view feminism in a more favorable fashion would most likely be more supportive of Clinton's political agenda that revolves around social justice issues. They may also view her in a more approving fashion due

to the fact that she was a woman running for the highest office in the nation that has yet to be occupied by a woman. By contrast, those who do not harbor positive feelings toward feminism would be more likely to view Hillary Clinton's social justice oriented policies with disdain and may harbor no feelings or even hostile ones toward a woman running for President of the United States. Both of these reasons would account for variation in the feelings toward Hillary Clinton during this past election cycle.

Other possible sources of variation could be party affiliation as well as feelings toward the status quo at the time of the election, which I aim to control for in this paper. Due to strong partisan ties in American politics, people may view a candidate as being more favorable to them if they share the same party—in this case it would be the Democratic Party. The Democratic Party is also more friendly toward social justice issues, so it is quite likely that party affiliation would correlate with feelings toward both Hillary Clinton and feminism which could make it appear that there is a relationship between the two when there is in fact none. A person's approval rating for President Obama is telling of their feelings toward the status quo at the time, which Clinton campaigned on the basis of extending. This may correlate with party affiliation as well and similarly correlate with both feelings toward feminism and Hillary Clinton, which could also make the relationship appear stronger than it actually is.

In the ANES 2016 pilot study data, 1,200 completed surveys by individual provide all the data and thus it is broken down by individuals as units of analysis. The completed surveys came from all of the different states and thus the data should be representative of the entirety of the nation. However, the age of the participants was strictly those over the age of 18, or voting age, and thus anyone younger than that is not represented within the data set. However, the sample was not collected by random sampling and instead was through an opt-in process on the internet.

This means that the participants could be different from the general populous due to their willingness to participate and also because all of them had some kind of internet source. Overall, this data is most likely generalizable to the American population of voting age that have access to the internet, though the opt-in aspect makes it a bit more dubious as to whether or not it is truly representative.

Data: Operationalization and Description

Both the independent variable (IV) and the dependent variable (DV) were relatively easy to translate into real terms thanks to two items on the ANES 2016 survey. Feeling thermometers were taken over both participants' feelings toward Hillary Clinton and feminism on a scale of 0 to 100, with 0 being very cold or unfavorable and 100 being very warm or favorable. This method of converting a qualitative feeling into a quantitative scale is a rather effective method for operationalizing both the IV and DV. By looking at a person's feelings toward these issues, it allows for a translation of the theory into an applicable form for the real world. This translation process for both the DV and IV comes at the sacrifice of gaining insight into the specifics as to why a person feels a certain way about the candidate. Furthermore, it only provides for a rigid structure for a subject that can be incredibly wide ranging from one person to another and can vary in understanding between individuals. While this method of translation provides for the most reliable way to take qualitative emotions and turn them into something quantitative, there are still things being lost. Similarly, the control variable that measures the contentment with the status quo, the individual's approval rating of President Barack Obama, also takes something more complicated and intangible and turns it into a very simplistic scale that only ranges from 1 to 7. For party affiliation, more specifically whether or not a participant identifies with the Democratic Party, turning it into a quantifiable variable causes one to lose

sight of the political spectrum and sects within the political party. There is a great spectrum of conservatism to liberalism that can be found within the single entity that is not properly taken account of by simply using a measurement that requires participants to simply check off which party they belong to without delving any deeper.

In terms of recoding, the only major change made to both the IV and DV was the removal of responses that were missing, which were coded as 998 in both cases and easily recoded as being not available (NA) in the recoding process. The control variables required more work, as the status quo variable's values needed to be reversed for clarity and party affiliation was turned into a dummy variable. For the status quo, I used a variable that measured the participant's approval of President Obama's job performance. Originally, the data was coded with 1 being equivalent to strongly approving and 7 being equivalent to strongly opposing, thus I reversed the data so that 1 was equivalent to strongly opposing and 7 was equivalent to strongly approving. For political affiliation, I was interested in whether or not the participants identified with the Democratic Party and thus coded the Democratic Party option as being a 1 with Republican, Independent, and Other all being 0.

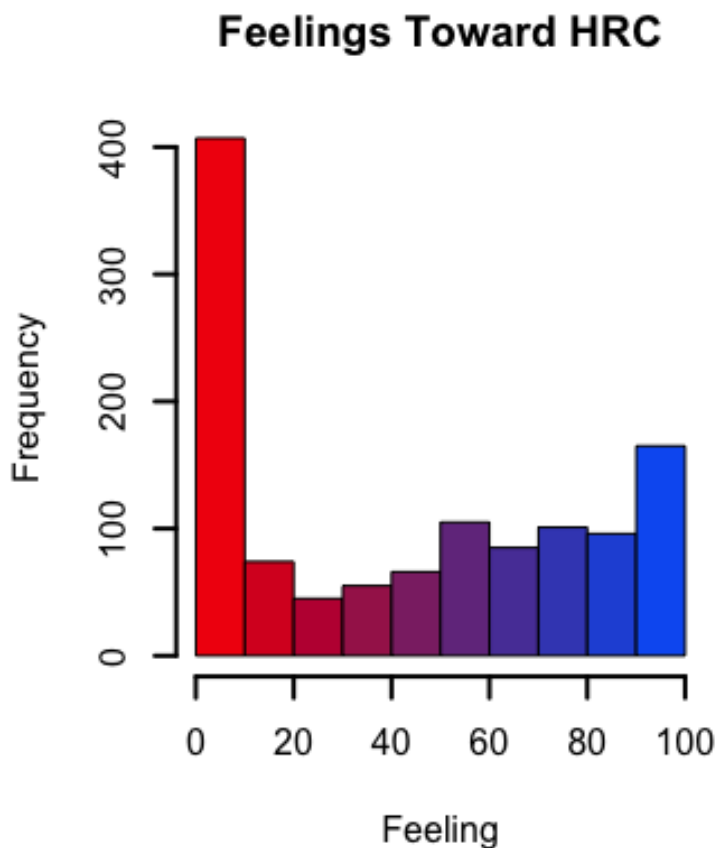
The data used within this paper originates from the ANES 2016 pilot study, which was sent out during the month of January from 2016. The respondents who participated in this study were all U.S. citizens that were at least 18 in age, overall 1,200 participants were used in this study. These 1,200 individuals came from all over the States and were part of an internet opt-in program, meaning that they had some form of access to internet be it of their own or at some local source. The data was collected via a survey on the internet that contained a variety of different kinds of questions, using both quantitative and qualitative data. This leads to data ranging from continuous variables to categorical variables and everything in between.

Table 1: Variable Summary

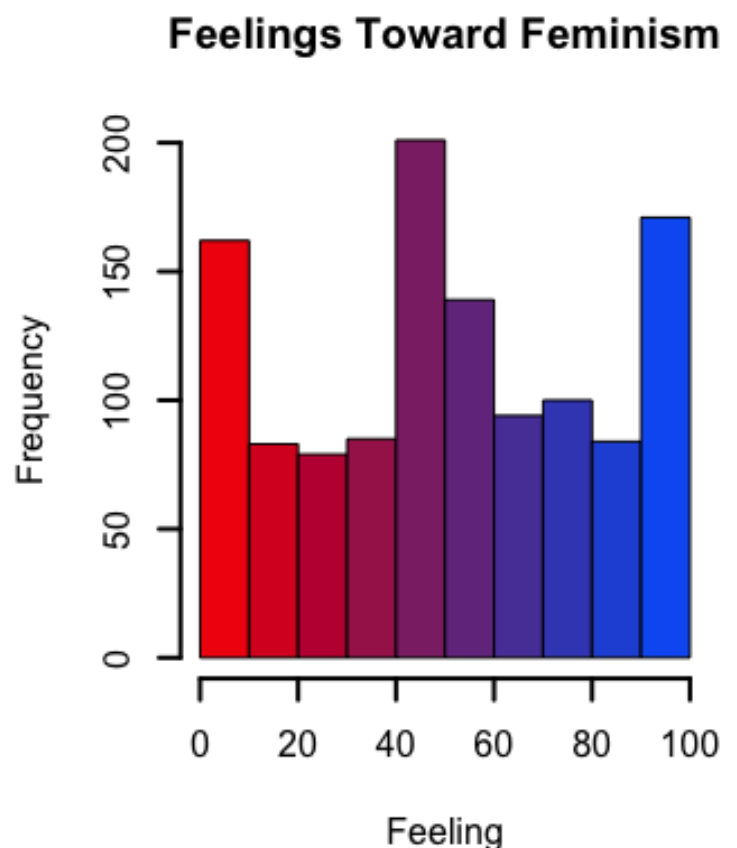
Variable	N	Mean	SD	Min	Max
Feelings_Toward_HRC	1200	42.9942	36.5047	0	100
Feelings_Toward_Feminism	1200	51.4691	30.4008	0	100
Obama	1200	3.8058	2.3610	1	7
DemID	1200	0.3842	0.4866	0	1

The summary table shows some of the most important information for the variables: the number, mean, standard deviation, minimum value, and maximum value. From just this bare minimum, we can glimpse that the IV and DV share similar means and standard deviations, which may indicate that the two correlate. Furthermore, both are somewhat normal based on the summary table information, though it appears that the feelings toward feminism is a much more normal distribution since the mean is considerably closer to 50. It appears that the DV and Obama variable may be more positively skewed, though it is not completely clear based on just the summary table.

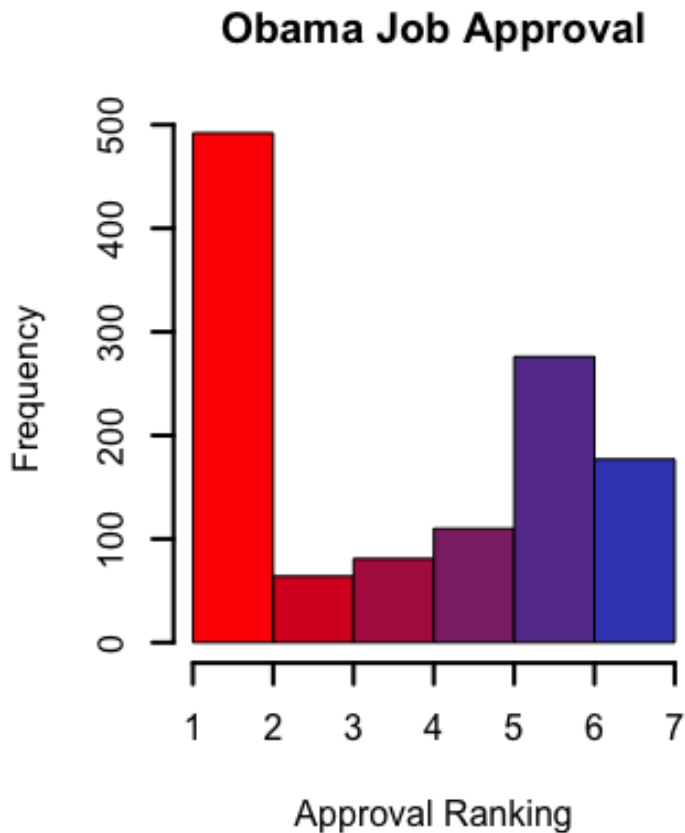
Graph 1: Feelings Toward HRC Histogram



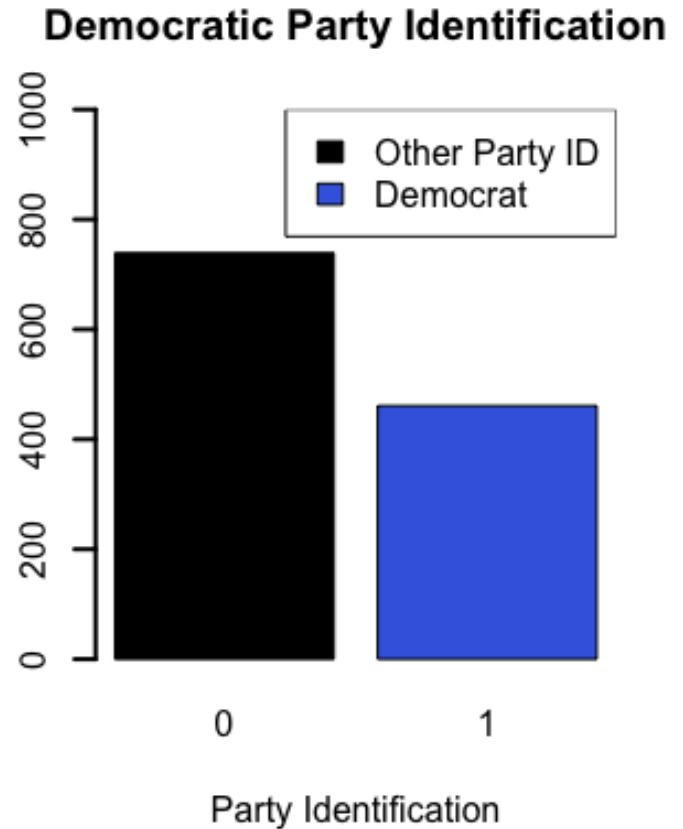
Graph 2: Feelings Toward Feminism Histogram



Graph 3: Obama Job Approval Histogram



Graph 4: Democratic Party Identification Bar Chart

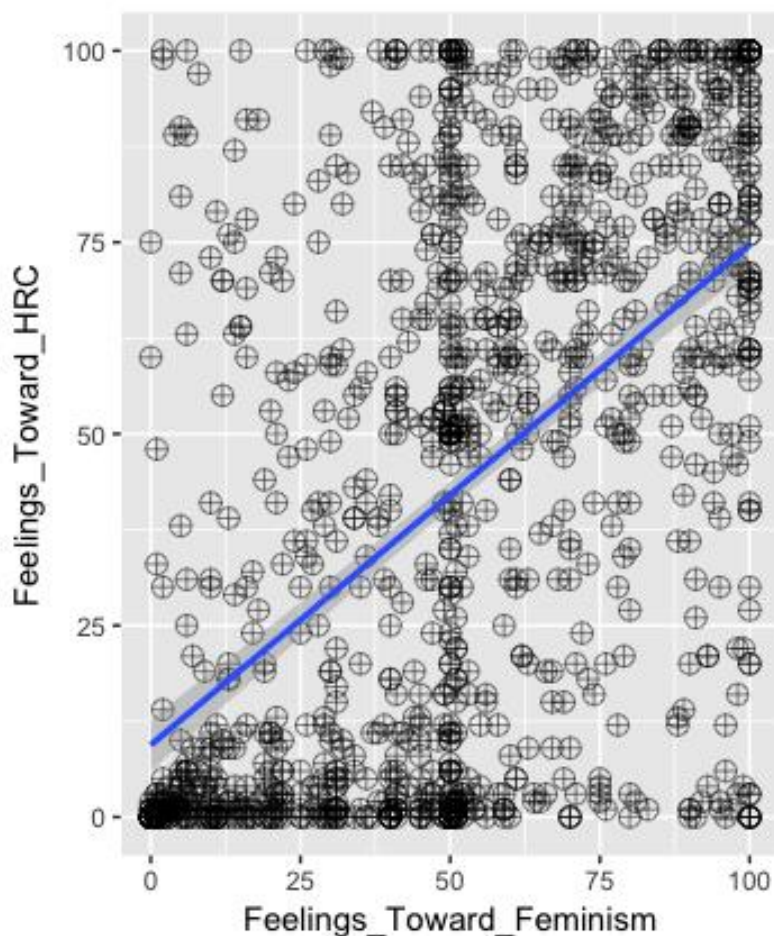


Graph 1 shows that indeed, the DV is relatively positively skewed as was expected from the summary table and thus is actually quite far from a normal curve. Graph 2 shows that the IV is a bit less normal than was expected based off of the summary table, as both the minimum and maximum are more frequent than most of the numbers in the middle, save for the very middle. Graph 3 shows that the Obama job approval control variable is also somewhat positively skewed, perhaps slightly more than the DV.

Bivariate Analysis

Analyzing the relationship between the experimental IV and DV reveals that there is a fairly strong, positive correlation between the two.

Graph 5: Feminism and HRC Bivariate Regression



This strong, positive correlation appears to affirm my hypothesis that the more positively the participants felt about feminism, the more positively they felt toward Hillary Clinton. But just how significant is this apparent relationship? The Pearson's R was less than $2.2e^{-16}$, which qualifies as being significant at the 0 level and is thus significant at the $p < .05$ level. The correlation coefficient given was also 0.5430713, also indicating that this is a rather strong correlation. This Pearson's R value essentially indicates that this relationship is incredibly significant in the form of a bivariate regression.

Due to the incredible strength of this correlation, it seems unlikely that the implementation of the control variables would completely cause the relationship pictured here to

disappear. What is more likely is that the control variables will cause the relationship to diminish, but probably not enough to make it significantly insignificant. So the odds appear quite favorable that the relationship will weaken, but overall will not drastically change or vanish.

Multivariate Regression

The two control variables that will be added in the multivariate regression model are President Obama's job approval rating and Democratic Party identification. The relationship between Obama's job approval rating will most likely influence the DV in a similar fashion as the experimental IV: the more they like the job President Obama did (the status quo), the more favorably they view Hillary Clinton. In other words, it is another positive correlational relationship. The hypothesized mechanism behind this relationship is that the way a participant feels toward the status quo will influence how they feel about Hillary Clinton because she campaigned as being an extension of the Obama era as well as tying herself to President Obama. This creates a close association between the two, which would affect how Clinton is perceived. For Democratic Party identification, those who identify as members of the Democratic Party are more likely to like Hillary Clinton. The mechanism behind this relationship is that people are more inclined to like her policies if they share the base issues from the Democratic platform and the issue of partisanship also would play a role in it. Much like both the experimental IV and other control IV, this too would have a positive relationship with the DV. The fact that both the control IVs share a positive relationship indicate a potential confound in the relationship between the experimental IV and DV, since they may make it look like there is a positive relationship where there otherwise is none.

The formula for a multiple regression with three independent variables:

$y = \alpha + \beta_1 X_i + \beta_2 Z_i + \beta_3 Z_i + u$. What this is showing is that the independent variable equals the y-intercept plus the experimental variable ($\beta_1 X_i$), the two control variables ($\beta_2 Z_i$ and $\beta_3 Z_i$), and the “random” part of the relationship (u). Essentially what this means is that β_1 is the slope between X and Y when controlling for the effects of Z, whereas β_2 and β_3 are the respective slopes for the relationships between Z and Y when controlling for the effects of X. So you add them together as well as the random part of the relationship in order to calculate the value of your DV, y.

Table 2: Bivariate and Multivariate Regression Models

Var	Model 1	Model 2
Feelings_Toward_Feminism	0.65254***	0.19417***
Obama		10.66956***
DemID		0.09985
N	1200	1200
R ²	0.2943	0.6249

In Model 1 the slope coefficient for Feelings_Toward_Feminism would have an effect of a clean 0.6254, because there are no interaction terms, so for each one unit increase of Feelings_Toward_Feminism there's a 0.65254 increase in Feelings_Toward_HRC. In Model 2 the slope coefficient for Feelings_Toward_Feminism dropped to 0.18727, or an increase of .18727 in the DV for every one unit increase of Feelings_Toward_Feminism while the Obama job approval rating control variable had a coefficient of 8.61050, or an increase of 8.6150 in the DV for every one unit increase in the Obama job approval rating variable, and DemID had a coefficient of 0.09985, which means there would be a 0.09985 increase in the DV for a one unit

increase in the DemID variable. The experimental IV and the Obama job approval IV were both statistically significant at the 0 level, however the Democratic ID independent variable proved to be completely insignificant.

In the population, it would mean that if you rate President Obama's job as being positive and feel positively toward feminism, you are more likely to view Hillary Clinton in a positive light with party affiliation not being a solid indicator. A person's feelings toward President Obama's job performance is a stronger indicator for a person's feelings toward Hillary Clinton than their feelings toward feminism overall. Because two of the IVs proved to be statistically significant, the null hypothesis for the Obama control variable is able to be rejected as it does have a significant impact on an individual's feelings toward Hillary Clinton. However, the null hypothesis about party affiliation is indeed true in that it does not positively influence a person's view toward Hillary Clinton.

R squared is essentially a measure of how close the data is to the regression line, which also explains how much of the variation is accounted for by the independent variable. In Model 1, approximately 29% of the variation can be accounted for by Feelings_Toward_Feminism. In Model 2, 0.6249 or 62.5% of the variability in the DV overall. However, for feelings toward feminism alone, the coefficient was reduced in Model 2 to about a third of what it was in Model 1, so it wouldn't make much sense for it to triple. Instead, in Model 2 the R2 value for that aspect alone would be approximately .085, or 8.5%. So the difference between the two models ended up creating a decrease in the experimental IV's influence, as had been predicted. While there was a lack of support that party ID influenced a positive view of Hillary Clinton, both of the other IVs managed to hold up and the support for the experimental IV was only slightly diminished in Model 2. However, there was no change in significance to be found between the

two models. What this minor change means is that when taking the variable of Obama's approval rating into consideration, the effect of Feelings_Toward_Feminism diminishes slightly as President Obama's approval rating had been contributing to the correlation coefficient in the bivariate model.

Conclusion

What this study served to reveal is that a person's feelings toward feminism was a decently important variable in the span of things, accounting for approximately 8% of the variability in the DV. However, it shows just how greatly other variables can mask the effects of others, as with the Obama variable not being accounted for the experimental IV was three times the amount it really was. And while the DV and experimental IV are undoubtedly correlated, it should be noted that the causal mechanism between the two could very well be caused by some extraneous variable that links the two of them that was not present to test within the ANES 2016 pilot study as the amount of variables was quite limited. However, based on the evidence presented in this study, the null hypothesis can be rejected as the relationship between feelings toward feminism and feelings toward Hillary Clinton shared a significant and positive relationship, even when factoring in other variables. It may perhaps be too strong of language to assert that positive feelings toward feminism caused positive feelings toward Hillary Clinton though, so it would be best to reword the theory into something more open and correlation based. The reason why this is important is to avoid asserting a causal relationship where there may not be one, which in this case would primarily be caused due to a lack of demographic data and other kinds outside of pure policy that could have perhaps influenced the relationship between feminism and Hillary Clinton. Future research should incorporate more variables to see how well this relationship actually holds up and whether or not it becomes insignificant statistically

with the introduction of some other variables. Future researchers could utilize this study as merely a starting point into exploring this kind of question with far more complex and accurate statistical analyses. While this was a fairly basic study, the relationships between Hillary Clinton and feminism, while not as strong as they were in the bivariate regression, are important nonetheless. The fact that President Obama's approval rating was a stronger indicator was quite interesting, though it made sense because of how heavily Clinton relied on marketing herself as an extension of his administration. What was most puzzling was that party ID had little effect, none significant, on the view toward Hillary Clinton.