**Russian Facebook Propaganda Analysis Proposal**

Donald Cooper, Paul Franklin, Jan Danel, and Tiger Hu

3/5/2020

**Introduction**

On February 16, 2018 Special Counsel Robert S. Mueller III indicted 13 Russian individuals and three Russian organizations for engaging in operations to interfere with U.S. political and electoral processes, including the 2016 presidential election. This was a significant step forward in exposing a surreptitious social media campaign and holding accountable those responsible for this attack. The indictment spells out in exhaustive detail the breadth and systematic nature of this conspiracy, dating back to 2014, as well as the multiple ways in which Russian actors misused online platforms to carry out their clandestine operations.

Our motivating question going into our own analysis of the Russian Propaganda involved the efficacy of advertisements by the Russian “Troll” Farm that targeted the American public through social media platforms. We originally intended to observe how effective the ads were as a political weapon leading up to the 2016 General Elections. The major factor we considered to help answer this question was the change in voting pattern due to the Russian propaganda. In order to analyze a change in voting patterns, we did a comparison from the election cycles during both 2016 and 2012 using the latter as the control group for the study. We also considered taking into account a few of the past presidential elections to build a time-series analysis. Considerations were also made as to whether data should be gathered on whether the IRA chose to create advertisements in prior years before the 2016 elections (i.e. our control year of 2012) as another baseline of comparison for ad effectiveness. However, our research question has since evolved as we found no significant evidence of a potential relationship between voter turnout and click through rate (i.e. effectiveness) across the country. This contradicts the findings of Spangher et al., which was an influence of ours going into this project[[1]](#footnote-0). We instead decided to ask whether we can predict the effectiveness of Russian ads given a combination of predictors like location, ad theme, and demographic information about the populations targeted among different states. We plan to measure effectiveness with Click Through Rate.

Spangher et al.’s *Analysis of Strategy and Spread of Russia-sponsored Content in the US in 2017* provides evidence that Russians targeted unregistered U.S. Citizens after the election. However, the literature did not seem to have evidence to whether the advertisements had any effect on unregistered U.S. Citizens before the election. Our model is intended to give an estimate on how effective the advertisements were on unregistered citizens and whether the ads provoked a significant percentage change in registered voters from unregistered voters from 2012 to 2016 (see Spangher p8). In addition to testing the effectiveness of the advertisement, we wanted to measure the effectiveness by location, preferably at the city or state level. This will help us see if some locations are more susceptible or were targeted more by the Russian ads.[[2]](#footnote-1)

**Data and Data Collection**

The data we selected to use was the same data collected and used during the The House Intelligence Committee Minority Investigation by Special Counsel Robert S. Mueller who indicted 13 Russian individuals and three Russian organizations, including the IRA, for engaging in operations to interfere with U.S. political and electoral processes such as the 2016 presidential election. Facebook released “a total 3,519 total advertisements [that] were identified to have been purchased” with over “11.4 million American users exposed to those advertisements”. The data has 25 different variables which include AdIDs, Adtext, Clicks, Impressions, Locations, CreationDate, EndDate, and AdSpend, all of which we use in our analysis[[3]](#footnote-2). In order to interpret and use most of the data we convert the variables in character or string format to categorical data. In addition to the 3,519 advertisements from Facebook made public by the House Intelligence Committee Minority, we have also gathered data from the U.S. Census Bureau and the U.S. Bureau of Labor Statistics that mostly has to do with demographics, voting statistics, and median household income measured by state. This additional data was selected by states in order to combine this data with the Facebook Ad data which we decided to break down to state-level, as we will explain.

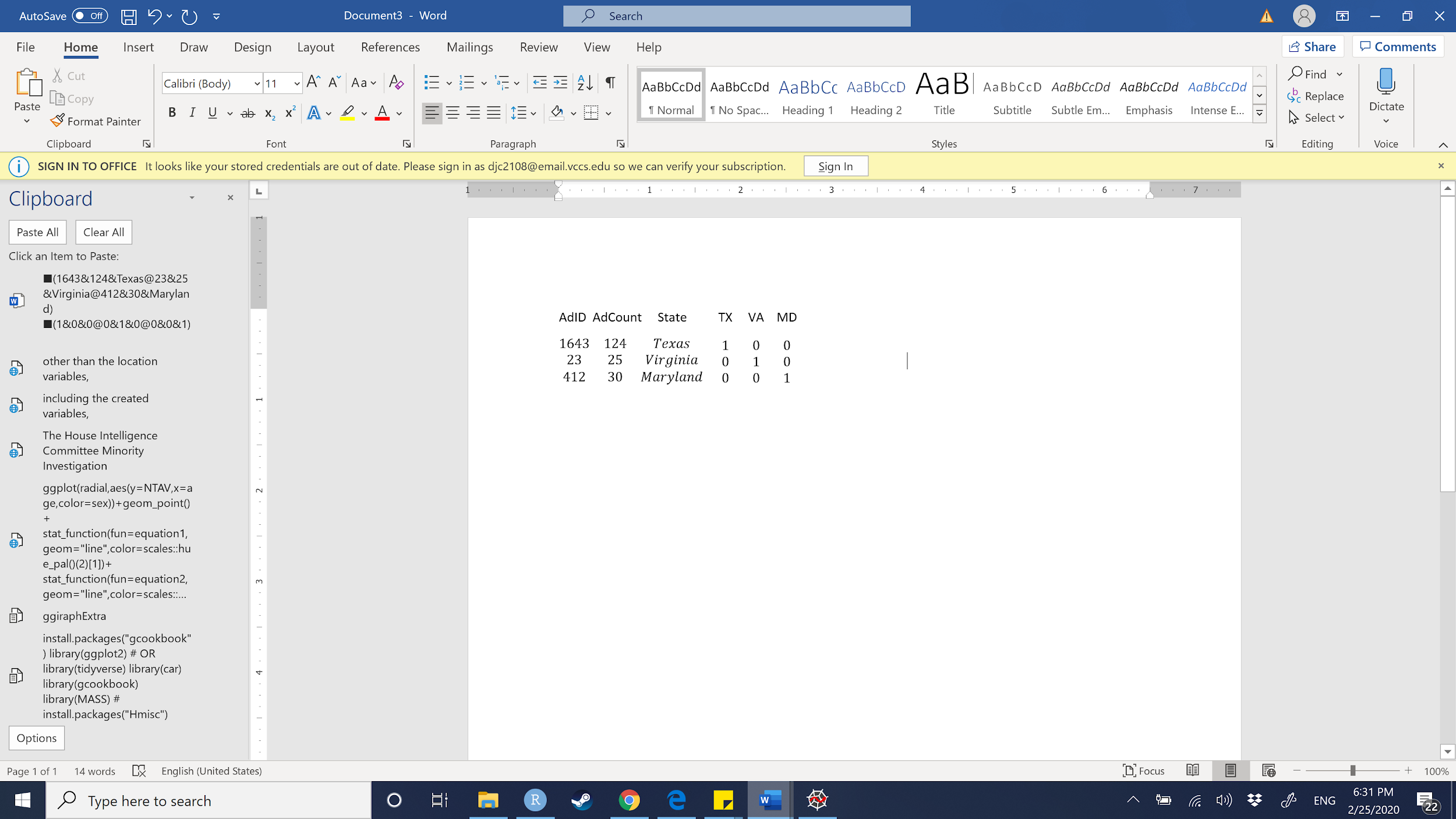
**Method of Analysis**

For the analysis, we needed to measure ad effectiveness on different locations/demographics. In other words, we needed dummy variables to differentiate ad effectiveness by each unique location. While there are many different location subcategories, such as town or city, we opted to use state level location parameters because of the abundance of data available through government websites such as the Census Bureau, Bureau of Labor Statistics, etc. Advertisements would then be categorized by the states they targeted. Some ads included only one individual state while others needed to be split from other coded values, such as having a city coded in next to the state name which identified that the ad targeted a particular city within that state. There were 982 entries that included a state level characteristic remaining after the initial cleaning. Other ads that were either from a different country or did not contain any state level location were removed from the dataset. Some ads happened to have multiple states that they had targeted. To account for ads with two or more targeted state level locations, the states were separated and advertisement IDs were counted the number of times it was included within a different state.

In addition to identifying which ad appeared in which state, a method needed to be developed in order to test the advertisement efficacy within each state individually for comparative purposes. For each state, a dummy variable was created to account for each location in which an ad had appeared. For the purposes of our model, the state dummy variable would account for when a specific ad had appeared or not for that respective state[[4]](#footnote-3).

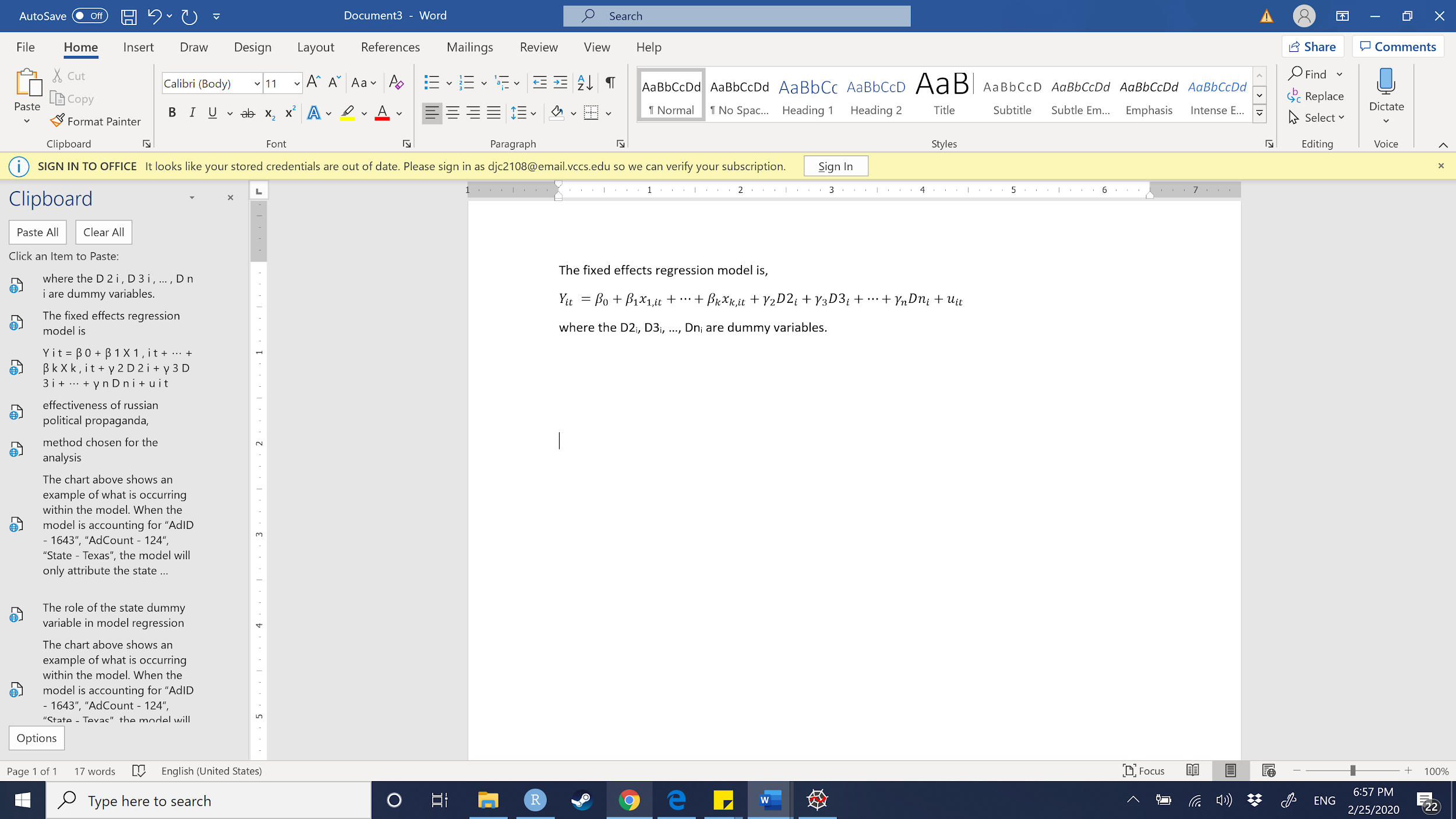
Chart 1: Example of the role of the state dummy

variable in model regression.



Many of the variables that were needed were created through coding the original Russian Ad Dataset, many of the variables were also kept. The original variables we included were Clicks (Clicks), or the number of clicks that Ad had received, and Impressions: “an impression is counted as the number of times an instance of an ad is on screen for the first time” (Impressions)[[5]](#footnote-4). The creation date (CreationDate) as well as the end date (EndDate) of the ad were kept in the dataset. For the advertisements that did not include an end date, we decided to add the end date of the ad as the date when the House Intelligence Committee Minority asked Facebook for the data. This happened sometime in January, therefore we opted to use the end date January 1st, 2017[[6]](#footnote-5). We decided to do this because we assumed that all the ads that Facebook provided to the House Intelligence Committee Minority were taken down no later than when Facebook turned the information over. The count of each advertisement (AdCount), as well as the cost for each ad (AdSpend), was included for each reported ad ID (AdID). The text (AdText) for each advertisement was also included. We created our response variable out of dividing the amount of clicks over the number of impressions for a click-through-rate (CTR). The CTR is the amount of clicks each ad had received in respect to the number of impressions it had garnered (i.e. number of clicks ad *i* received divided by the number of impressions ad *i* received). From both the creation and end date, we created a duration variable (AdDuration) that gave the time in which the advertisement was active in days. From the ad text variable, we created an ad word count (AdWordCount) variable that counts the number of words of that particular ad. Other than the location variables (Location), all of the original data, including the created variables, were quantitative variables. In addition to the original data, other outside data was gathered to boost the effectiveness of the analysis. State level data regarding voting statistics, demographics, and population data were merged with the current Russian Ad data. Most of the useful predictors from this gathered data were quantitative in nature.

It was agreed upon that using a fixed-effects multiple regression model[[7]](#footnote-6) was the best method for the analysis in examining the effectiveness of russian political propaganda by accounting for the many differences between the states while simultaneously measuring the effectiveness of ads. As discussed before, the formula for a fixed-effects regression model is as follows,



and will account for all states for individual comparison to a baseline. Essentially, the model categorizes indicators for each subject, in this case each state, in the model. This makes each state a categorical variable in order to test the significance of our other predictors of the model within each state. For example, the categorical variable listed for Virginia is “VA”, “MD” for Maryland, and continues through all the states included within the dataset including the District of Columbia (i.e. “DC”). Currently, we have discussed using Maryland as the baseline for our fixed-effects regression because Maryland was targeted by the Russian Ads the most. However, we have opted to use New York as our baseline because it is the median AdCount total of 85 different facebook ads. We concluded that using New York would offer a reasonable baseline for comparison because of the high concentration of ads as well the high concentration of population (StatePopulation). For example, a state could have been affected by ads more than New York or affected by ads less than New York based on the model prediction.

We will be running stepwise regression techniques to justify our choices in predictors. We will be calculating the Variance Inflation Factor (VIF) for all of our predictors as well. If we see large VIFs within most of our variables, we may be inclined to use several augmentation methods like Ridge and Lasso regression to specialize to our data. This would introduce bias to our data and may not be needed as we have few predictors. Evidence of multicollinearity among our predictors will allow us to proceed with a multi-fold cross validation to pick our tuning parameters for Ridge and Lasso. However, if we see small VIF’s throughout most of the predictors (i.e. only a few VIFs are large and little to know multicollinearity), this may not be a major concern and we would avoid Ridge and Lasso Regression. We also plan to conduct a Box-Cox transformation of the response in the event that our variance is nonconstant. We also ruled out some nonparametric statistical methods like regression splines and smoothing splines. Other methods were not suitable given our objectives for the following reasons.

**Supporting the Model**

We believe a multiple regression fixed effects model to be the best algorithm for answering our research question. This is so because we are trying to account and control for the differences across states in order to make them as similar as possible. So, when we test for the efficacy of the ads, we will get the result of how individual ads affected states differently or similarly, and if they had a greater or lesser impact in specific states. Multiple regression is also most relevant here because we have a continuous, quantitative response variable. This conclusion let us rule out classification techniques. Another advantage of multiple regression is that we can include categorical predictors. Including categorical variables (as dummies) will allow us to explain differences in CTR between locations. If we do not control for the differences between them we would be facing Omitted Variable Bias which could cause the interpretation of our result to be invalid. This happens due to the fact that our model could be significant but only because the states are so different that they return significant differences, and if we say that ad efficacy was the only reason for the difference between states then we would be greatly mistaken. In particular, outside research has identified multiple variables that correlate with the effectiveness of russian political propaganda, yet have not tried to predict effectiveness by combining all of these variables into a model. We believe that given new data, such a model could accurately predict the effectiveness of any newly detected advertisements (given some reliable predictors). This way, public officials and social media sites can narrow their search to the most effective types of advertisements or vulnerable populations, and hopefully be more successful at eliminating it in the future.

Since we had a response variable in mind from the beginning, we did not conduct Principal Component Analysis (PCA), which is unsupervised. This method would be informative if we wish to consider other response variables in the future, but given time constraints and our desire to discern which ads were most effective, we skipped PCA. Regression decision trees were ruled out as they are generally less accurate in predicting quantitative response variables than multiple linear regression. Still, for the future a tree could let us compare the importance of one of our IVs to every other IV. A regression decision tree would also be advantageous if we wanted an algorithm that would produce an easily interpretable tree diagram. Such a diagram would be a good visual if we’re fortunate enough to report our findings to the public. In the future, we might also consider a hierarchical or K-means clustering technique to identify subgroups of people most susceptible to ads. This could also be a valuable way to narrow the search for propaganda in the future, as the IRA will most likely try to exploit those subgroups it successfully influenced in the time frame of our data.

**Potential Objections**

A possible objection to how we manipulated our data would involve the end date of the advertisements that were unknown, and our assumption that they were taken down on the day that Facebook gave the House Committee the report. In other words, we assumed that the end date for all missing “end date” values was the last day Facebook compiled the ads before giving it to the House Committee.

We could not assume normality of individual ads concerning their distribution of clicks and impressions across different states which is why we created individual rows for a single ad which had multiple locations in order to have ad information by individual states. This could be thought of as manufacturing points since we could be counting multiple variables more than once. A problem with this is that we have a certain number of clicks for a specific ad and we cannot assume homogeneity across the states; citizens of one state may react differently when seeing the same ad that another citizen from a different state causing clicks to also be different between the states. The way we addressed this problem was by creating the Click Through Rate of the ad which is a rate and thus should be the same across heterogeneous states.

Not all of our regressions to date include ads that appeared nationally. These ads could have appeared in any of the states for which we have dummy variables, but we’re unable to train the dummies on those clicks since we do not know where clicks of a national-level ad occurred locally. In this way, our dummy variables may not represent the true change in effectiveness when a state is targeted.

For comparing the effectiveness of the Russian Facebook ads, we wanted to gather either not political or simply generic ad data from facebook between the years of 2015 and 2017. Unfortunately, Facebook does not publish ad statistics from before May of 2018. Thus, we could not get an accurate date range for ads during this period that were not from the IRA. We have decided to use ad statistics from May 2018 to May 2019 and calculate the effectiveness of these ads to compare with the 2015 and 2017. This comparison may not be as significant due to various changes in the confounding factors we are measuring between the 2015-2017 period and the period of 2018-2019. This happens to be the only date range of non-political ads that we could use to compare with the Russian ads are the 2018-2019 ads. So we will use these ads as a control group.

States are very heterogeneous which makes running regression between them difficult due to multiple variables that could be omitted. The reason why we are conducting a fixed effects model is so we can account for this and control for the major differences between states such as politics, demographics, income, and population size.

1. It should be noted that while we did not find any significance in voter turnout based on the data. Others, such as Spangher et al., may have found significance using different predictors or other statistical modeling techniques that we have not considered, nor have the time to implement. This could be a follow-up task within the parameters of our analysis if given the opportunity to extend our period of data analysis. [↑](#footnote-ref-0)
2. Others, notably Dutt have measured effectiveness at the state level using ad language. However, we hope to measure it at the city, state, and national level. Similar to Dutt, we might consider conducting many Wilcoxon tests with Bonferroni adjustments or bagging.) [↑](#footnote-ref-1)
3. Note that these variables were renamed from the original dataset to fit our needs in R-Studio. [↑](#footnote-ref-2)
4. Chart 1 on the following page shows an example of what is occurring within the model. When the model is accounting for variables “AdID” = 1643, “AdCount” =124, “State” = Texas, the model will only attribute the state dummy created for Texas (i.e. TX in the model) with “AdID” = 1643, “AdCount” = 124, and “State” = Texas. This works for all other states. [↑](#footnote-ref-3)
5. <https://www.facebook.com/business/help/675615482516035> [↑](#footnote-ref-4)
6. We believe that choosing this date in January would not be significantly different from choosing a random date within this month when running this variable in our model. Many of the ads that we chose to have this end date already exceed triple digits in days, so adding at most 30 days will not significantly change the predictive model. [↑](#footnote-ref-5)
7. Having individual specific intercepts *𝛂i*, *i = 1,...,n,* where each of these can be understood as the fixed effect of entity *i,* the model will have fixed effects on the response variable. For the formula above, *𝛂i* are entity-specific intercepts that capture heterogeneities across entities and are denoted as 𝜸*nDni* (Econometrics with R, 10.3). [↑](#footnote-ref-6)