Using Regression Analysis to Compare the returns and volatility of Different S&P500 Industries' during the COVID-19 Pandemic.

Using regression analysis to compare the volatility of different industries market return during COVID-19 compared to the past

Arie Hendrikse[*1]
[1]Faculty of Information Technology, Monash University, Australia 3800 [*]Corresponding author:ahen0012@student.monash.edu

Author contributions.

A.H. proposed the research question, designed the study, conducted the survey, wrangled the data, analysed the results, drafted and edited the paper.

Abstract.

Keywords.

python, numpy, pandas, computer science, COVID-19, finance, SPY, regression, analysis,markets,nasdaq

Introduction.

The COVID-19 pandemic has seen a sharp increase in the volatility of the NASDAQ market (Z. We, I. Florescu, 2020). The S&P 500 has seen some of the sharpest increases as well as declines since the 1987 stock market crash, also known as Black Monday (NASDAQ, 2020). Every day new COVID-19 developments are releases, causing mass uncertainty.

With this paper I provide a detailed analysis of the Beta Coefficient for each individual sub-industry within the Global Industry Classification Standard (GICS). This will determine how industries' volatility compares to the rest of the market during unprecedented lockdown, social distancing, and states of emergency.
This will allow for investors, policy makers, financial analysts, as well as other economic researchers to better understand the volatility to allow for a more informed decision making process in the future.

How will this help the public?
Policy can be directed to mitigate volatility and ensure better returns from more vulnerable industries and highlight the industries in the most distress. Policy-makers reallocate resources such as social and corporate welfare, incentivise weaker industries, and run more in depth cost benefit analysis to determine where tax-payers' money should go. Furthermore, the average retirement 401(k) balance was $106,000 in Q2 2019, with 196,000 Americans holding over $1m in their retirement fund (Fidelity Investments, 2019). Fund managers could use this research to ensure that these large amounts of capital are not lost for a large number of people.

The Standard & Poor 500 index was studied, which covers about 80% of the available market capitalisation (S&P 500, 2020). As the inclusion of stocks in the S&P500 are dynamic, historical data is used for comparison.

The null hypothesis is that the volatility would remain in the same direction with similar magnitudes between each year.

[1] The original research question was "*How has the volatility between different GICS sectors within the S&P500 changed during the COVID-19 pandemic, compared to the same time last year?*"

Method.

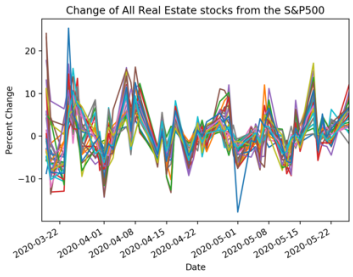
Categorising stocks into particular industries.

GICS is arguably the most widely used standard, with an advantage of being consistent across different applications for effectively grouping stocks with similar operating characteristics (Hrazdil, Trottier, & Zhang, 2013).

I scraped Wikipedia table (S&P 500, 2020), creating a data-frame using the Pandas python library that represented this same table. I then filtered the data to find distinct sub-industries for the stocks that were present within the S&P500. In total, the number of distinct categories was 126, out of a total of 504 stocks.

Downloading the data.

After scrubbing the S&P500 data off of online tables (List of S&P 500 companies, 2020), the stock tickers were then filtered into GICS sub-categories. The percentage change between each Adjusted Close was calculated for each ticker in that industry, then the mean of these percentages was taken. As seen in Fig.1 and Fig.2.



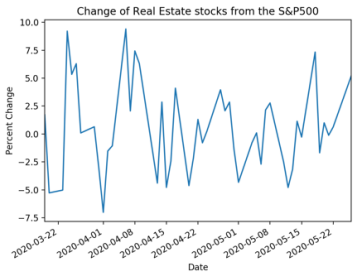


Fig. 1 shows the daily change for each individual Real Estate Stock.

Fig. 2 shows the mean of all Real Estate stock's change.

Date Range for Fig.1 and Fig. 2: 19/03/2020 through 26/03/2020
Pricing time for Fig. 1 and Fig .2: Adjusted Close

Separating the population.

Once the percent change between Adjusted Closings were calculated for each individual stock (we will refer to this as *P*) the sum of each row in *P* was divided by size of *P* (e.g. 'Real Estate' has 31 individual stocks) in that category to determine the mean of percentage changes between Adjusted Close.

The same was done for every stock that was not a part of the category, this will be referred to as *P'* and will be used as the benchmark to analyse against. The reason for comparing an industry to all the other stocks excluding the current industry was because some industries, such as real estate, take a bigger proportion of the S&P meaning they, as previously mentioned, would contribute a greater weight creating bias.

Regression Analysis.

The beta coefficient β of stock is used to measure how volatile it is against the market during a particular time-series. β x 100 shows the overall percentage of volatility citation of the stock with regards to the overall market. If β>1 then the stock is more volatile than the market, and vice versa. The beta coefficient was calculated through an ordinary least squares regression model.

$$\text{Beta coefficient } (\beta) = \frac{Covariance(R_P, R_{P'})}{Variance\ (R_{P'})}$$

$\mathbf{R_P}$ = Return of an industry
$\mathbf{R_{P'}}$ = Return of everything not within the industry.

Results.

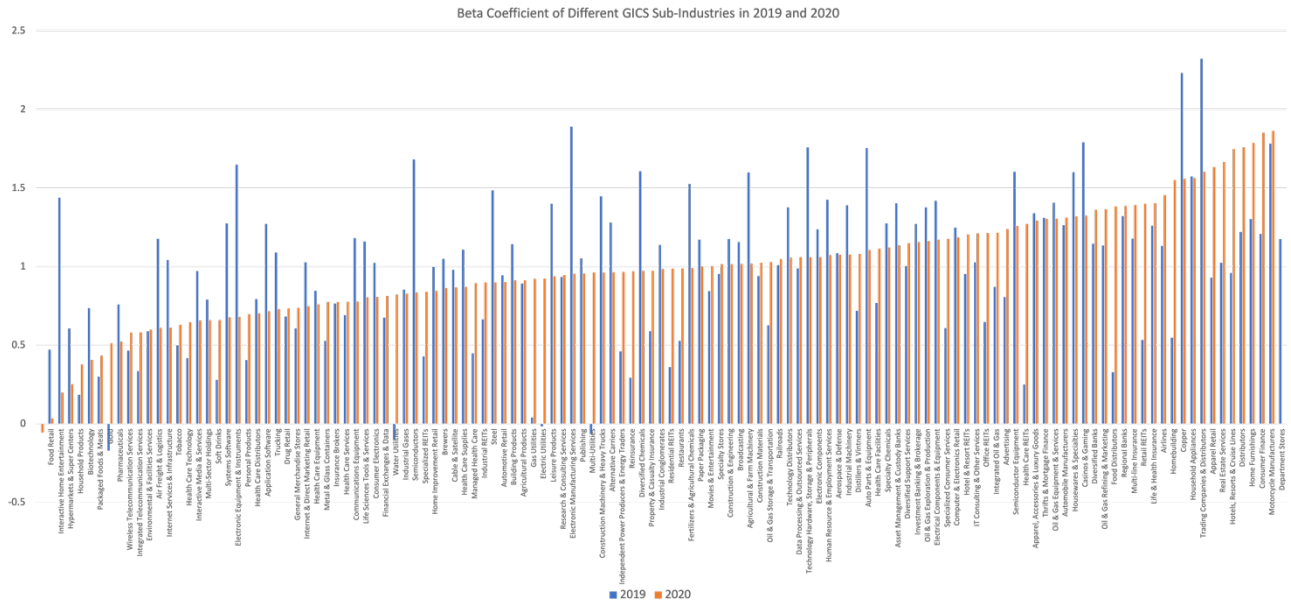


Fig.3 shows the beta coefficient for all 126 sub industries within the S&P500

Some industries in Fig.3 showed large $\Delta\beta$, while others stayed similar. With 28.6% of industry having ($\Delta\beta > .5$) Multi-Utilities, Gold, and Water Utilities volatility turned from going against the market with $\beta < 0$ indicating a negative relationship, to going with market behaviour as indicated by $\beta > 0$.

Discussion.

Gold has been recognised as safe haven in extreme stock market conditions in the past (Baur & Lucey, 2010), however, during COVID-19 the stock has seen an increased risk, now being 43% the volatility of the market.
On the other end of the spectrum, some past volatile stocks have seen less movement with the market;
Air Freight and logistics isn't moving with the market as much as it used to. This could be attributed to cheaper fuel (Albulescu, 2020) and larger supply of planes due to reduced air travel meaning the industry is less affected.
Another noticeable low volatility industry is Interactive Home Entertainment, this could be attributed to stay-at-home boredom increasing the necessity of at home entertainment services (Taylor, 2020), but the lack of a negative Beta could indicate that as consumers are earning less money due to record unemployment (Coibion, Gorodnichenko, & Weber, 2020), and with lockdowns already being reduced households may soon view at home entertainment as a luxury expense (Garnett, 2020). Studies in the future could use this data to link market caution with changed consumer perception of good and services.
Low betas might also be associated a long road to recovery regardless of current market events (e.g. Restaurants).
High volatility industries such as Department Stores and Home Furnishings indicates that these are the most affected by uncertainty, and Governments possibly need to convey more accurately what restrictions will be placed on these in the past or future. High volatility could also indicate that the stocks are more so affected during announcements such as lockdown, stimulus, social distancing laws, and supply chain strain. The chain reaction of supplies being unavailable is expected to last well beyond the pandemic (Ivanov, 2020), and this could explain why manufacturers have seen increased volatility.

Conclusion.

Every day COVID-19 creates new speculation on how the market will perform.

Acknowledgements.

AH wishes to acknowledge the support of the FIT2083 staff, for providing the skills necessary to write this paper. He also wishes to acknowledge the authors at Investopedia, developers of pandas, and Marcel Masque for teaching me data science in such a short time period.

Bibliography.

References

Scott R. Baker, R.A. Farrokhnia, Steffen Meyer, Michaela Pagel, Constantine Yannelis (2020). HOW DOES HOUSEHOLD SPENDING RESPOND TO AN EPIDEMIC? CONSUMPTION DURING THE 2020 COVID-19 PANDEMIC. National Bureau of Economic Research. Available at: https://www.nber.org/papers/w26949.pdf

Ye, Ziwen and Florescu, Ionut, (2020) COVID-19 and the Equity Market . A March 2020 Tale Available at http://dx.doi.org/10.2139/ssrn.3566281

E.A.Farmer, R. (2012). The stock market crash of 2008 caused the Great Recession: Theory and evidence. Journal of Economic Dynamics and Control.

Available at: https://www.sciencedirect.com/science/article/pii/S0165188912000401

Hong-Yi, C., Sheng-Syan, C., Chin-Wen, H., & Cheng-Few, L. (2013). Does revenue momentum drive or ride earnings or price momentum? Journal of Banking & Finance

Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transportation Research Part E: Logistics and Transportation Review.

Available at: https://www.sciencedirect.com/science/article/pii/S1366554520304300

Scott R. Baker, R. F. (2020). HOW DOES HOUSEHOLD SPENDING RESPOND TO AN EPIDEMIC?

CONSUMPTION DURING THE 2020 COVID-19 PANDEMIC. National Bureau of Economic Research. Available at: https://www.nber.org/papers/w26949

Sizzlers and Fizzlers. (2020, April). Retrieved from SPI Indeces:

https://us.spindices.com/indexology/djia-and-sp-500/sizzlers-and-fizzlers

Leigh, A. (2012). How Much Did the 2009 Australian Fiscal Stimulus Boost Demand? Evidence from Household-Reported Spending Effects. The B.E. Journal of Macroeconomics.

Available at : https://www.degruyter.com/view/journals/bejm/12/1/article-1935-1690.2035.xml.xml

NASDAQ (2020) *https://www.nasdaq.com*