Effect of Ideological Difference on Campaign Contribution Amount

Hursh Desai

New York University Leonard N. Stern School of Business

Abstract

For campaign contributors contributing to a candidate's campaign, does more distance on the ideological scale between the contributor and recipient affect contribution amount? This is an important question when it comes to Campaign Contribution Theory, how individuals and corporations decide to give their money to politicians' campaigns. The results of this study show that more contributions happen at smaller distances between contributors and recipients, however individuals tend to give higher **amounts** of money to candidates the further away that candidate is from them in ideological score.

Purpose of Research

The results of this research will advance the field of Campaign Contribution Theory. There has already been quite a bit of research done to validate Voting Theory however, equally as important if not more so is how and why people donate to campaigns. Some scholars believe that contribution amounts can determine the results of the election. If this is true then analyzing the reasons why an individual would contribute more money to a candidate is a worthwhile endeavor. According to the executive director of the Center for Responsive Politics, Stephanie Krumhols, "more than 90 percent of candidates who spend the most win." (Koerth, 2018) However, Adam Bonica and Richard Lau, a professor of political science at Rutgers, both agree that:

> "the strong raw association between raising the most cash and winning probably has more to do with big donors who can tell that one candidate is more likely to win — and then they give that person all their money."(Koerth, 2018)

We are already seeing a variable other than ideological score that is dictating how much money a person might donate to a campaign. However, we will analyze confounding variables and whether or not this is one later on in this paper.

Studies have already shown that contributors place a great importance in the ideology of the candidate who they are donating to, when they contribute to a campaign. There is however, a split when it comes to contributor type. PACs and Interest groups did not discriminate which campaign they would donate to based on the legislator's ideology whereas this was the most important point when it came to how individuals would donate. (Barber, Michael. 2016) However, how does that play out in the amount they are donating. Does it make a marked difference in the amount of money that they donate to different campaigns.

The topic of this research also has value for election prediction and for politicians/campaign managers. The results of this research could show predictive power of difference in ideological scores between the candidate and the electorate in which they are running, on amount donated or effectively amount of money raised. If this is true that difference could become another variable to keep in mind prior to a campaign raising money when trying to see the viability of success for a candidate in a certain race. Politicians themselves as well as their campaign managers would find this dynamic important to include when doing their analysis.

Data Discussion

The data I used comes from a database put together by Adam Bonica at Stanford called the DIME (Database on Ideology, Money in Politics, and Elections) database. He collected this data from the FEC official website because the FEC requires some personal information to be disclosed on any contribution greater than $200. After compiling and cleaning the raw data, the

database now houses over 130 million political transaction records. The most important aspect of this specific aggregation of data is that Adam Bonica has developed a method to calculate the political ideology of both the contributors and receipts through just the contribution records. Though the specific mathematical formulation done to come up with these CF scores is outside of the scope of this paper, the methodology has been put through a number of internal and external validity tests to ensure that these numbers are reliable. (Bonica, Adam, 2019) This method has even been shown to be able to predict roll call scores using supervised learning. (Bonica, Adam, 2018) They have been shown to be highly correlated with DW-Nominate scores which are highly regarded when it comes to ideological scoring. DW-NOMINATE scores have been used widely to describe the political ideology of political actors, political parties and political institutions. (Poole Keith, 2005)
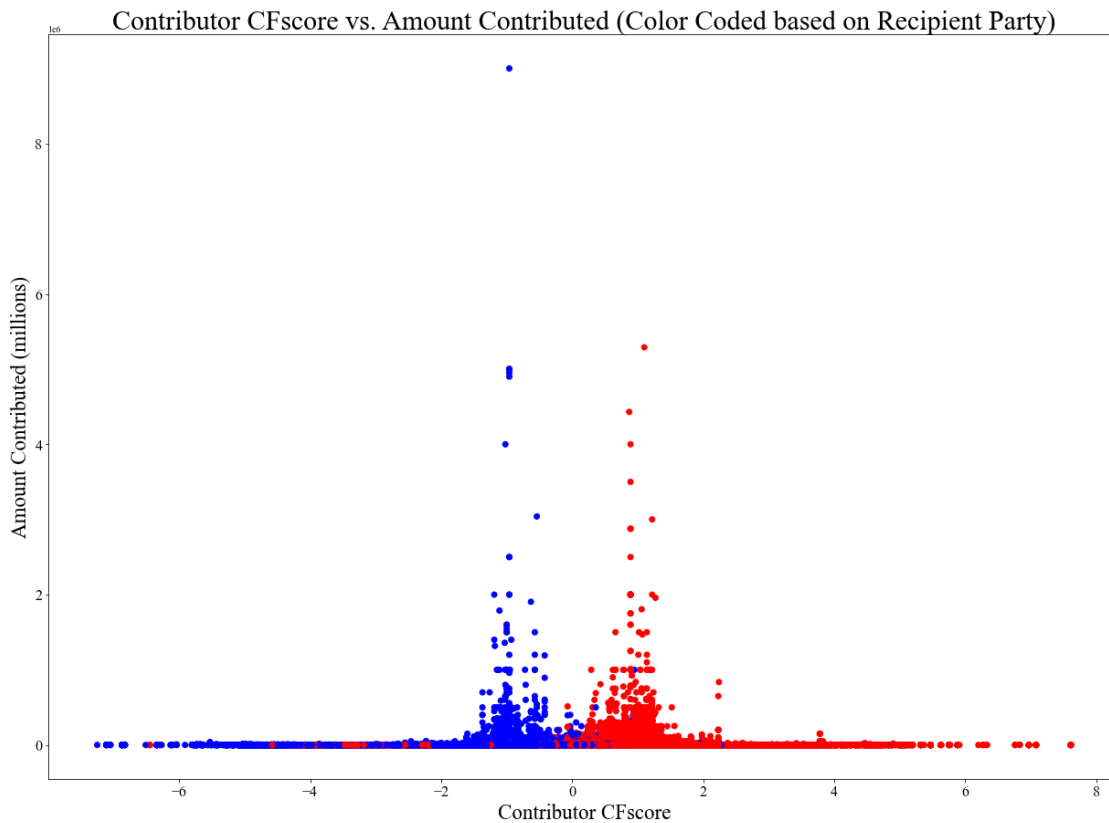

Figure 1

Even from this initial visualization in Figure 1 of contributor CFscore against the amount contributed we can see that the method used to calculate CF score is dividing the data set almost perfectly along party lines. There also appears to be a significant increase in the amount contributed at more neutral ideological scores, seemingly suggesting that the less extreme your political views are the more likely you are to donate larger amounts of money.
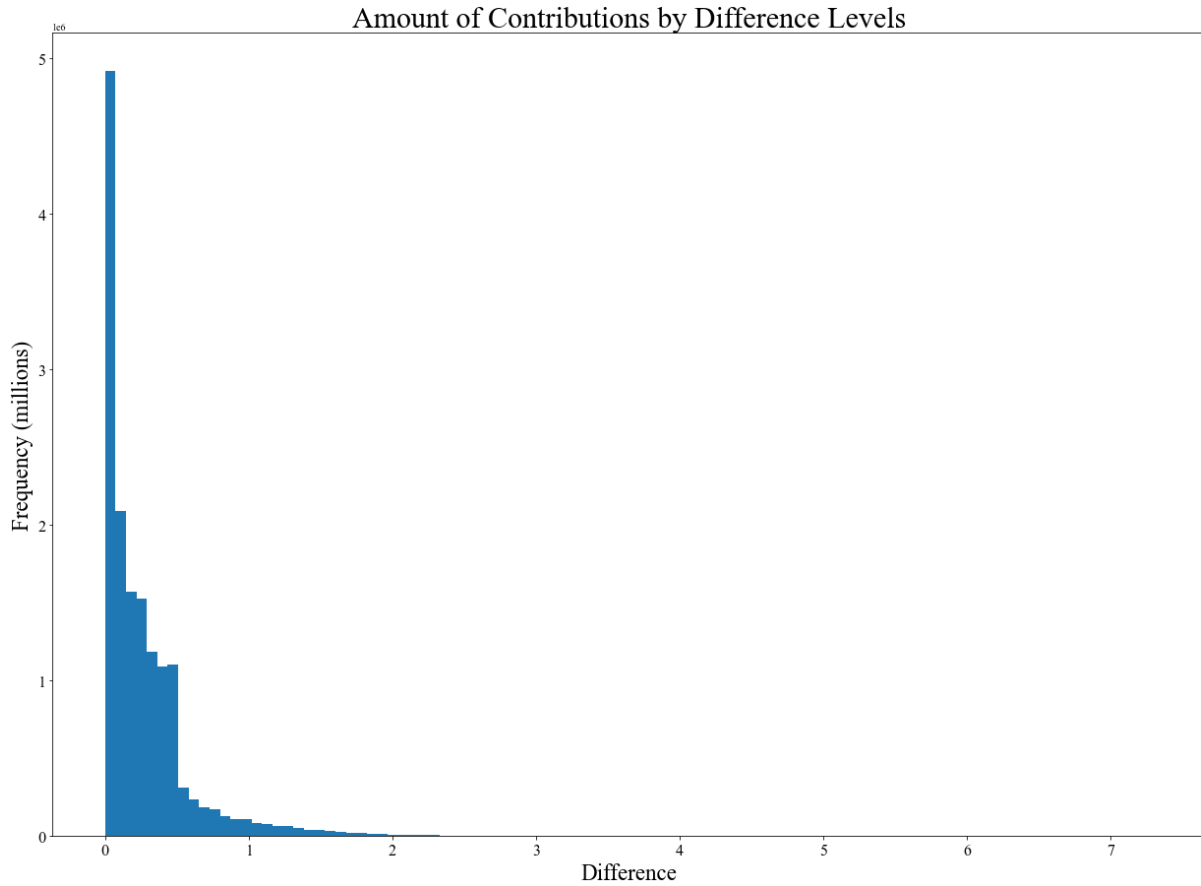

Figure 2

Even a preliminary look at the frequency of data in Figure 2 shows that a bulk of the contributions being made happen at lower amounts. In fact 93% of contributions made are within those first 10 bins (0 - 0.518 differences).
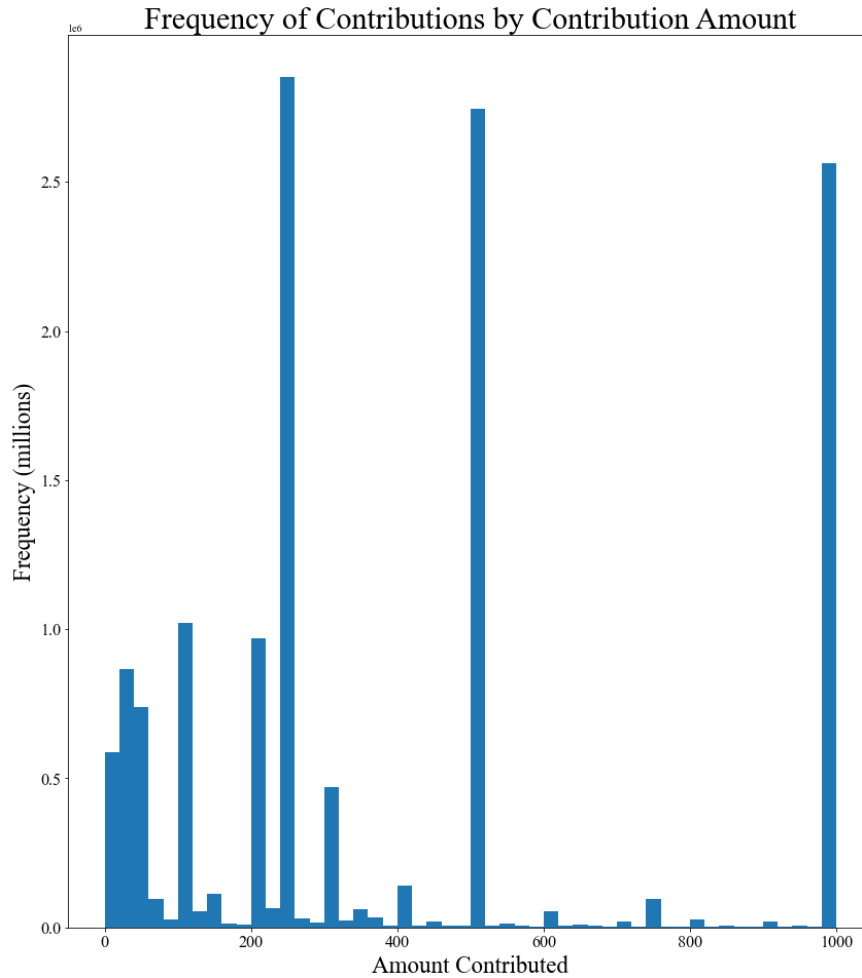
Frequency of Contributions by Contribution Amount

Figure 3

In Figure 3 we have a histogram that shows the frequency of contribution amounts under $1000 which is where 91% of the data is coming from. Quick analysis of the graph suggests that there is abnormal amount of contributions being made at what seems to be around $250, $500, and $1000.One hypothesis is that campaign websites give those optional amounts to donate at the start which anchors people to those amounts. Additionally,these numbers are also round numbers which could attract people to those amounts.
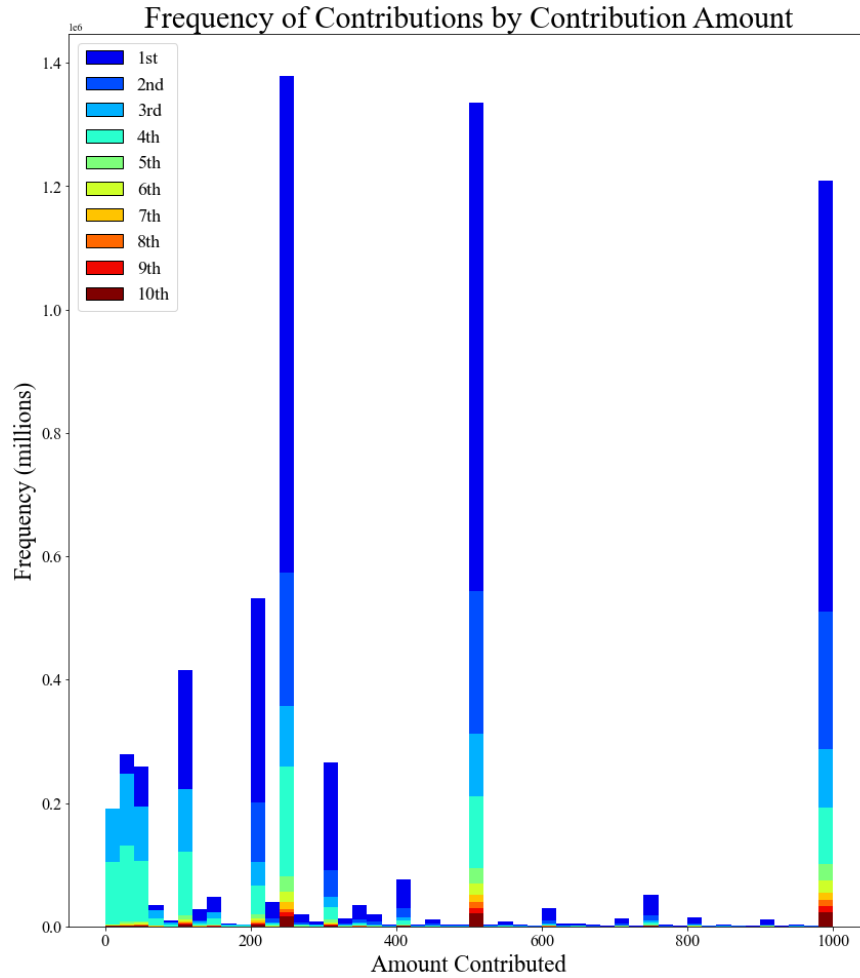
Figure 4

From Figure 4, there does not appear to be a significant difference between different ideological differences bands. Using the first 10 bins given by the Histogram in Figure 2 I constructed a stacked histogram to see if any certain band of ideological difference contributed in a higher frequency than the others. However, we find no evidence of that in this data.

The most difficult part of working with this data and running it through my regression model is the level to which it is concentrated within an unusually small range. 93% of the contributions were made within a 0.5 range of difference in ideological score and 91% of the contributions were made under $1000.

Data Cleaning

A substantial portion of the data cleaning came from analyzing whether the data was relevant to the regression. There were many amounts contributed that were coded as negative amounts - which represented campaign contributions that were refunded over this period of time. I removed those amounts since they did not add to the understanding of the relationship being tested. Due to the large size of the data file and my limited computing capability, I also made the decision to pare down the rows to only include those that would be most useful to the regression. Since a lot of the variables had categories that contributed very little to the regression, I began to take out categories that had few transactions under them. For example, I removed all states that had under 10,000 transactions which was 48,288 transactions out of around 15.28 million, still retaining 99.68% of the transactions. I also only included contributions that were being made to the two dominant parties, Republican and Democrat, for similar reasons.

I only included contributions that were made during primary elections to give more variation on the recipient side of the transaction. In addition, when I ran a regression using only general election data, the p-value for the difference variable was 0.309, indicating that the beta coefficient -6.12 wasn't statistically significant and no conclusion could be drawn for this relationship in the case of the general election since that relationship could be found in completely random data as well. This might be because

> "If you focus on general elections, he said, your view is going to be obscured by the fact that 80 to 90 percent of congressional races have outcomes that are effectively predetermined by the district's partisan makeup."(Koerth, 2018)

Such an influential confounding variable could affect the relationship between difference in ideological score and contribution amount as well. If a district's partisan makeup has such a

strong influence on the outcomes of elections then it could also have an influence with similar strength on who is donating and how much they are donating.

Lastly, of course I removed all rows of data that had any nulls or were left empty because it would render the regression useless. By the end, I was left with about 15.23 million rows of contribution data to work with which was more than enough.

Methodology

Clearly it would be difficult to find counterfactuals for a person donating to a campaign to make a causal connection between the amount of money they contribute and the candidate's ideological score. Even if we found a significant enough number of people who had donated to multiple campaigns to act as a counterfactual, we couldn't be sure that they changed their donation amount based on the fact that they are donating to multiple campaigns. There is no perfect counterfactual so I have assumed that people with the same difference in ideological score between themselves and the candidate they are contributing to, will have similar donation tendencies over election cycles.

Some unobservable confounding variables that could bias the relationship are increasing polarization of the electorate over time. This would cause people to only contribute to candidates who are very close to their ideological score and due to the increased intensity that comes with polarization it could very well cause people to donate in higher contribution amounts. Figures 1 & 2 give us a somewhat muddled view on this dynamic since it appears that people with relatively neutral CFscores tend to donate more money and yet the smaller the difference in ideological score between contributor and recipient the higher the contribution amount. However, this turns out to have a simple answer.

	Q1	Median	Q3	Mean
Candidate CFscore	-1.12	-0.42	0.95	-0.11
Contributor CFscore	-1.18	-0.28	0.97	-0.14

Figure 5

The distribution of CFscores for both the contributor as well as the candidate are almost identical. Even if the electorate was particularly fervent in their political beliefs, since both them and the candidates they are donating to are so similar in ideology, it wouldn't lead to a large differential in ideological score. Regardless this effect would not play a big part in this relationship since the number of large amount donations made were very few.

Similar dynamics that are at play in voting strategy could play out campaign financing as well, so that one might donate to a candidate slightly farther away from their ideological score in the hopes that this candidate is more likely to beat the candidate from the opposing party in the general election. This is because even though a strategy like this shouldn't necessarily affect the amount contributed, a significant amount of non uniformly distributed increase in the number of contributions could skew a regression model testing contribution amount. In general, many of the confounding variables that can be found during the analysis of how and why voters vote the way they do can be found at play in the analysis of how and why voters contribute to campaigns the way they do. Previously, I mentioned that big donors only donate to candidates that they think are going to win could possibly be a confounding variable. However since it doesn't affect the difference in ideological score and the possibility of big donors having a significant impact in the number of contributions is limited, it is unlikely that it is a confounding variable.

BecauseI am doing this study over time, these are not the same people or even cohort of people participating in the study throughout the election cycles. However, it would be too computationally expensive to include fixed effects for all 15 million people throughout the years,

so I have included fixed effects for certain attributes that could code for these individuals/cohorts over time such as gender, election cycle, contributor type, and contributor state.

Results

First and foremost, even after putting in all of these variables the R-squared remained at only 0.001. This means that none of the variables included in the regression contributed significantly towards the amount contributed. There is about 99.9% of the effect on contribution amounts that is not being taken into account by this regression. Confounding variables, or dynamics such as the previously mentioned big donors donating to candidates they think are going to win, seem to have played a heavy role in this relationship. This result was somewhat surprising seeing how we already know that the ideological stance of the candidate is of great importance to those contributing to campaigns. It would seem to follow that the amount donated to these campaigns should be largely influenced by how ideologically different the candidate they are donating to is from them. However, it appears that there are many other variables that dictate the amount of money that one contributes. This is where the confounding variable of income most probably comes into play. Although finding any level of income data, that would add to the regression, for the individuals in this dataset proved to be quite difficult, had I included a fixed effect for the income level I am certain that it would increase the quality of the regression significantly. Even if the income level of the contributor had a slight relationship with the difference in ideological score, the relationship between income level and amount contributed would have been very strong and so the resulting end effect would have been significant.

However, there is a relationship between the difference in contributor and recipient CF score. The p-value for the difference variable in this regression is *0.000* which means it is

statistically significant. It has a beta coefficient of 34.32, which is also surprising given my

hypothesis. This beta coefficient should be interpreted as, all else being equal, each one unit

increase in CF score difference between contributor and receipt is associated with a $34.32

increase in contribution amount. I expected that the smaller the difference the larger the

contribution amount should be (negative relationship) or maybe even no relationship, however I

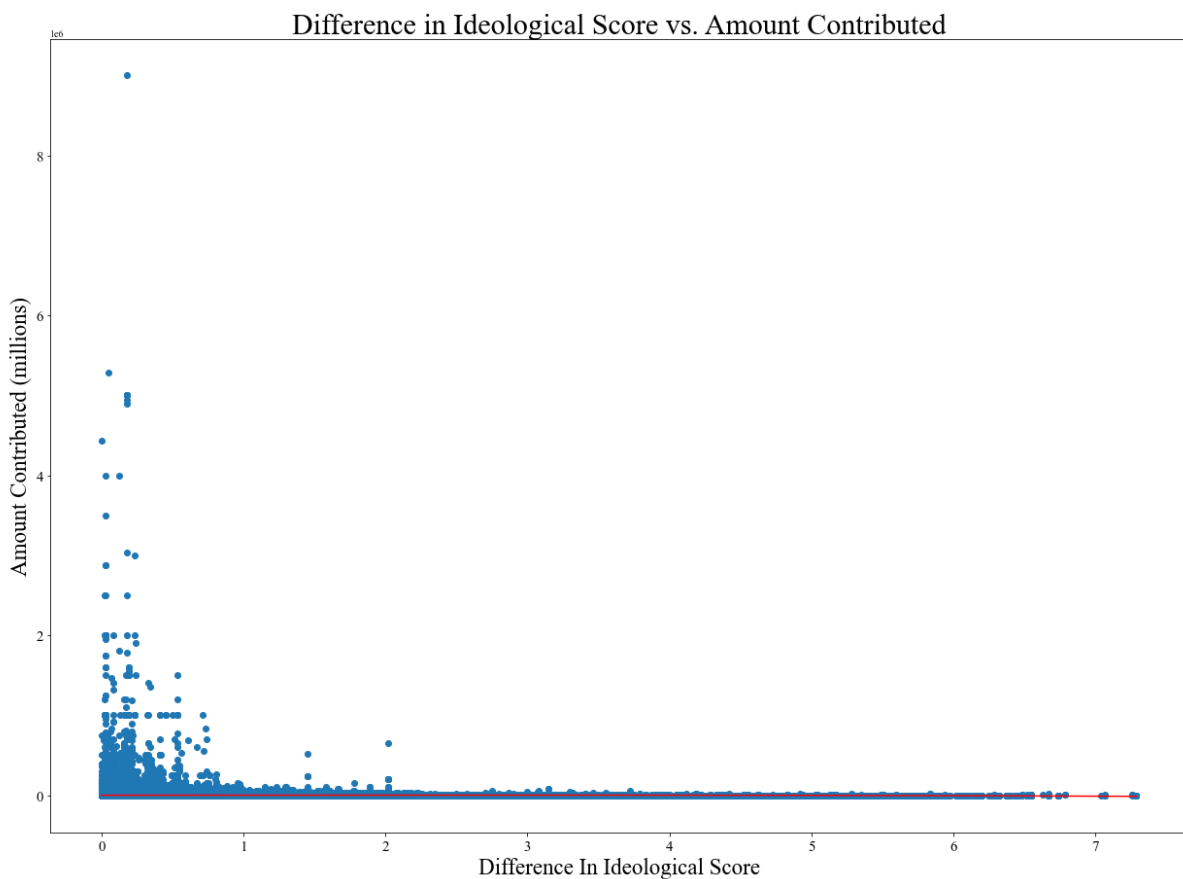did not expect there to be a positive relationship between the two variables.

Figure 6

My preliminary visualizations at first made these results quite confusing, because visually

there does seem to be evidence of a negative relationship between contribution amount and

difference in CF score. However, plotting the line of best fit gives what seems to be a straight

line at the very bottom of the plot. However, I created a histogram and then because most of the data is in essentially the first 7 bins I created boxplots for the data within those bins. This revealed to me that the sheer volume of contributions made below $1000 is what is actually driving the regression. Although visually there seems to be a negative relationship, this visualization is misleading because the impact of millions of rows of data coming in within such a small space cannot be accurately visualized so readily. Through the mean given by the box plots we can see that all of the data points that on the naive visualization seem to create a negative correlation are actually all outliers. The relationship that the regression is essentially picking up on can be seen through the means given by these box plots.
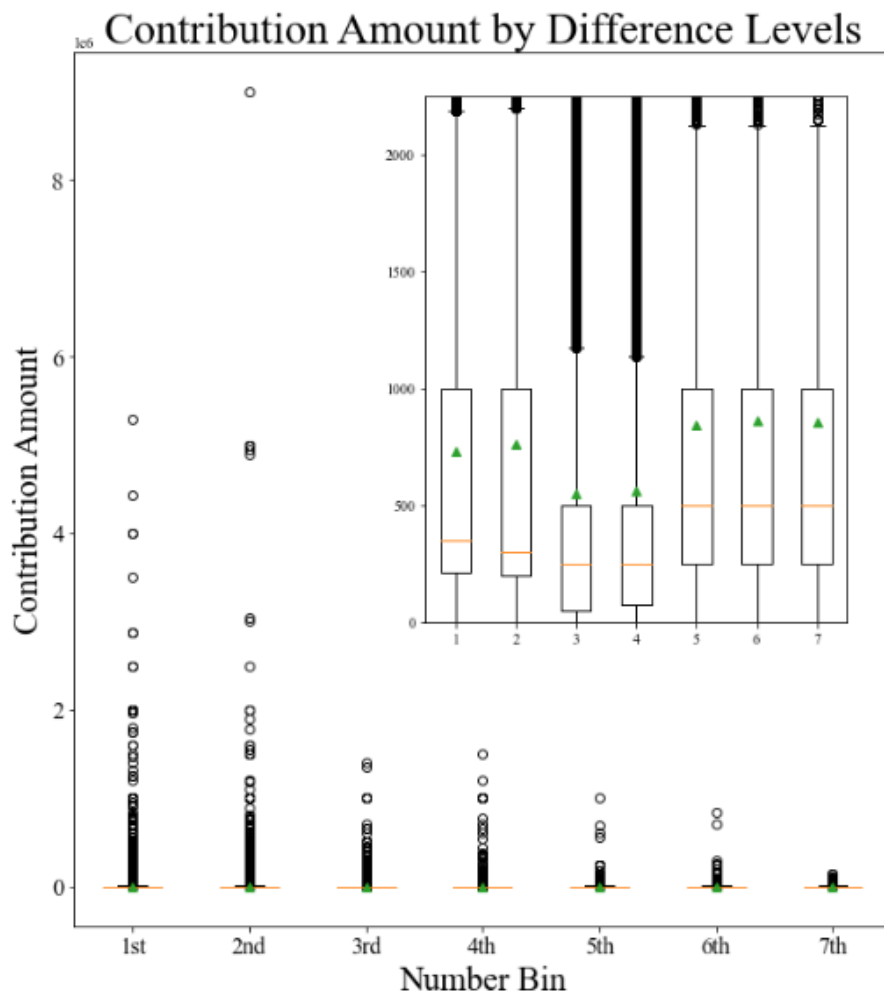

Figure 7

As we can see, the average is essentially steadily rising as the difference increases. Although the first bin has the greatest outliers, it also has the greatest volume of low amount contributions, causing the average contribution size of that bin to be one of the lowest of the first seven bins. The best explanation for the line of best fit in Figure 6 is that it follows the averages indicated by the green triangles in Figure 7, but without zooming into those particular data points it looks like a simple straight line.

The Durbin-Watson test is at 1.838 meaning the autocorrelation at lag 1 is on the low end of the spectrum ranging from 0 to 4. Plotting the residual values vs fitted values was almost just as useless as the original scatter plot because almost all of the analysis you would be able to see is of the outliers and not the smaller amount contributions where the surfeit of data was. Doing a robustness check by taking out the fixed effect of the contributor state revealed that the difference in ideological score is still statistically significant but now with a beta coefficient of 55.07. This is still similar enough to the original regression including a state fixed effect that we can consider this regression robust to the difference variable. Adding in an interaction effect between election cycle and difference in ideological score did not yield a statistically significant result. Doing the regression without any of the fixed effects and then another one with the year fixed effect gives statistical significance to the difference variable and a beta coefficient of 19.11 and 48.64 respectively. Adding in the time varying variable also increased the regression's R-squared from 0 to 0.001. We can say from this that including a time varying fixed effect of the election cycle might actually account for a lot of the variance that we were witnessing in the contribution amount. Further research will be necessary to determine the internal validity of this relationship between ideological difference and contribution amount, is this increase in

contribution amount as ideological score increases actually do to the variance in ideological score.

<u>Conclusions</u>

There is a positive relationship between the difference in ideological score between the contributor and recipient and the amount of money contributed. This relationship is derived almost solely from the millions of contributions coming in under $1000. The relationship itself may seem slight with a one point ideological difference associated with a $47.17 increase in donation amount. However, this is quite an increase seeing how the amount of money averages about $703. Even though this relationship was quite an interesting one considering I had hypothesized the opposite, the few insights I found along the way were equally as curious. Such as the sheer volume of small donations, making any donation outside of $1000 an outlier since 91% of the donations that were made were under that number. Another being that the greatest number of donations come from contributors with less than a -0.5 difference in ideological score between them and the candidate they are donating to. It is curious to see the dichotomy between amount contributed and contribution amount. It seems that for some reason people contribute more to candidates closer to them in ideology and yet contribute more **money** to people further away from them in ideology. I cannot even come up with a theoretical interpretation of these results, so I conclude that there must be some confounding variable I have not taken into account in this regression that is causing millions of individuals to donate to candidates that are quite distant from them on the ideological scale.

References

Barber, Michael. "Donation motivations: Testing theories of access and ideology." *Political Research Quarterly* 69.1 (2016): 148-159.

Bonica, Adam. "Mapping the ideological marketplace." *American Journal of Political Science* 58.2 (2014): 367-386.

Bonica, Adam. "Inferring Roll-Call Scores from Campaign Contributions Using Supervised Machine Learning." *American Journal of Political Science* 62.4 (2018): 830-848.

Bonica, Adam. "Are donation-based measures of ideology valid predictors of individual-level policy preferences?." *The Journal of Politics* 81.1 (2019): 327-333.

Koerth, Maggie. "How Money Affects Elections." FiveThirtyEight. FiveThirtyEight, September 10, 2018.

https://fivethirtyeight.com/features/money-and-elections-a-complicated-love-story/.

Poole, Keith T. *Spatial models of parliamentary voting*. Cambridge University Press, 2005.