Analysis of SPDR S&P 500 Trust (SPY)

Email: bsmcconathy@cpp.edu

Alan Perez
Department of Mathematics and Statistics
Email: alanjperez@cpp.edu

Brandon McConnathy
Department of Mathematics and Statistics

Kalvin He
Department of Economics
Email: khe@cpp.edu

***Abstract*— The SPDR S&P 500 Trust (SPY) is an exchange-traded fund (ETF) that tracks the S&P 500 index. It is one of the most popularly traded funds as it replicates the index at 10% of the S&P 500. Given that this stock is a strong indicator of how the market is trending, we were interested in analyzing its relationship between SPY and other global factors.**

I. Introduction

As investing becomes more popular and accessible to the masses, a deep understanding of the stock market and knowledge about its potential is desperately sought after. As monumental world events are becoming more common, we would like to inquire about the most popular form of investing and saving: the stock market. Our goal is to perform some data analysis on SPY to further our understanding on the way the stock market moves.

SPY is an ETF that aims to follow the S&P 500. An ETF is a collection of stocks and the S&P 500 is a collection of the 500 largest publicly traded companies in the United States. So SPY is made up of these top 500 companies in which it has holdings in all of them. The percentage that makes up SPY varies based on what company it is. For example, Apple makes up 22.8% while Tesla makes up 5.6%. Since SPY is made up of the top 500 companies, it is a really good indicator to tell us what the market as a whole is trending towards. SPY is also one of the most popular funds among investors. This is because, on average, it increases in value by 7-10% per year. This makes it extremely popular among lots of investors since it has such a large return per year. SPY is also a very safe investment. Since it is made up of the top 500 companies in the United States, it is a very safe investment. If one stock inside of SPY decreased a lot in value, the overall value of SPY will not decrease by a lot since it has other holdings that influence its price. This safety also contributes to its popularity among lots of investors.

Seeing as how important SPY is for the entire market, we are also interested in what factors affect it. These factors could include volatility, inflation, mortgage rates, etc [4]. For the sake of brevity, we only explored the effects of volatility and average mortgage interest rates on the stock.

With respect to the stock market, volatility is how quickly the stock goes up or down in value. In order to track volatility, we use the Chicago Board Options Exchange (CBOE) Volatility Index - also known as the VIX [5]. We choose to explore volatility in this paper because we are interested in the magnitude by which fear and world events affect the stock market. In our exploration, we expect to see SPY and VIX generally move in opposite directions because world events typically trigger large sell-offs.

We are interested in this relationship because we know that rising interest rates impede stock market growth. Also, rising interest rates usually cause fixed-rate mortgages to increase in the housing market. Thus, we wondered if fixed-rate mortgages also had an effect on stock market growth or if they were both byproducts of rising interest rates with no relationship between them. We predict a relatively strong negative relationship between average mortgage interest rates and the adjusted close price of SPY.

II. Data Source

A. *Source*

All stock data came from Yahoo! Finance[3]. This includes the data collected on SPY, VIX, and the stocks within SPY: Apple, Microsoft, Facebook, Amazon, Google, Tesla, Nvidia, and J.P. Morgan. Our data about the 15- and 30-year fixed-rate mortgage average in the United States came from the Federal Reserve Bank of St. Louis website [1][2].

B. How the Data Was Accessed

Throughout the project, we used a number of different datasets. Some of which were accessible only through .csv files while others were read and accessed using built-in python packages.

To access our stock data we imported the pandas_datareader.data package and then used the

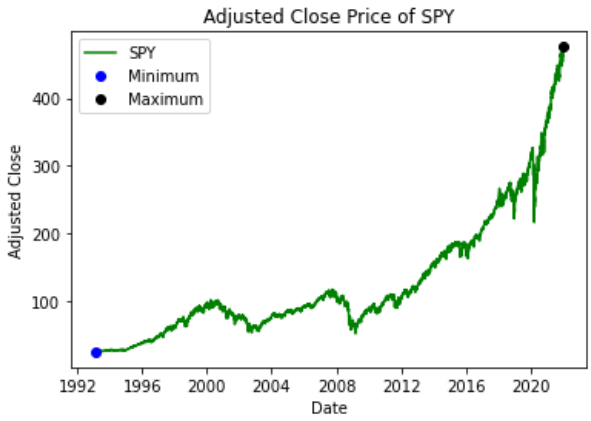
.DataReader() method. To read CSV files we also used the built-in pandas.read_csv() function.

The mortgage interest rate data was first downloaded as a CSV file [1][2] and then loaded using pd.read_csv(). We then searched through the data and fixed any discrepancies within the data that were ruining the plot.

We used the DataReader() function within the pandas_datareader package to read in any stock data. This was used for analysis of SPY, ^VIX, and the stocks within SPY directly from Yahoo! Finance[3].

III. Diagrams and Models

First we want to look at the historical adjusted closing price of SPY. As seen below in Figure 1, the minimum price of SPY was around when it started at about 25 and



increased to its max at about 476.

Fig 1. Historical Adjusted Close Price of SPY

We now look at the historical volume of SPY which can show us interest in a stock and it can also increase by large amounts during times of uncertainty where large price changes happen.

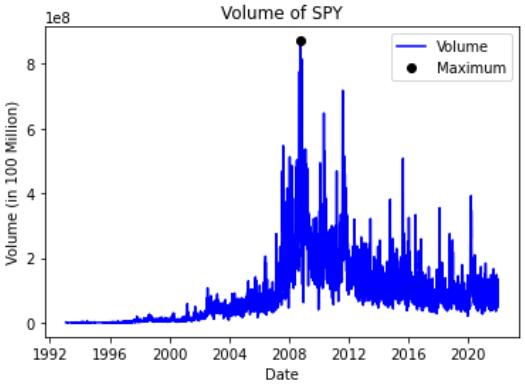


Fig 2. Historical Volume of SPY

It is also important to look at the daily returns of SPY which can show us how much the price increased or decreased on a given day. This is a good indicator to see how safe a stock is.

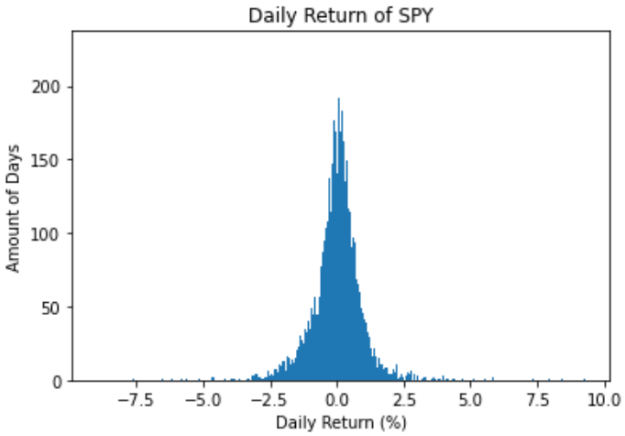


Fig 3. Daily Return of SPY

Since we are exploring the relationship between SPY and volatility, we first want to explore how VIX has moved over the years, as seen in Figure 4.

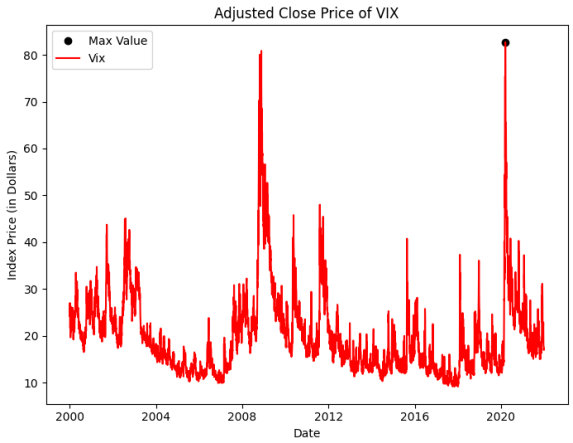


Fig. 4 Adjusted Close Price of VIX over 20 years

Now we examine how SPY has moved in comparison to VIX over the past 20 years. We hope this will reveal some trends between the two stocks. Given that SPY and VIX are very different scales with SPY being worth several hundreds while VIX tends to hover around the 15 to 30 dollar range, we plotted their scaled values. This was done by taking the first recorded value of both SPY and VIX in the year 2000 and then dividing the rest of the values in the respective stock by that number. This presented us with more obvious trends as shown in Figure 6.

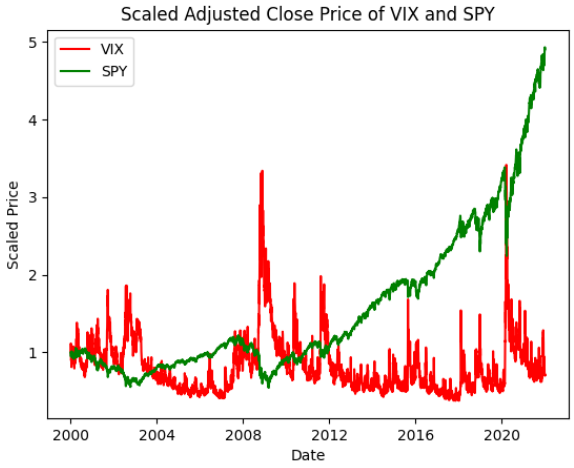


Fig 6. Scaled Adjusted Close Price of VIX and SPY

To further explore the relationship, we created a heat map between SPY Daily Return and VIX Daily Return (Fig. 7).

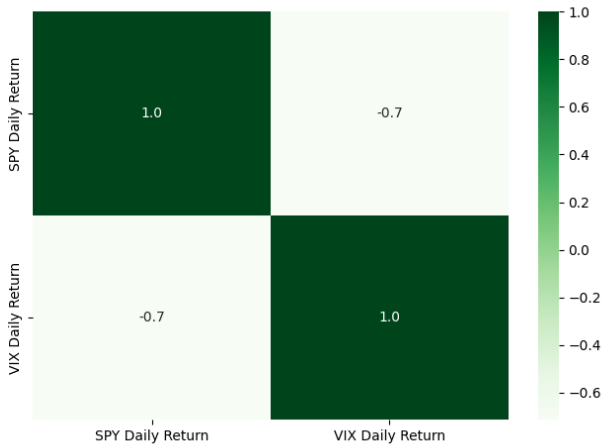


Fig. 7 Heatmap between SPY Daily Return and VIX Daily Return

This presented a fairly strong negative correlation value of -0.7 which will be explored in the Findings and Suggestions section.

Our last inquiry in regards to volatility is its effect on volume. This can be demonstrated using another heatmap by comparing SPY Volume to VIX Daily Return. Figure 8 shows this and depicts a moderate relationship (correlation of 0.5) between the two.

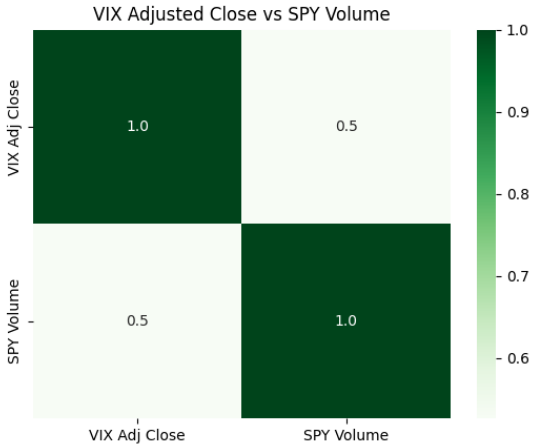


Fig 8. VIX Adjusted Close and SPY Volume

After performing some data exploration on the 15- and 30-year fixed-rate mortgage average, we found that data was only recorded once a week. This means that we are unable to compare the two data sets since they are of different sizes. Thus, to conduct any analysis we have to rely on our plots.

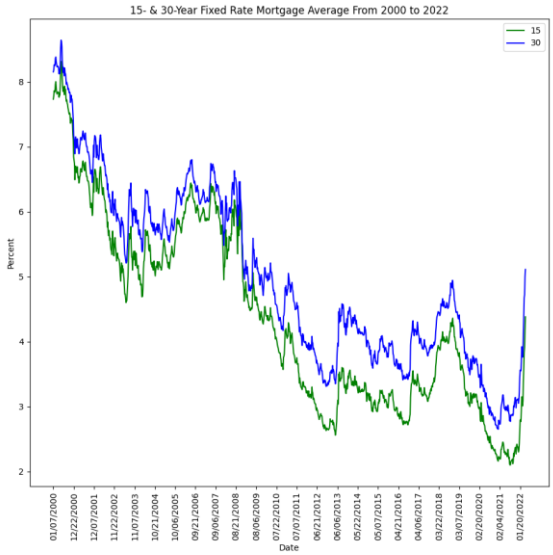


Figure 9. 15- & 30-year Fixed-Rate Mortgage Averages

IV. Findings and Suggestions

We predicted to see a strong negative relationship between SPY daily return and VIX daily return. We also predicted a really strong positive relationship between VIX adjusted close and SPY Volume.

First, we look at how VIX has moved over the last 20 years (Fig. 4). We see a number of large spikes over the

years with two very prominent ones in 2008 and in 2020. The first spike in 2008 is a result of the Great Recession, and the second spike in 2020 is a result of global COVID shutdowns. The all-time highest VIX Close was in March of 2020 with a value of 82.69.

Figure 6 shows some obvious trends in the price movements. Typically when volatility spikes up, SPY adjusted closed values spikes down. To analyze this quantitatively we create a heatmap of both SPY Daily return and VIX daily return (Fig. 7). This shows a correlation of -0.7. This is a strong negative relationship between the two and means that on average as VIX daily return is higher, we can expect SPY daily return to be lower.

Our second heatmap (Fig. 8) shows a correlation of 0.5 between VIX adj close and SPY volume. This is a moderately strong statistical relationship between the two variables.

To examine how average mortgage interest rates affect stock market growth, we explored data recorded over the last 20 years. The 15- and 30-year fixed-rate mortgage averages were plotted in Figure 9. At first, we were interested in also plotting SPY adjusted close price on the same figure, but found that fixed-rate mortgages were only recorded once a week while stock data was recorded daily. Thus, we resorted to analyzing the plots instead. By looking at both Figure 1 and Figure 9 we see no obvious trend between the two.

V. Components of SPY

As stated, SPY follows the exchange tradable funds SPDR S&P 500 which consists of the top 500 profitable companies within the stock exchange. To create a visualization in regards to the companies within SPY, Fig. 10 shows the top 9 holdings and the percentages of each individual company. Apple, Microsoft, and Amazon hold most of the weight within the portfolio. This means that major fluctuations within either Apple or Microsoft are going to reflect on the movement and price of SPY.

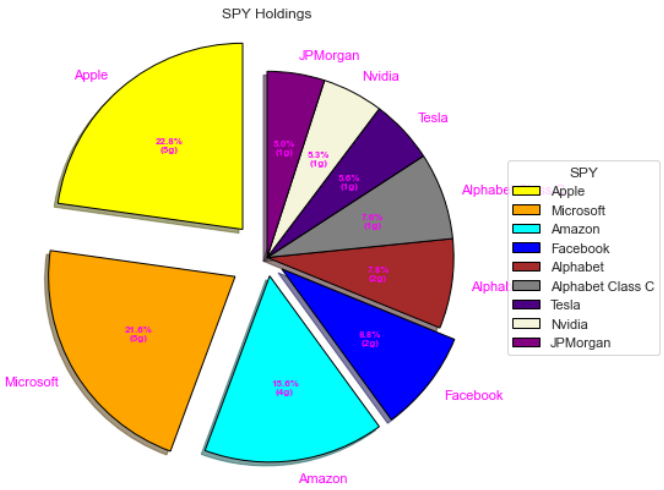


Fig. 10 SPY Holdings Pie Chart

VI. Movement in SPY Compared to Holdings

SPY's portfolio consists of about 500 different companies, which means that the individual performance of each company has an effect on SPY's movement and pricing. To minimize risk, SPY diversifies its holdings with a combination of different stocks from different sectors within the

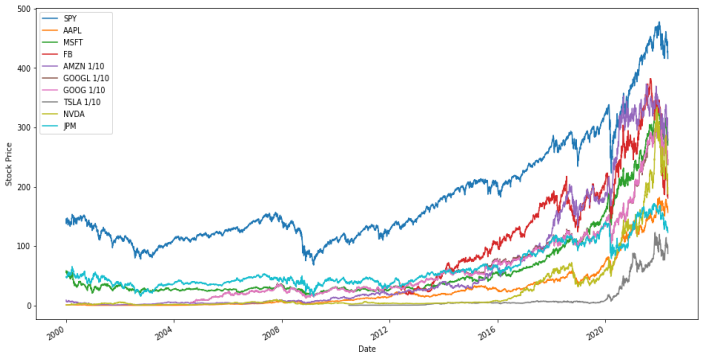


Fig. 11 SPY vs Stocks

market. As a way to help visualize the trend of SPY compared to its holdings, Fig. 11 shows the top 9 stocks and their movement. Stock prices of Amazon, Google, Google Class C, and Tesla were manipulated to be 1/10 of their original price. This helps with fitment where it is easier to identify a pattern. The pattern is faint where you can see some correlation between the different stocks and the performance of SPY. However, in 2020 during COVID the patterns are more concrete and you could see how each individual stock seemed like they were mimicking each other's movements.

VII. OLS Regression to Create a Model

In order to find a relationship between the stock prices and SPY, OLS regression is used to try and find the effects on SPY. With linear regression we are able to take all the variables into consideration when we regress it onto SPY. Fig. 12 shows us the data after running OLS regression.

OLS Regression Results

Dep. Variable:	SPY	R-squared:	0.982
Model:	OLS	Adj. R-squared:	0.982
Method:	Least Squares	F-statistic:	2.250e+04
Date:	Tue, 03 May 2022	Prob (F-statistic):	0.00
Time:	14:30:25	Log-Likelihood:	-9685.5
No. Observations:	2501	AIC:	1.939e+04
Df Residuals:	2494	BIC:	1.943e+04
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	120.7141	0.626	192.723	0.000	119.486	121.942
AAPL	0.0903	0.031	2.953	0.003	0.030	0.150
MSFT	0.0444	0.026	1.678	0.093	-0.007	0.096
AMZN	0.0094	0.001	7.613	0.000	0.007	0.012
FB	0.2339	0.011	22.146	0.000	0.213	0.255
GOOGL	0.2172	0.018	11.866	0.000	0.181	0.253
GOOG	-0.1434	0.018	-7.841	0.000	-0.179	-0.108

Omnibus:	34.088	Durbin-Watson:	0.029
Prob(Omnibus):	0.000	Jarque-Bera (JB):	21.460
Skew:	-0.063	Prob(JB):	2.19e-05
Kurtosis:	2.564	Cond. No.	6.68e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly speci
[2] The condition number is large, 6.68e+03. This might indicate that there are
strong multicollinearity or other numerical problems.

Fig. 12 OLS Regression Results

From the results we can deduce that our model for this project is :

SPY = 120.714 + 0.090(AAPL) + 0.044(MSFT) + 0.009(AMZN) + 0.233(FB) +0.217(GOOGL) - 0.143(GOOG)

Notice on Fig. 13, the second graph shows the residuals of AAPL. The lack of any patterns indicate that there is no homoscedasticity and heteroscedasticity problem within the model. The model is regressed traditionally without doing robust regression due to the fact that there are no heteroskedasticity problems within the model.

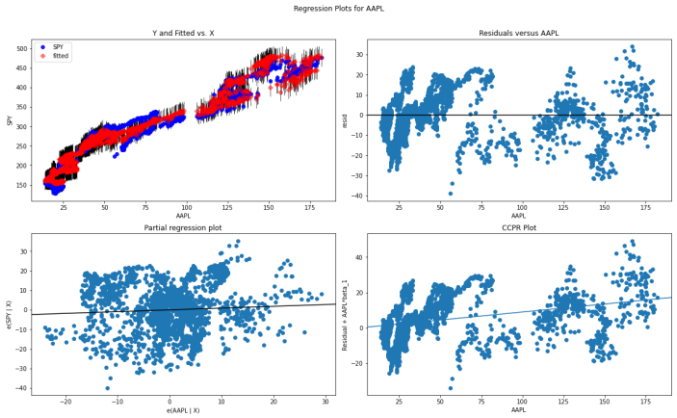


Fig. 13 Regression Plots for AAPL

The regression analysis on Fig. 13 tests for the validity of each variable. In this case we're seeing the residual of AAPL when it gets regressed into SPY. The first graph shows the fitment between the variable and the regressed line. It holds true, similar to the OLS regression which produces a very high R-squared. This shows that the variables have little variation between each other so the outcome is somewhat accurate. The problem with this lies in the fact that due to low variation there is a high bias within this model. The bias comes from the relationship between AAPL and SPY and how AAPL makes about 23% of SPY's holdings. With a bais it means that the predicted variable which would be SPY has a dependence on the explanatory variable which would be AAPL.

VIII. Conclusion

In this research project, we utilized different analytical concepts to understand how SPY functions in the market. By looking at its volume and seeing its average daily returns, we understand why SPY is one of the most preferred ETFs in the stock exchange. The price of the volatility index VIX has minor effects on the price of SPY, unless it is during a time where uncertainty runs high. Normally, the increasing volume while daily returns range from -2.5 to +2.5 shows liquidity within the stock which cements it being a safe bet among investors.

REFERENCES

[1] "15-Year Fixed Rate Mortgage Average in the United States." *FRED*, https://fred.stlouisfed.org/series/MORTGAGE15US. Accessed 9 May 2022.

[2] "30-Year Fixed Rate Mortgage Average in the United States." *FRED*, https://fred.stlouisfed.org/series/MORTGAGE30US. Accessed 9 May 2022.

[3] (n.d.). Yahoo Finance - Stock Market Live, Quotes, Business & Finance News. Retrieved May 10, 2022, from https://finance.yahoo.com

[4] *Factors Affecting Stock Market - News, Economy, Pandemic & more*. (2022, March 4). Top 10 Stock Brokers in India. Retrieved May 11, 2022, from https://top10stockbroker.com/share-market-education/factors-affecting-stock-market/

[5] Kuepper, J. (n.d.). *Cboe Volatility Index (VIX) Definition*. Investopedia. Retrieved May 11, 2022, from https://www.investopedia.com/terms/v/vix.asp