Predicting Market Risk for Sectors During Presidential Elections

Alexander Sahinidis

Economics and Data Science (73-265)

December 20, 2020

Professor Portnykh

Introduction

In 2020, the world was faced with a crippling pandemic, COVID-19. Almost 100 million people have acquired this life-threatening virus, and almost two million people have died from it. National lockdown enforcements have been put into place by many countries, and millions of workers have lost jobs. Ultimately, the world is witnessing a major disruption to routine life as we have seen it, which is also causing severe economic distress. This has been evident in the stock market, where the Dow Jones Industrial Average fell over 30% in March 2020. Since then, the market has been on a rise, but has proved extremely volatile. News of the pandemic has dominated the psychology of investors, and therefore the prevalence of COVID-19 has created many rare market returns over space and time in the market. With some stocks going up 10 or 20 percent a day, others go down 10 or 20 percent a day. Although the market has seen some recovery since the 2020 market crash, volatility in the stock market is as prevalent as never before.

The year 2020 has seen many lows due to the pandemic, and a sense of normalcy seems far away. Although, for the past 200 years, the American people have witnessed the United States presidential election every four years, this year's was much different than those of the past. Since September, the presidential election's presence has dominated the media. Because the pandemic has emphasized an unusual aspect of strong candidacy for frontrunners in the election, the pandemic's effect on the economy and the presidential election seems to be closely linked.

Perhaps, the link between the economy and the presidential election is not limited to this year only. By examining data of the past, we might be able to make a prediction about the presidential elections and the stock market for the future. Due to the volatility and unpredictable

recent behaviour of the stock market this year, this paper focuses on the link between market risk within the eleven sectors of the economy and presidential election party winners.

Methods and Procedures

Overview

This paper focuses on the market risk of the eleven sectors of the economy: consumer goods, real estate, telecommunications, utilities, technology, industrial, financial, energy, consumer discretionary, health care, and materials. The Greek beta measures the market risk, which is defined as the connection between the returns of an asset and the returns of the market. Therefore, it is imperative to wield a tool that allows us to quantitatively assess the market risk of a given asset. Beta does exactly that. We will use the beta calculations to quantitatively measure the market risk of the eleven sectors of the economy, and compare these values to make conclusions about our data within the election timeframes that we choose.

Data

There are two main datasets used: presidential election data that gives information on individual county voting for each election from 2000 to 2016, and market and sector returns from 2000 to 2016. All of the calculations will work with only these two, but large data sets. The presidential election data is from the MIT Election Data and Science Lab. This dataset gives information on county-level returns, such as the total number of votes the candidate voted for, the party voted for, and so on. The data ranges from 2000 to 2016, or, from George W. Bush to Donal Trump. The data on the market returns is computed from the S&P 500 Index, a major and prominent indicator of the stock market. The data on each individual market sector is computed from the highest total asset exchange traded fund (ETF) for the relevant sector. For instance, the indicator for the consumer goods sector of the economy is IYK, the iShares U.S. Consumer

Goods ETF which has the highest amount of assets within its fund. Similarly, each other market sector follows the same guidelines: an ETF with the highest total assets with data that ranges from 2000 to 2016. The range of the data is chosen to be from 2000 to 2016, because 2000 is the earliest date for which the currently significant ETFs existed.

Calculations

As mentioned, beta is the indicator that will be used to quantitatively determine the market risk for a given sector of the economy. Each ETF tracked from 2000 to 2016 will have a corresponding beta calculation. The formula used to calculate the beta is the following:

$$\beta = \frac{Cov(r_i r_m)}{Var(r_m)}$$

where r_i is the expected return on an asset i, r_m is the expected rate of return on the market, *Cov* is the covariance, and *Var* is the variance. This single formula is the most relevant calculation done. In this case, i is the ETF analyzed for each market sector, and m is the S&P 500 Index. To interpret each beta, consider the following. If beta = 1, then if the market is expected to increase by 10%, so will the asset that is being studied. If beta = 1.10, then if the market is expected to increase by 10%, the asset is expected to increase by 11%. If beta = -1.10, then if the market is expected to increase by 10%, the asset is expected to *decrease* by 11%. This is the information that beta tells us, and comparing the different betas for each sector can tell an interesting story.

Data Analysis

To configure and make calculations with the data, a statistical computing and graphics programming language, R, will be used. Doing so will allow visualizations of the data configurations, in addition to the data wrangling applied to result in more informative data. Packages used include *reactable*, *tidyverse*, *readxl*, *ggplot2*, *ggthemes*, *gridExtra*, *sparkline*, *dplyr*, and *car*. The *reactable* package produces visually appealing tables and are "reactable" with the mouse pointer. The *ggplot2* package produces all of the graphs, and therefore is extremely important to visually analyze the eleven market sectors. The *car* package is used to apply the linear regression models and perform backtesting by assessing the linearirity, exogeneity, and residuals of each model. Thus, linear regression models will be used to allow conclusions to be made when comparing two or more variables. The linearity of these models will be assessed to ensure a thoroughness and robustness of each model.

Procedure

The results of the project will be separated into three sections: deriving which sectors are most politically significant, the effect of landslide election victories, and pre and post election market differences.

Political Significance

Political significance is a measure of the difference between the market risk of a market sector when a Democratic or a Republican candidate wins the election. In the data, 2000, 2004, and 2016 are Republican winners. 2008 and 2012 are Democratic winners. A beta is calculated for every election year, and an average beta is calculated for the years corresponding to

Republican winners. Similarly, an average beta is calculated for the years corresponding to Democratic winner. Comparing these averages for each market sector will allow us to gauge which sectors are most politically significant, meaning that the political party that wins the election in a given year will have a significant impact on the market risk of a sector of the economy.

Landslide Election Victories

Landslide election victories represent how much of a margin a candidate won each presidential election by. This data is formulated through the county voting statistics. A linear regression model is used on the beta of each sector versus the margin percent that a Republican candidate won by to assess if there exists a relationship between these two variables. With this information revealed, it is possible to make a more thorough conclusion about the market risk of each sector if examining pre- and post-election beta calculation differences when a party wins by a large or small margin.

.

Pre- and Post-Election Market Differences

Studying the market risk prior to an election versus after an election gives an indication of the reaction of the economy due to a specific party winning the election. Doing so will enable a more thorough analysis by studying solely the reaction to a party winning, rather than the expectation that a party will win. In other words, the presidential election is very unpredictable leading up to the election. The market risk during that time could be misleading. However, the

change in the market risk after a candidate wins could supply information on which sectors of the economy benefit more with one party winning over another.

Results

Before showcasing the results of the three main sections, first it is important to illustrate the beta calculations. Table 1 provides the beta calculation for each sector.

Table 1: Beta Calculation for Each Sector per Election Year

Election Information		Beta										
Year	Winner	Consumer Goods	Real Estate	Telecommunicat...	Utilities	Technology	Industrial	Financial	Energy	Consumer Discr...	Health Care	Materials
2000	Republican	0.3	0.21	0.95	0.13	2.13	0.83	1.05	-0.09	0.98	0.7	0.34
2004	Republican	0.79	0.7	0.74	0.42	1.19	1.09	1.03	0.58	1.03	1	1.07
2008	Democrat	0.7	1.37	1.05	0.8	0.93	0.81	1.42	1.36	0.9	0.6	0.94
2012	Democrat	0.83	0.52	0.82	0.39	1.1	1.05	1.14	1.17	0.92	0.75	1.29
2016	Republican	0.93	0.97	1.3	0.89	1.01	1.04	1.05	1.22	1	1.05	1.06

Furthermore, Figure 1 provides a visualization for the beta calculations for the years 2000 to 2004. A line of best fit shows that overall market risk is increasing.

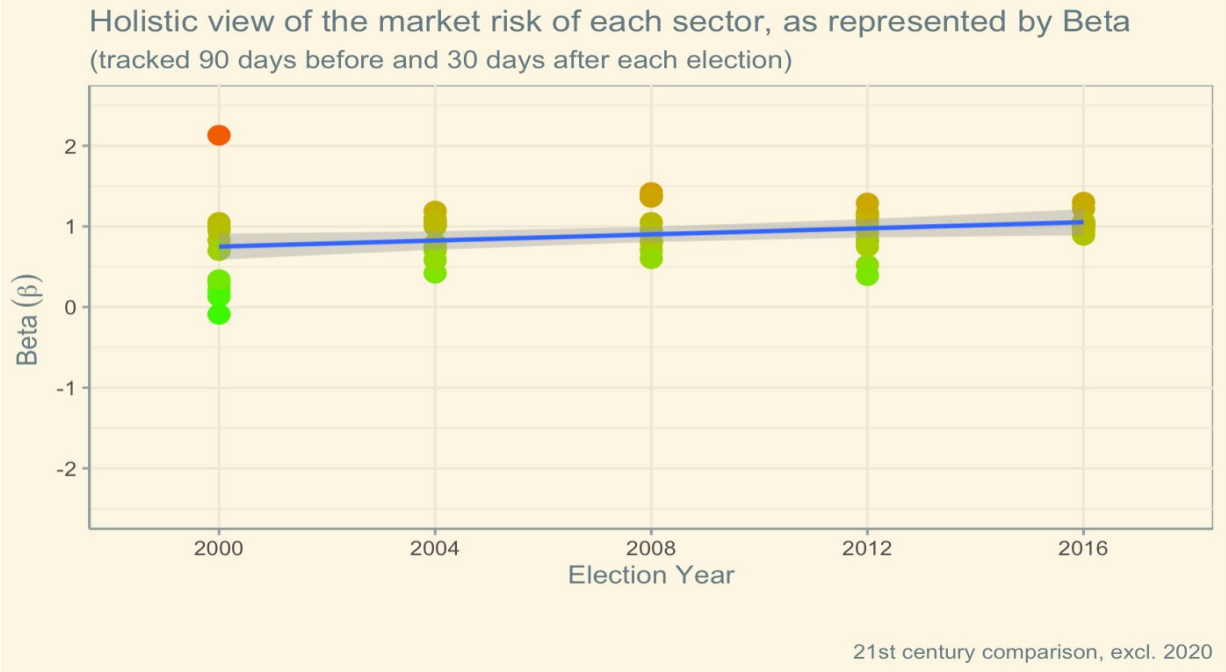

Figure 1: Aggregated Beta Visualization per Election Year

As evident, most sectors have steadily increased their market risk during presidential elections throughout the 21st century (with the exception of 2020). The financial, technology, and consumer discretionary sectors seem to have lower market risk over time. The overall spread of market risk between the sectors seems to lower over time, as well, with 2000 yielding a sector beta of over 2 as well as a sector beta in the negatives. The years following have fewer variations. Interestingly, the highest beta for the year 2008 was the financial sector, during which the housing crisis of 2008 occurred. Banks were the significant cause for this recession due to their shady tactics to finance loans for housing units.

Political Significance

Table 2 provides supplementary information that allows us to compare the average market risk depending on which political party wins the presidential election.

Table 2: Market Risk of Sectors With Political Party Comparison

Sectors	Difference (R-D)	Average	Std. Dev.	Avg. w/ Democrat	Avg. w/ Republican
Energy	0.70	0.85	0.60	1.27	0.57
Technology	0.42	1.27	0.49	1.02	1.44
Real Estate	0.32	0.75	0.44	0.95	0.63
Materials	0.29	0.94	0.36	1.11	0.82
Financial	0.24	1.14	0.16	1.28	1.04
Health Care	0.24	0.82	0.20	0.68	0.92
Utilities	0.12	0.53	0.31	0.60	0.48
Consumer Goods	0.09	0.71	0.24	0.76	0.67
Consumer Discretionary	0.09	0.97	0.05	0.91	1.00
Telecommunications	0.06	0.97	0.22	0.94	1.00
Industrial	0.06	0.96	0.13	0.93	0.99

From the table, which is ordered in a descending fashion for the absolute difference of market risk between the two political parties, it is obvious that the energy, technology, real estate, materials, financial, and health care sectors have the highest differences between the average beta when a republican candidate wins and the average beta when a democratic candidate wins.

Landslide Election Victories

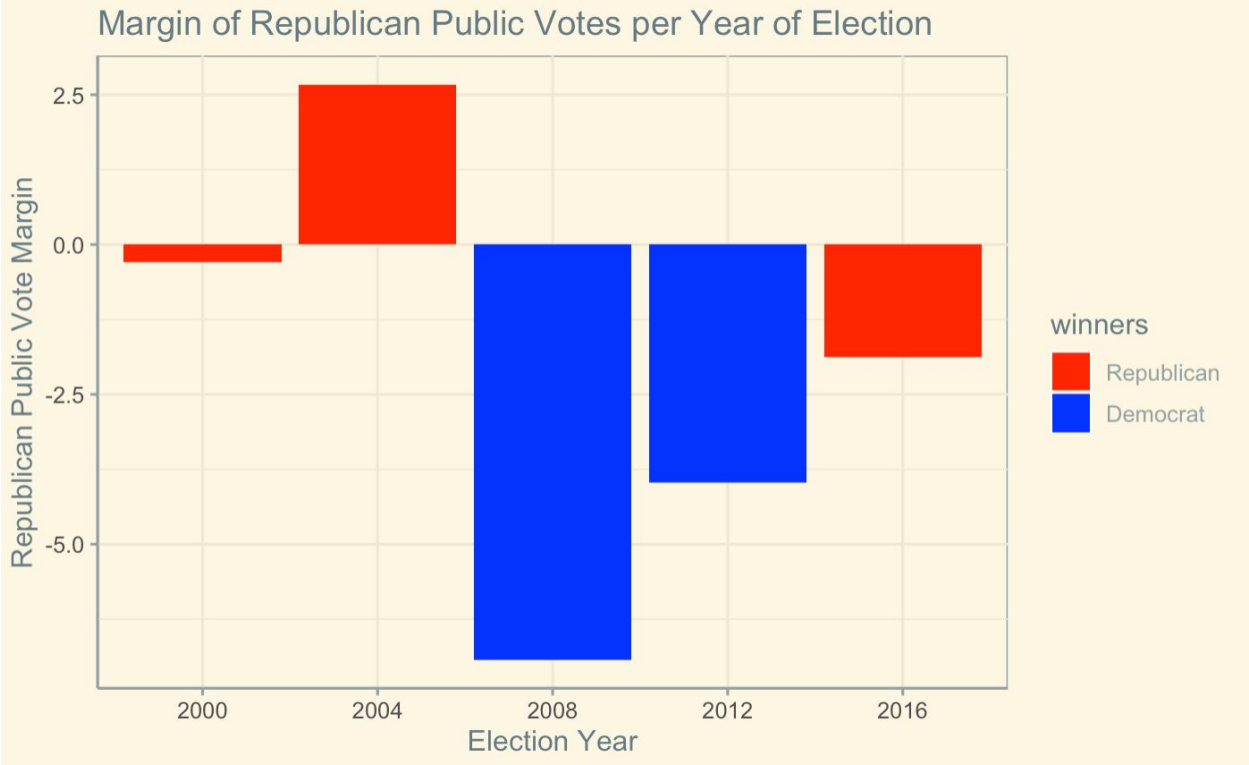

Figure 2: Margin of Republican Public Votes per Year of Election

Figure 2 shows the margin by which a candidate won for each election. This information is used to create linear regression models between individual sector market risk and the margin by which a Republican candidate won (or lost).

Landslide election victories represent how much of a margin a candidate won each presidential election by. A linear regression model is used on the beta of each sector versus the margin percent that a Republican candidate won by to assess if there exists a relationship between these two variables. With this information revealed, it is possible to make a more thorough conclusion about the market risk of each sector if examining pre- and post-election beta calculation differences when a party wins by a large or small margin.

Table 3: Linear Regression Intercepts and Slopes of Market Risk and Party Voting Margin

Sectors	Regression Intercepts	↓ Regression Slopes
Technology	1.39	0.06
Financial	0.89	0.04
Health Care	0.90	-0.02
Real Estate	1.06	-0.04
Energy	0.61	-0.07
Materials	0.61	-0.11

Table 3 shows the linear regression model coefficients of the politically significant market sectors. It is ordered in descending fashion by the regression slope. Clearly, the technology and financial sector yield positive regression slopes, while the rest of the sectors yield negative slopes. The technology sector, additionally, yields the highest regression intercept.

Pre- and Post-Election Market Differences

In this section, the beta values are computed 90 days before each election and 30 days after. Each figure below shows the difference in pre- and post-election market risk in each politically significant sector of the economy.

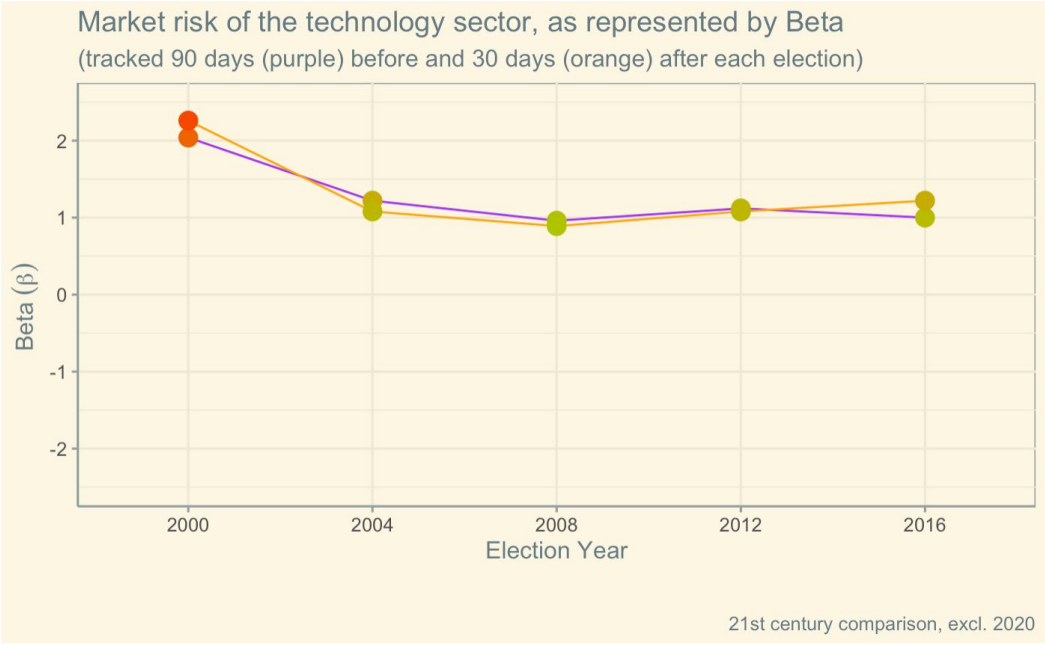

Figure 3: Pre- and Post-Election Market Risk of the Technology Sector, as Represented by Beta

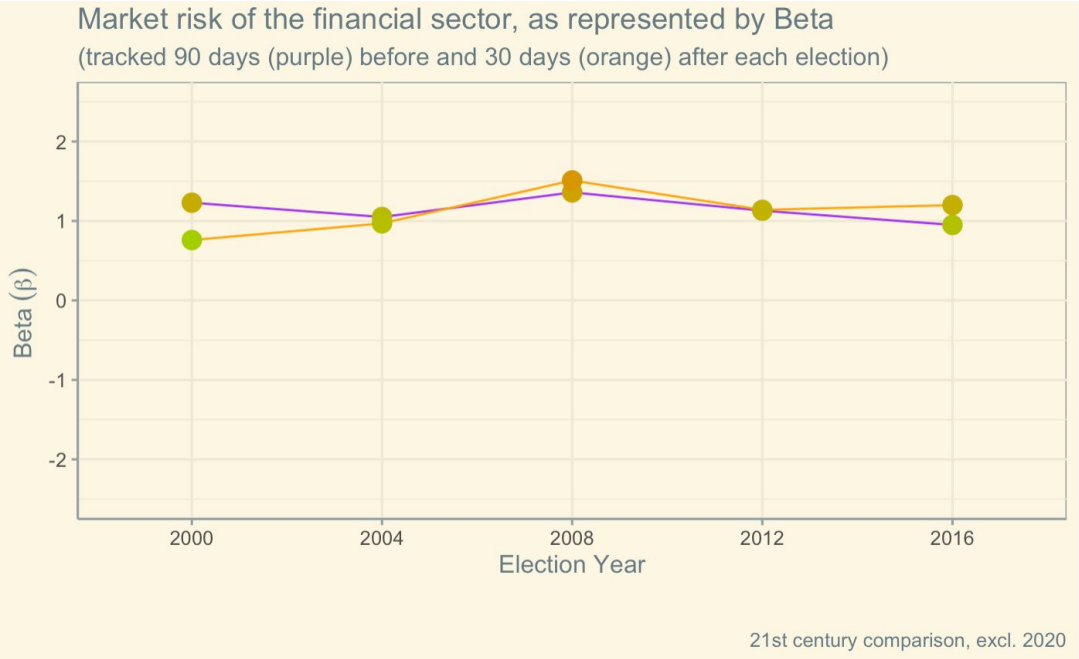

Figure 4: Pre- and Post-Election Market Risk of the Financial Sector, as Represented by Beta

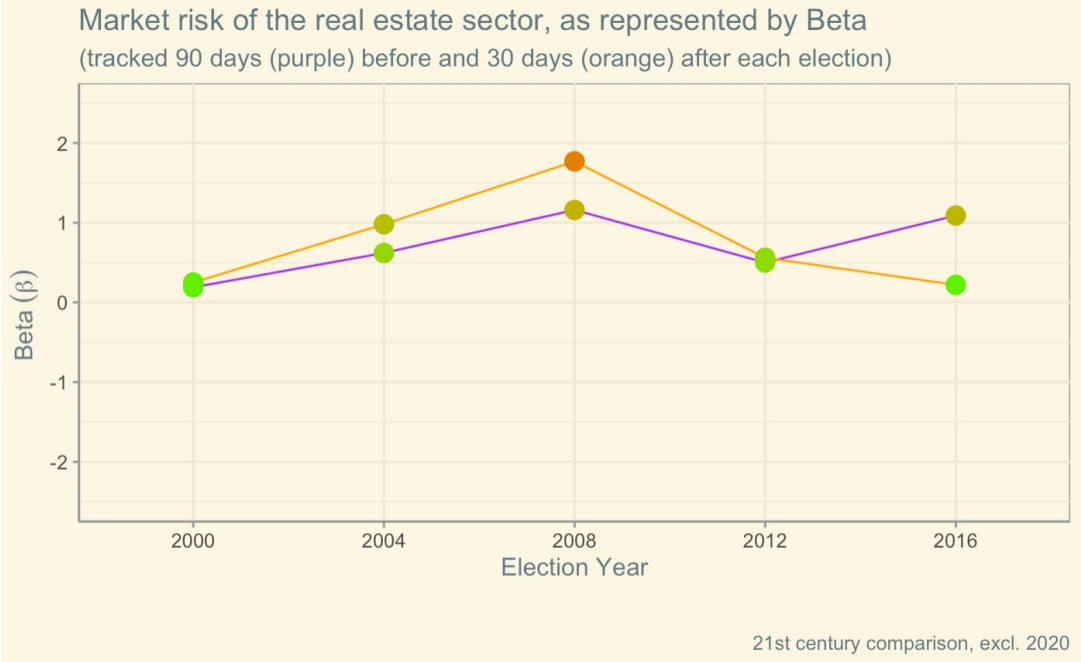

Figure 5: Pre- and Post-Election Market Risk of the Real Estate Sector, as Represented by Beta

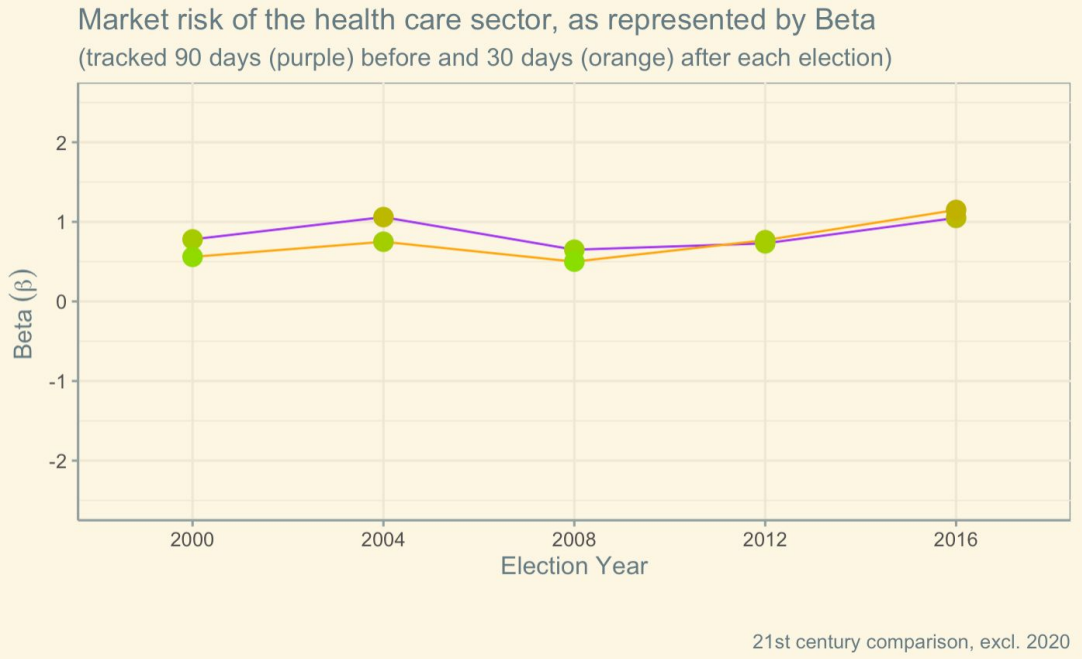

Figure 6: Pre- and Post-Election Market Risk of the Health Care Sector, as Represented by Beta

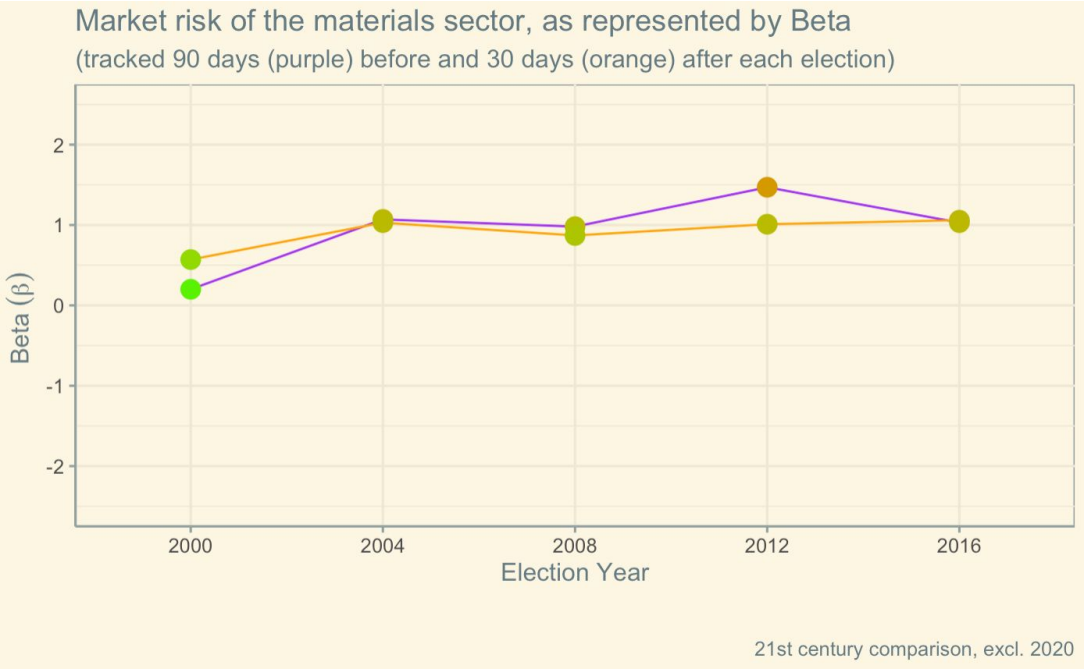

Figure 7: Pre- and Post-Election Market Risk of the Materials Sector, as Represented by Beta

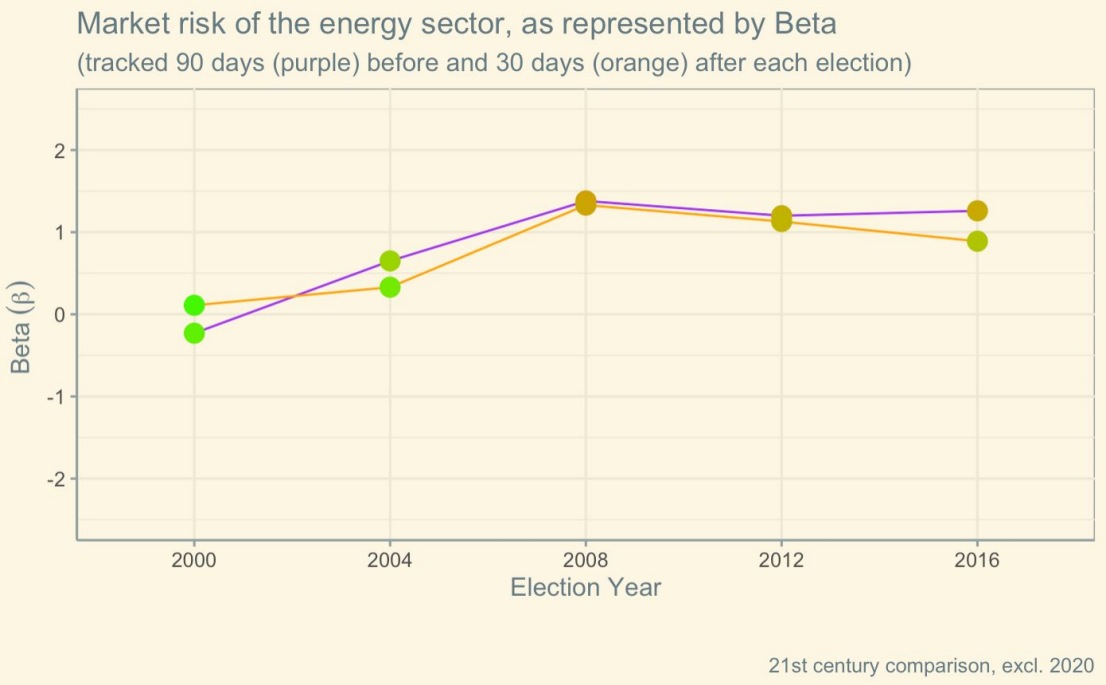

Figure 8: Pre- and Post-Election Market Risk of the Energy Sector, as Represented by Beta

From Figures 3-8, the largest each sector has large differences between pre- and post-election in the majority of the election years. The most noticeable differences in 2000 include the energy, materials, health care, financial, and technology sectors. In 2004, the most noticeable difference is in the energy, health care, and real estate sectors. In 2008, the most noticeable difference is in the real estate and financial sectors. In 2012, the most noticeable difference was in the materials sector. In 2016, the most noticeable difference is in the energy, real estate, financial, and technology sectors.

Finally, it is imperative to display the returns on the overall market, or the S&P 500 Index, as done so in Figure 9. This allows an understanding of how well the overall market is performing.

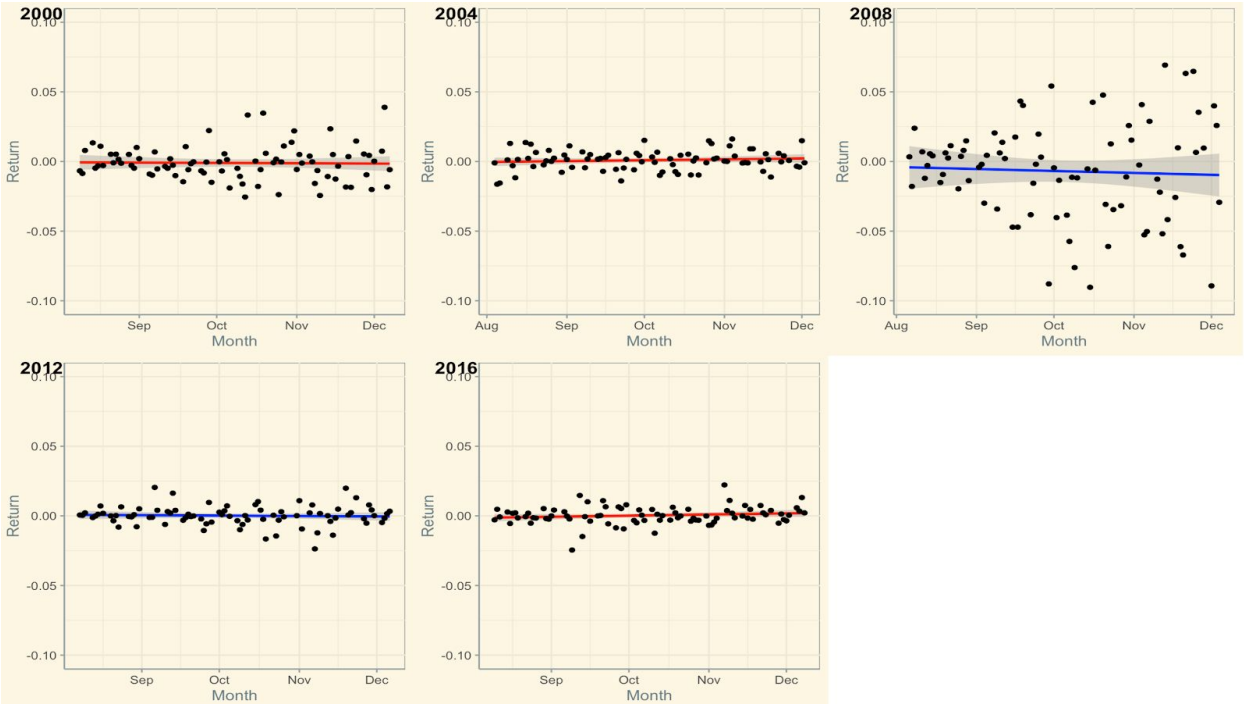

Figure 9: Holistic View of the Market Returns During Elections per Year

Assessing Linearity

When assessing linearity, there are three assumptions to test for the regression models between the percent margin won by a Republican and the market risk: a linear relationship between the dependent and independent variables, exogeneity, and residuals. All sectors provide similar results with linearity, and are tested in the code file provided. Below is an example of the linearity assessment for one of the sectors (real estate):

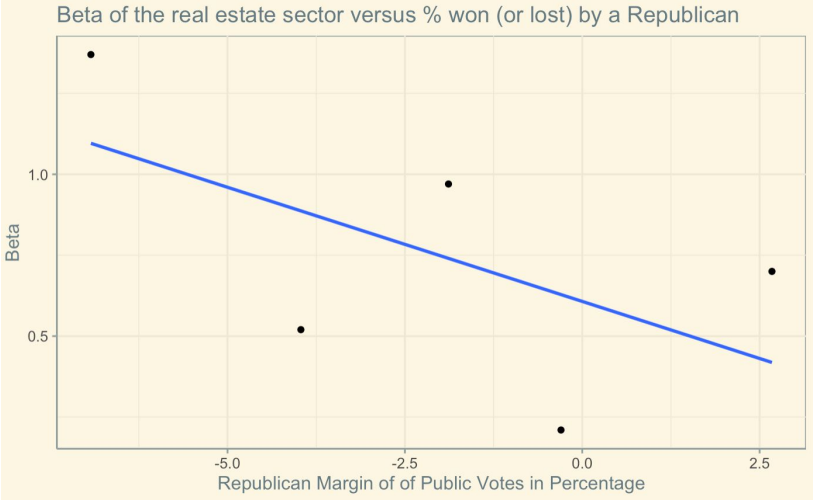

Figure 10: Linear Regression Model of Real Estate Sector

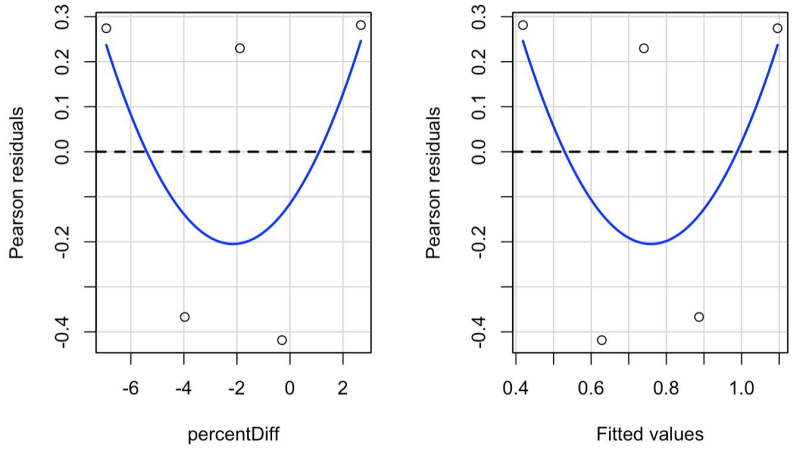

Figure 11: Residual Plot of Linear Regression Model of Real Estate Sector

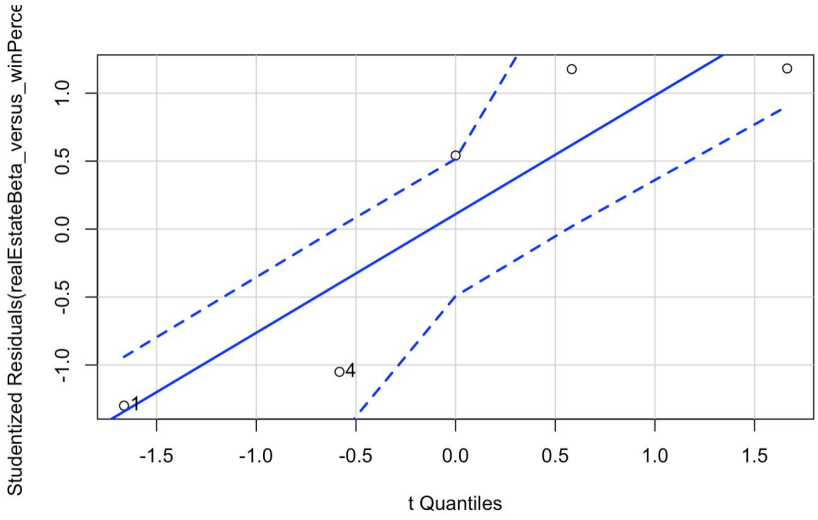

Figure 12: Outlier Plot of Linear Regression Model of Real Estate Sector

It is clear that a linear relationship forms between the two variables. Additionally, the residual plot shows a balance between the residuals, further predicting a linear relationship between the two variables. The outlier test seems to not pass with this sector if the point at t Quantiles 0 is considered an outlier. Thus, the outlier test does not pass. Finally, it is not right to say that the linearity assumptions are satisfied. Not only does the outlier test fail, but the exogeneity circumstance fails certainly, since many more variables impact the two variables studied. In the code provided, each other sector's linearity assumptions are tested. All provide similar results.

Conclusion

Summary

Although the results were not surprising, the models created provide an enticing claim on how much each individual market sector is affected by presidential elections much differently than the other sectors. Moreover, Table 3 provides interesting insight on how the margin of an election relates with the market risk. For instance, in the technology sector, a one percent increase in a Republican margin yielded a beta that was 0.06 higher, on average. Thus, if a Republican candidate is to win by a landslide, and the investor is risk-averse, then the technology sector is not the right sector to invest in for the needs of the investor. If, on the other hand, a Democratic candidate is to win by a landslide, this risk-averse investor is better off investing in the technology sector rather than the materials sector (which has the lowest beta per percent Republican margin).

Simply looking at the overall market index to make informed investment decisions is not robust enough. In 2008, there was a lot of volatility in the market as a result of the housing crisis. However, this does not mean that there are no risk averse investment decisions. The health care beta in 2008 was lower than 1, meaning that there was little market risk. Even in the pre- and post-election data, the health care sector had negligible differences. As a result, a risk averse investor would be able to invest with lower risk than initially assumed by looking at stricly market returns.

The direction that I wanted to take with this project initially was to do the following: use primary voting as a predictor of what candidate will win and if the same thinking of the project is

applied backwards, use the market risk to predict a presidential candidate. The project would be much more interesting with this direction.

Limitations of the Study

There are many issues with this study. The first, and most significant, is that there is limited data that is worked with. The election years 2000-2016 do not provide much clarity on the topic at hand. For example, in two elections, a Democratic candidate won and the candidate was the same person (President Barack Obama). The data could be skewed, therefore, because President Obama's performance would have an affect on the market after his reelection. This is the same case with President George W. Bush. Therefore, our sample size is very low and the project would be improved immensely if data was collected for many more elections. However, this was not possible due to data needed for ETF returns, which did not date to elections prior to 2000. Furthermore, the ETF used is not the best indicator for each sector. Dow Jones provides the best indicators for individual sectors, but the data does not track back far enough to be able to include enough data for presidential elections. Lastly, the one-third of the United States Senate is reelected every four years, during the time of the presidential elections. The United States Senate has a strong influence on policy making, and therefore could be a significant factor in the market.

Overall, more data should be incorporated into this study, and as a result more concrete and statistically significant results can be obtained.