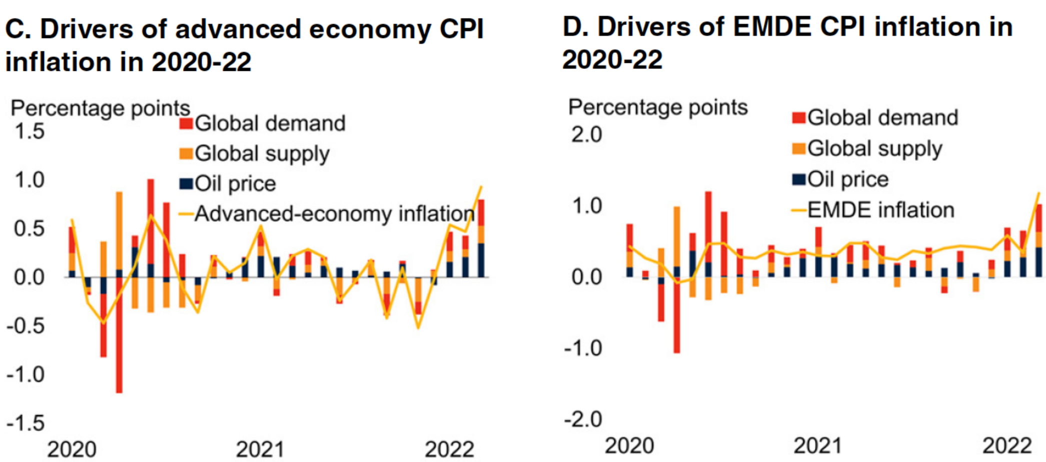
Group Project: The Effect of Macroeconomic Indicators on Stock Market Performance
Team 011: Thomas Byrne, Liangqu Chen, Jari Oinas, Anuj Shelat, David Chang

Project Report

1. Overview

In this project, we analyze the relationship between various macroeconomic factors during 1967 – 2022 and the performance of the S&P 500 stock market index in the United States. Our primary research question is stated as follows: “What is the relationship between the U.S. stock market and macroeconomic indicators for the United States?”

Since 2020, global stock markets have been characterized by high volatility due to market disturbance caused by the COVID-19 pandemic (2020-Present) and the Russo-Ukrainian War (2022-Present). The current macroeconomic environment has been characterized by persistent, high inflation driven by increases in oil, food prices, and other supply/demand shocks.



Sources: Ha, Kose, and Ohnsorge (2021a); U.S. Bureau of Labor Statistics; World Bank.
Note: CPI = consumer price index; EMDEs = emerging market and developing economies.

Subsequent fiscal policy has tightened at an unprecedented rate, driving consumer sentiment to a record-low. As a result, the S&P 500 market index has had its worst first half of a year since 1970.

As our project proposal stated:

Our hypothesis is that a strong economy (as indicated by strong or healthy macroeconomic indicators) causes strong market performance. The underlying assumptions behind our hypothesis are the mainstream economic theories of price equilibrium, rational expectations, and the efficient market hypothesis. Given that the analysis will be done with data over a long period of time based on quarterly returns of major equity indices, the research group assumes that individual stocks will not have an effect and the market will be efficient. We assume that investors are rational and make rational investment decisions, and that stock market prices will move towards a price equilibrium that is defined by economic fundamentals such as supply and demand. In efficient markets, macroeconomic factors should be the only predictors of the value of market indices. Therefore, the regression analysis done by the team should result in a strong correlation and that we will be able to reject the null hypothesis.

In this project, we will build and fit a linear regression model taking the form of:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{i1} + \beta_{2j}X_{21} + \dots + \beta_{pj}X_{ip} + \varepsilon_{ij}$$

using S&P 500 stock market return data and examine the model to see if there is a linear relationship between the stock market and macroeconomic predictors in the United States.

2. Literature Review

The link between stock market behavior and macroeconomic variables, have been illustrated by different theories in the academic literature, including the Efficient Market Hypothesis, the Dividend Discount Model and the Arbitrage Price Theory (Alshogeathri 2011).

The Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH), is a hypothesis that states that share prices reflect all information (such as the macroeconomic environment) and consistent alpha generation is impossible, regardless of the investment strategies used. According to the EMH, stocks always trade at their fair value on exchanges, making it impossible for investors to purchase undervalued stocks or sell stocks for inflated prices. (Fama, 1970) The only way for an investor to obtain higher returns is by purchasing riskier investments (Investopedia).

The Efficient Market Hypothesis can be distinguished into three forms: weak, semi-strong and strong market efficiency (Alshogeathri 2011):

- The weak form: current stock prices incorporate all relevant past information
- The semi-strong form: current stock prices fully reflect all available public information such as the past price of the stock, how the company is performing, expectations regarding macroeconomic factors as well as public information about GDP, money supply and interest rate etc.

- The strong form: in addition to past and public information, stock prices also reflect private information about the specific company

This hypothesis is integrated into the Capital Asset Pricing Model. In the Capital Asset Pricing Model, beta is a measure of the asset's sensitivity to undiversifiable market risk or volatility (Sharpe, 1964; Markowitz, 1952).

Dividend Discount Model

The Dividend Discount Model (DDM) is a quantitative method used for predicting the price of a company's stock based on the theory that its present-day price is worth the sum of all of its future dividend payments when discounted back to their present value. (Gordon and Shapiro, 1956) Fair value of a stock is calculated irrespective of the prevailing market conditions (Investopedia).

DDM estimates the value of a common share at time t, using the relationship:

$$E_t[P_t] = \sum_{i=t+1}^{\infty} E_t[D_i] / (1 + r_t)^{i-t}$$

Where $E_t[P_t]$ is the expected intrinsic value or price that we would expect to pay for the share in the year t based on the information we have at the time. D_i is the nominal annual dividends we expect to be paid at time i, and r_t is the discount rate investors demand at time t (Foerster and Sapp 2005).

In further model development, macroeconomic factors are considered in dividend growth estimates. The growth rate of GNP is frequently argued to be the maximum sustainable growth rate for a firm's dividends. Other macroeconomic factors also may influence the firm's abilities to pay dividends. Booth (1998) forecasts the dividend growth using the model:

$$Divagro_t = \alpha_0 + \alpha_1 Yield + \alpha_2 Inflation_t + \alpha_3 GNPGrowth_t + \epsilon_t$$

where $Divagro_t$ is the dividend growth rate at time t, r_t is the yield on a long-term government bond to represent the opportunity costs for the firm and investors. $Inflation_t$ is the overall rate of inflation based on the year-over-year changes in the consumers price index (CPI) to capture changes in overall level of risk. $GNPGrowth_t$ is the year-over-year growth in GNP and is the residual error term.

The Arbitrage Pricing Theory

Arbitrage Pricing Theory (APT) is a multi-factor asset pricing model based on the idea that an asset's returns can be predicted using the linear relationship between the asset's expected return and a number of macroeconomic variables that capture systematic risk (Investopedia, Ross 1976).

$$R_{it} = r_i^f + \beta_i X_t + \epsilon_t$$

where R_{it} is the return of stock i at time t,R_i^f is the risk free interest rate or the expected return at time t, X_t is a vector of the predetermined macroeconomic factors or the systematic risks and β_t is the measure of the stocks sensitivity to each of these macroeconomic factors. ϵ_t is the error term representing unsystematic risk (Alshogeathri 2011).

The Arbitrage Price Theory does however not specify which or how many macroeconomic factors to include in the modeling of the stock return. Hence, a large number of different macroeconomic factors such as the interest rate, money supply, inflation and exchange rates have been included in a large body of empirical studies based on reasonable theory (Alshogeathri 2011).

Other research by the International Monetary Fund has shown that the stock market responds to or is a leading indicator to changes in macroeconomic factors. (Comincioli, 1996. IMF, 1998).

The aforementioned theories and models in the academic literature indicate that macroeconomic factors play a key role in affecting market returns.

3. Data Preparation

For this project, macroeconomic data and market return data was compiled from Federal Reserve Economic Data (FRED), yFinance, and NASDAQ using their respective APIs. Macroeconomic and finance data is from Jan 1967 until Mar 2022. The data points and data sources are shown in Appendix 1.

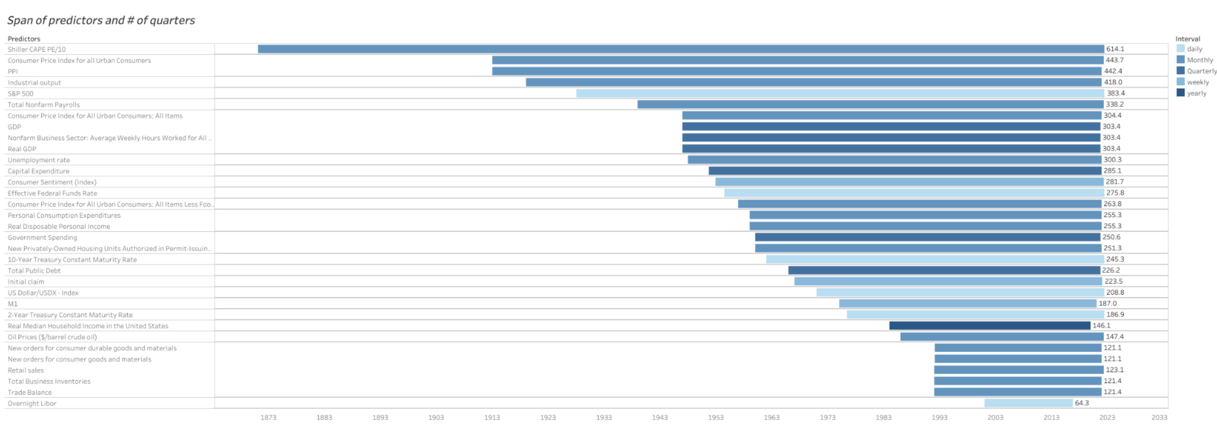


Figure 1: Time Span of Predictors and Number of Quarters of Data available for study

A key consideration was selecting a time frame we should study and which predictors we should include in our data set for our model. The data set we initially compiled included 29 variables (28 independent variables, 1 dependent variable for the S&P 500 market returns). The availability of data was different for each variable, with some variables having data going all the way back to 1873 (through present date) and others only as far back as 1993 (through present date). Ultimately, we decided on the time frame of 1967-2022 due to a desire to capture a wide range of different macroeconomic environments (such as the inflationary period during the 1970's, dotcom tech bubble in the late 1990's, etc) in our analysis of the stock market and build a multiple linear regression model with a high number of macroeconomic independent variables. The resulting data set contains 21 independent variables (macroeconomic factors) and 1 dependent variable (the S&P 500 market return).

The time frequency of our data ranges from daily to quarterly, depending on the variables. Our team has decided that we would fit the model to both quarterly and monthly time frames, because based on initial data exploration, it appears that we could achieve a better model fit, when using longer time frames.

We decided to transform our variables to month over month (MoM) and quarter over quarter (QoQ) and model the S&P 500 return using two different time intervals: monthly (MoM) and quarterly (QoQ). For monthly dataset, we forward filled all the null values for quarterly predictors

4. Exploratory Data Analysis

4.1 Trends 1967 - 2022

Trends for SP500 and macroeconomic variables are shown in Figure 2. From the trends we can see that many look similar to one another. Some variables look almost identical (GDP, CPI, Personal income, Personal consumption expenses, etc). We can already see that multicollinearity will be affecting some of the variables. The other more "cyclical" variables seem interesting for making shorter term predictions.

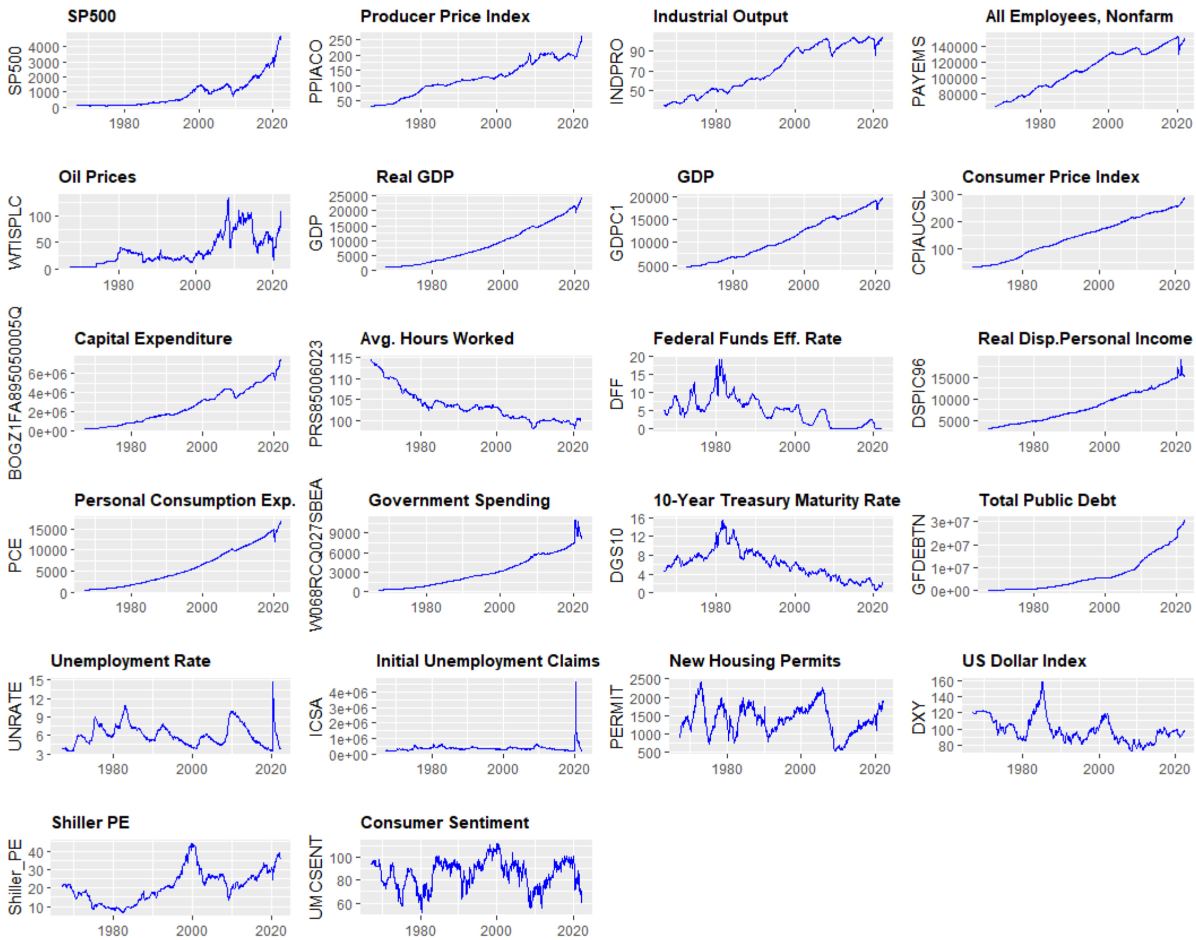


Figure 2: Time Series trends of SP500 and Macroeconomic variables

4.2 Correlation

Looking at the longer term trends, and also correlation between SP500 and the macroeconomic variables, the relationship seems pretty clear. During 1967 - 2022, there is a high correlation between SP500 and most of the variables (See Figure 3). When shortening the timeframe to quarter SP500 performance, and monthly SP500 performance compared to Macroeconomic variables, correlation becomes a bit weaker (See Figure 4).

Certain variables have stronger correlation with shorter time frames but not with longer time frame:

- Customer sentiment (UMCSENT)
- New housing permits (PERMIT)
- Unemployment claims (ICSA)

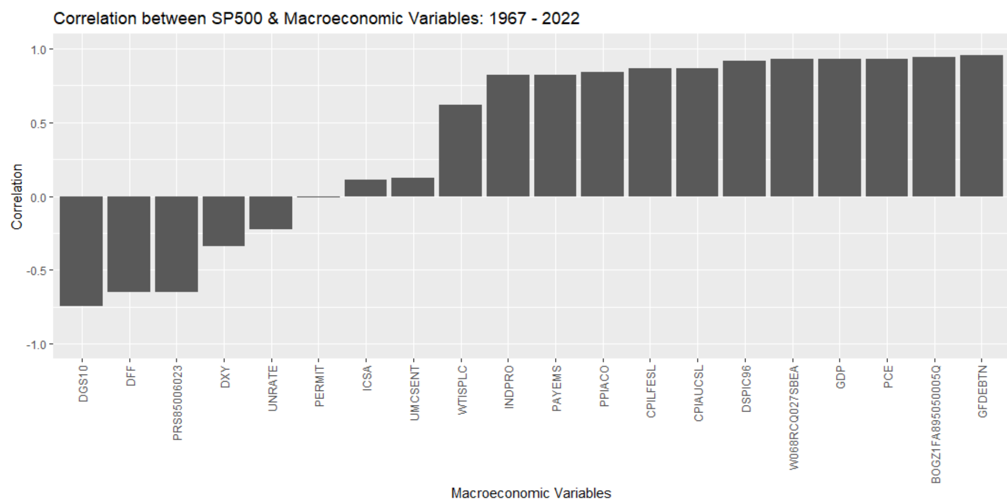


Figure 3: Correlation between SP500 & Macroeconomic variables: 1967-2022

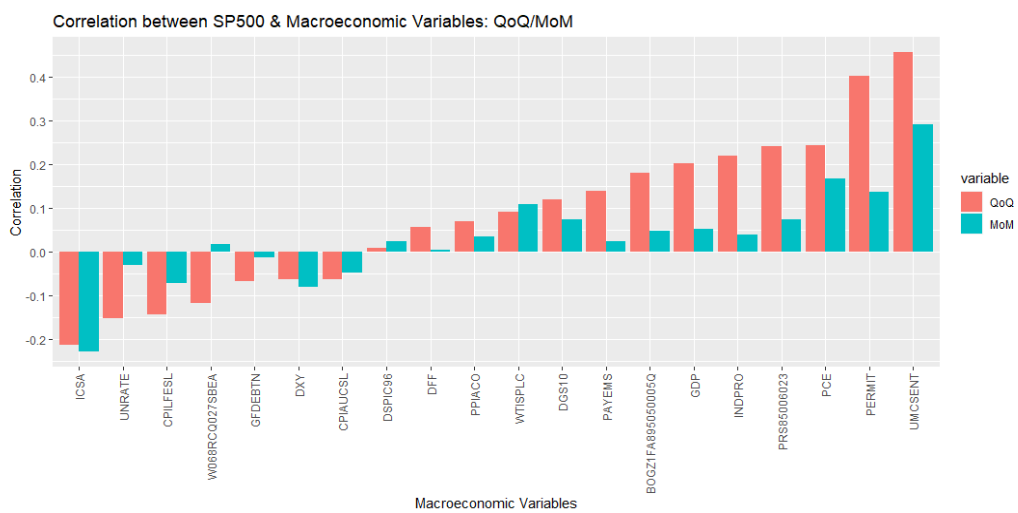


Figure 4: Correlation between SP500 & Macroeconomic variables: Quarter over Quarter (QoQ) and Month over Month (MoM).

4.3 Outliers detection

There are some outstanding points on the left of the chart with about -20% returns in month over month data, while quarter over quarter data shows potential data points on both tails, which also are verified in Q-Q plot, even though the data points in the middle are aligned straight to the line.

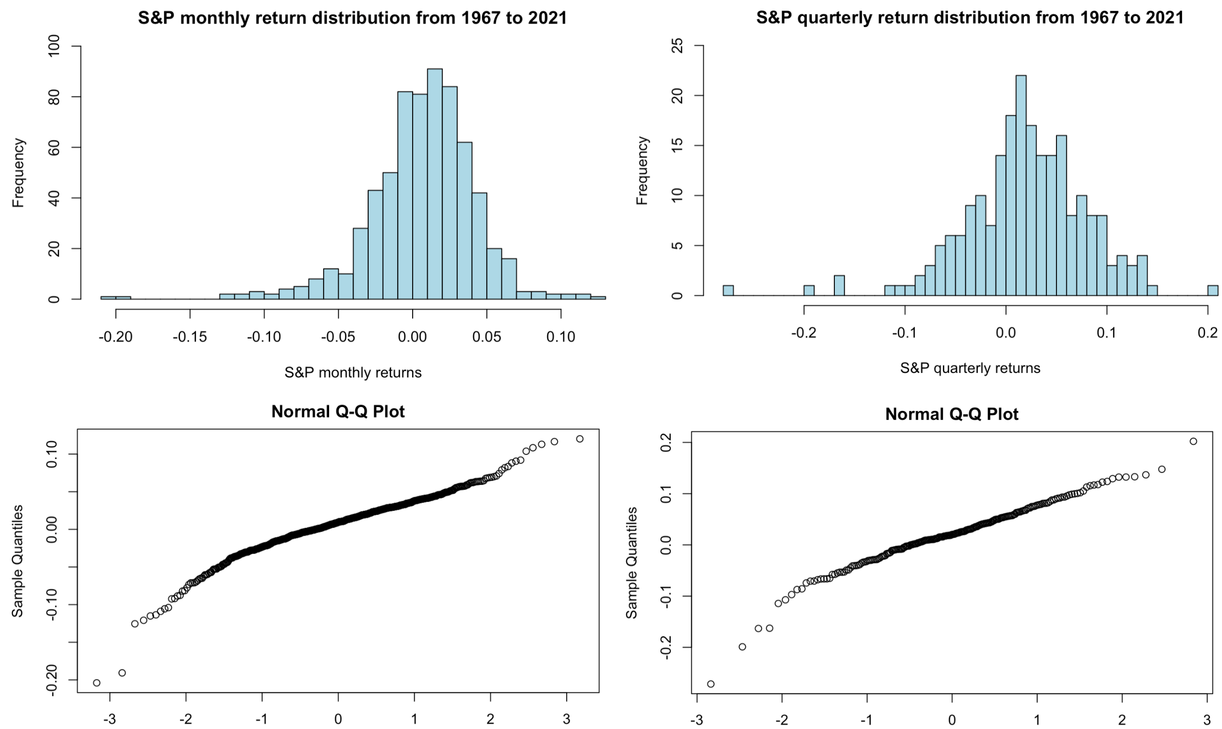


Figure 5: Histogram of SP500 Return and Q-Q Plots between SP500 & Macroeconomic variables: Quarter over Quarter (QoQ) and Month over Month (MoM).

Grubbs test is used for outlier detection. With a confidence interval of 95%, we find the below outliers in the datasets. One issue that arose was how we should handle outliers given that the datasets are made up for historical return and macroeconomic data; the outliers are not a result of user or measurement error. The below outliers were market crashes caused by Black Swan events, occurring in 1987 (Black Monday), 2008 (Subprime Mortgage Crisis), and 2020 (Covid-19). We ultimately decided not to remove the outliers because it is the study's goal to build a model that reflects the effects of macroeconomics in different economic conditions, even in extreme ones.

Date	**SP500**	**Dataset**
10/31/08	-0.204	MoM
3/31/20	-0.191	MoM
11/30/87	-0.1251	MoM
10/31/87	-0.1212	MoM
12/31/08	-0.272	QoQ
12/31/87	-0.1999	QoQ

4.4 Collinearity

In general, the quarterly data shows a stronger correlation between independent variables than the monthly dataset. From both correlation plots, PAYEMS, PCE and GDP have high correlation with other predictors. As shown in Variable Inflation Factor (VIF) values of each predictors, there are 5 and 9 predictors with VIF higher than 5 for the monthly dataset and the quarterly dataset respectively.

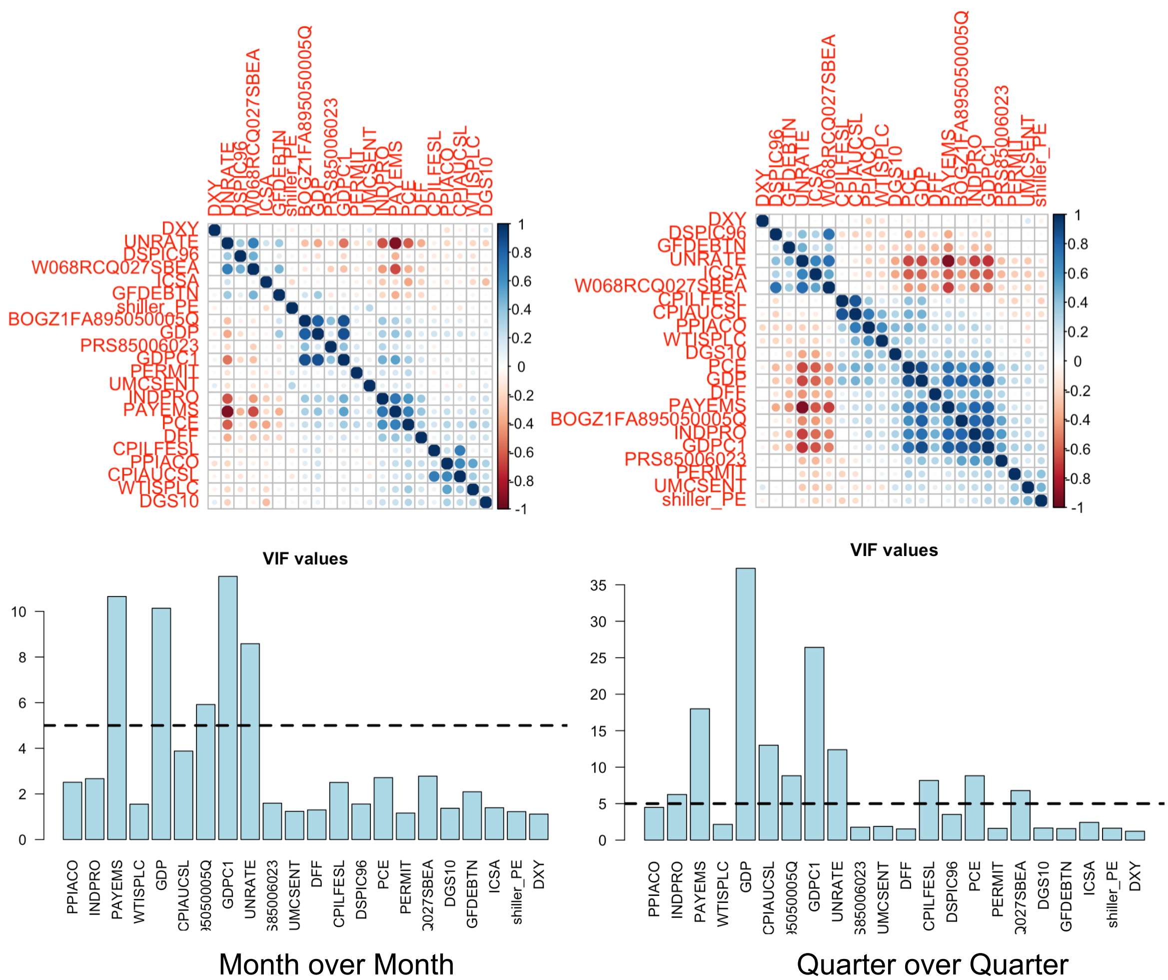
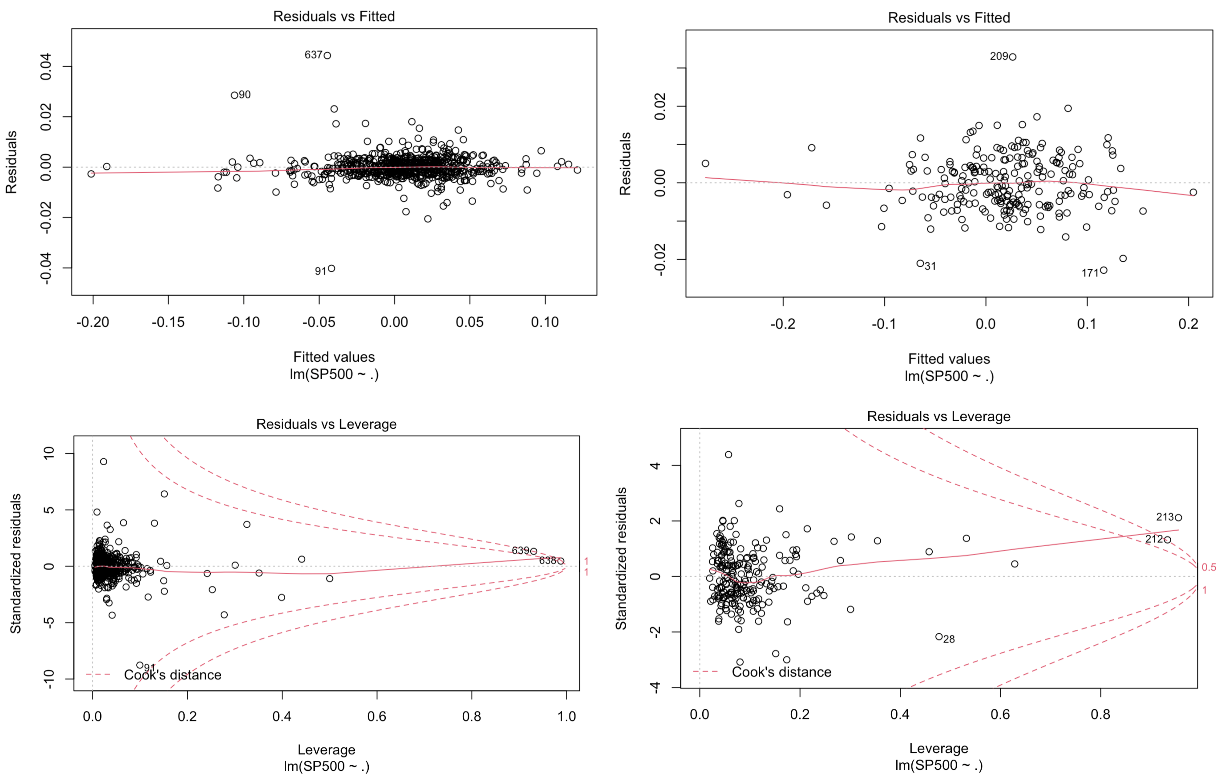


Figure 6: Correlation Plot for all variables (above). Variable Inflation Factor (VIF) values for independent variables (below). Month over Month is shown to the left. Quarter over Quarter is shown to the right.

4.5 Residual Study

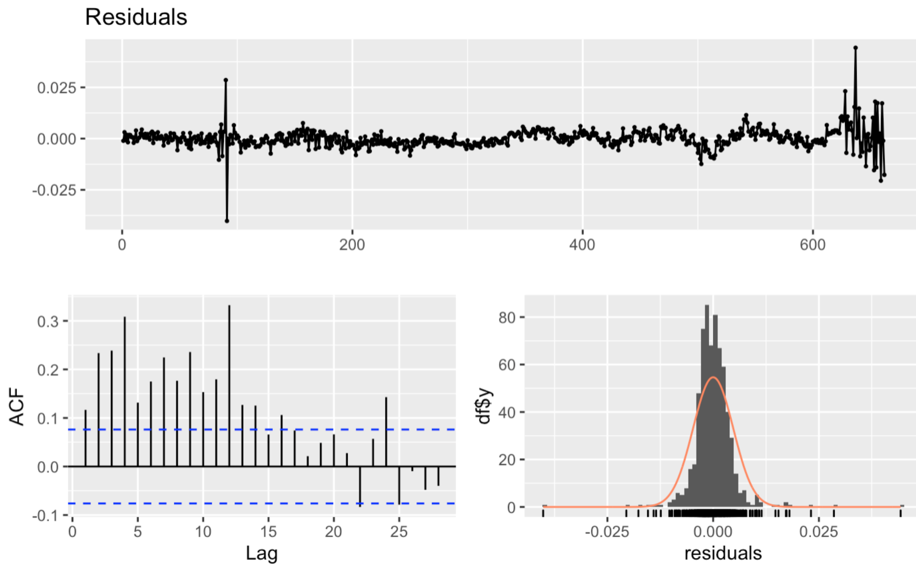
The full models of both Month over Month (MoM) and Quarter (QoQ) show a linearity relationship as residuals scatter on a horizontal line with fitted values with an average value 0. Variance of error from models trained by Month over Month (MoM) and Quarter (QoQ) datasets seems roughly constant. No outstanding high leverage points for the monthly dataset appear but there seems to be one high leverage point in the quarterly dataset as its cook's distance is higher than 1.



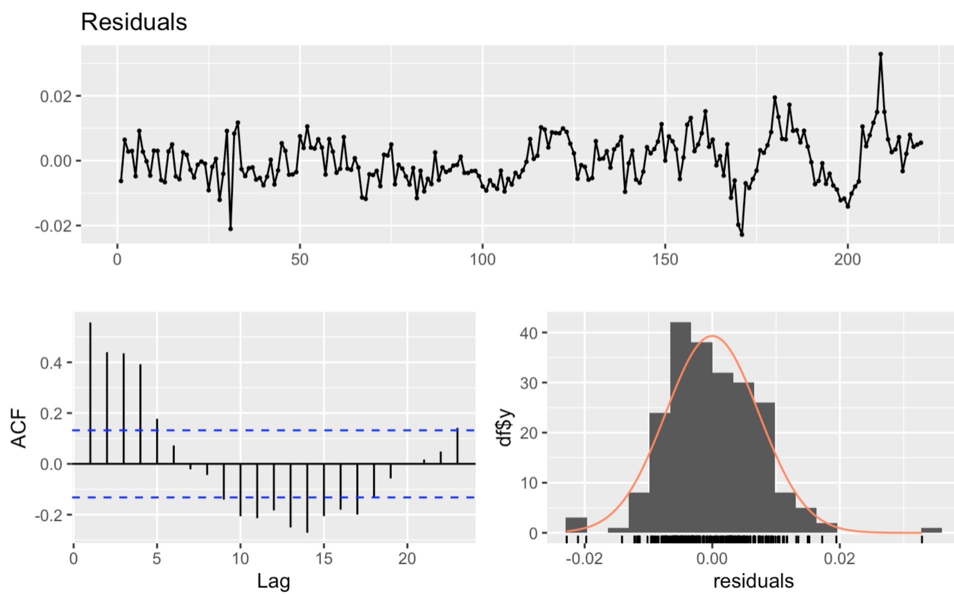
Month over Month Quarter over Quarter

Figure 7: Residual and Leverage Plots for Month over Month linear regression model is shown to the left. Residual and Leverage Plots for the Quarter over Quarter regression model are shown to the right.

As the linear model is built on a time series data, which is considered usually autocorrelated, the Durbin-Watson test is completed to test autocorrelation of error terms for full models of Month over Month (MoM) and Quarter over Quarter (QoQ) datasets. The DW of the monthly full model is 1.746, which is higher than the DW of 0.880 that is with the quarterly full model. This suggests that there is some information explained by predictors outside all the predictors in the full model, which is also supported by the ACF chart below, residuals lag chart and the Breucsch-Godrey test (refer to code file: exploratory data analysis.rmd for more details).



Month over Month



Quarter over Quarter

Figure 8: Residuals autocorrelation and distribution

The relationship between residuals and predictors are also studied by visualizing scatter plots of residuals and each predictors. For monthly data, PAYEMS, UNRATE, ICSA shows low correlation with residuals, while no pattern can be observed for other predictors and their variances remain constant. For quarterly data, PAYEMS, UNRATE, W068RCQ027SBEA, ICSA shows low correlation with residuals, while no pattern can be observed for other predictors and their variances remain constant. The scatter plots can be found in the exploratory data analysis markdown file.

5. Modeling and Analysis

5.1 Methodology Overview

To answer the question of “What is the relationship between the U.S. stock market and macroeconomic indicators for the United States?”, Team 11 modeled the relationship between the S&P 500 and macroeconomic predictors using the linear regression model. We completed model training and testing on two different sets of models: Month over Month (MoM) Linear Regression, Quarter over Quarter (QoQ) Linear Regression. Models were pre-processed with normalization (centering and scaling, as needed) and with and without Principal Component Analysis (PCA). For hyperparameter selection and training, we completed variable selection using stepwise regression and regularization techniques such as LASSO and ElasticNet.

We completed model training using train-test data split. We then evaluated each trained model using cross-validation to estimate each model’s R-squared value. We show the cross-validated model performance (R-squared values) for each set of models below:

Hyperparameters	**MoM cross-validated R Sq**	**QoQ cross-validated R Sq**
Full Model (least square estimation)	0.12	0.38
Stepwise Regression	0.10	0.22
ElasticNet	0.04	0.26
PCA	-0.46	0.23
PCA Stepwise	0.12	0.23
PCA ElasticNet	0.10	0.23
PCA Lasso	0.09	0.20

Table 1: Cross-Validated Model Performance (R-Squared) for Month over Month (MoM) and Quarter over Quarter (QoQ) Linear Regression Models.

The best performing model was the Quarter over Quarter linear regression model without PCA transformation or variable selection, even though the model is subject to multicollinearity problems. In general, we found that models trained using the Quarter-over-Quarter time interval dataset had higher predictive power than Month-over-Month time interval dataset models. Although models exhibited overfitting and multicollinearity, transforming the data set using principal component analysis did not improve fit. We also found that variable selection did not improve model performance, as the full linear regression model (without variable selection and without principal component analysis) had the highest R squared value for both month over month and quarter over quarter data.

We go into further detail for our modeling in the sections below. We discuss the resulting models, with and without principal component analysis and with and without variable selection.

5.1 Monthly Analysis

5.1.1 Month over Month (MoM)

We began by building a full linear regression model on the Month over Month data set, resulting in a 0.1845 R Squared value.

To try and improve the model performance, we attempted variable selection using elastic net regression, LASSO, and stepwise regression.

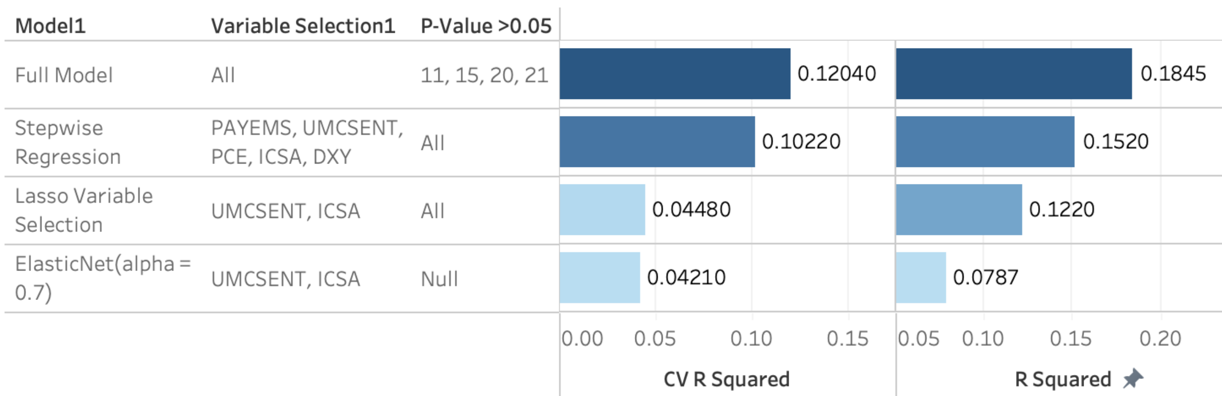


Figure 9: R squared and cross-validated R-Squared for month over month models

The best performing stepwise regression model resulted in selection of a 5-variable model. Both LASSO and ElasticNet regression models result in selection of a 2 variable model. In all three cases, we observe lower model performance and fit than the full model, as shown above. All three have comparatively lower cross-validated R-squared values. The poor model performance may be due to the monthly time interval data set being much noisier and exhibiting high variance.

5.1.2 Monthly Analysis with PCA

We used PCA transformation to help address the multicollinearity of the model. The below figure shows that increase in explained variance by number of principal components slows down at 5 and becomes even slower at 15. 5 principal components can explain about 60% of the variance and 15 principal components for more than 90%. Thus, 5 and 15 principal components are the number of PCs we try at first to fit a linear regression model. We also use backward stepwise, lasso and elasticNet for variable selection.

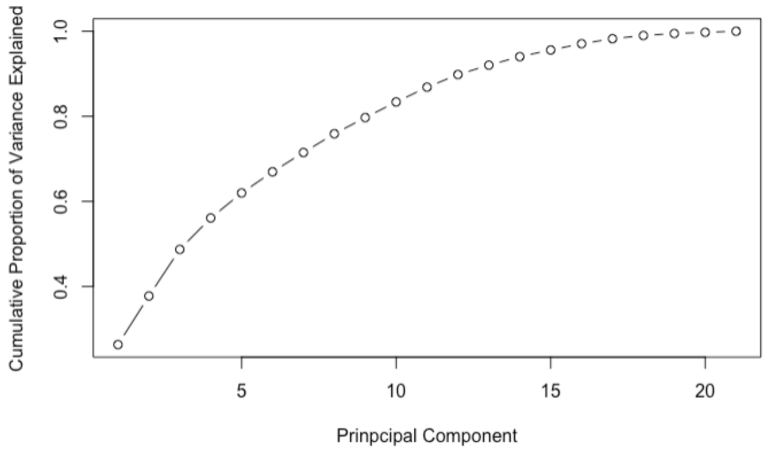


Figure 10: principal components and cumulative proportion of variance explained

The variable selection models lead to a better fit, when compared with the models built with the most variance-explained principal components. However, all of the models exhibit overfitting to some extent.

Model1	Variable Selection1	P-Value >0.05	CV R Squared	R Squared
First 5 PCs(60% variance explained)	1,2,3,4,5	2,3,5	-0.0178	0.06912
First 15 PCs(90% variance explained)	1,2,3,4,5,6,7,8,9,10, 11,12,13,14,15	2,3,5,11,13, 14,15	-0.4562	0.15380
PCA+stepwise	1,4,5,6,7,8,9,10,11, 12,16,19	5,11,19	0.1190	0.16320
PCA+ElasticNet(alpha= 0.1)	1,4,5,6,7,8,9,10,11, 12,13,16,19	5,13,19	0.1027	0.16420
PCA+Lasso	1,4,6,7,8,9,10,12	Null	0.0895	0.14950

Figure 11: R squared and cross-validated R squared of PCA transformation MoM Models

5.1.3 Monthly Analysis Summary

In general, month over month linear regression model fit was poor. The linear regression model without PCA transformation or variable selection model performed best., with a cross-validated R squared value of 0.12. Variable Selection and Principal Component Analysis did not result in a better cross-validated model fit compared with the full model. Next, we look at modeling using the Quarter over Quarter data set.

5.2 Quarterly Analysis

5.2.1 Quarter over Quarter (QoQ)

We first started by building a full linear regression model on the Quarter over Quarter data set, resulting in a 0.38 R2 value.

To try and improve the model performance, we attempted variable selection using elastic net regression and stepwise regression. With the forward direction in the stepwise regression, we yielded a resulting model of two predictors: Customer Sentiment, and New Privately-Owned Housing Units Authorized in Permit-Issuing Places. The stepwise regression in the backward direction yielded a model with eight predictors: Urban Consumer Price Index, Capital Expenditure, Real GDP, Unemployment Rate, Consumer Sentiment, Urban Consumer Price Index (Less Food & Energy), and New Privately-Owned Housing Units Authorized in Permit-Issuing Places. Doing the stepwise regression for both directions we achieved the same model as the forward direction model. Our Elastic Net model yielded a model with the four predictors, Real GDP, average weekly hours worked for all employed persons in the non-farm business sector, New Privately-Owned Housing Units Authorized in Permit-Issuing Places, and Consumer Sentiment.

Method	Predictors	R2	CV-R2
Full Model	ALL	0.36	0.38
ElasticNet	GPDC1, PRS85006023, UMCSENT, PERMIT	0.26	0.26
Stepwise (Forward)	UMCSENT, PERMIT	0.28	0.22
Stepwise (Backward)	CPIAUCSL, BOGZ1FA89505005Q, GPDC1, UNRATE, UMCSENT, CPILFESL, PERMIT	0.35	0.34

Table 2: Quarter over Quarter (QoQ) Variable Selection Model Performance Summary

The Consumer Sentiment was shown to have the strongest effect on the data, both monthly and quarterly, having a coefficient of 0.2368 for the quarterly model. The average weekly hours worked for all employed persons in the non-farm business sector had the smallest effect with a coefficient of 0.1570. Overall, the variable selection did not improve model performance but Quarterly data was able to obtain a model better fit than the Monthly data.

5.2.2 Quarterly Analysis with PCA

We then pre-processed the data with Principal Component Analysis to reduce multicollinearity. Data is first transformed using PCA, and then we iterate over the components to see how many PCs are needed for best results with linear regression. PCAs for quarterly data, and r-squared using 5-fold training and 5 times repeated cross-validated r-squared values are shown in the figure below.

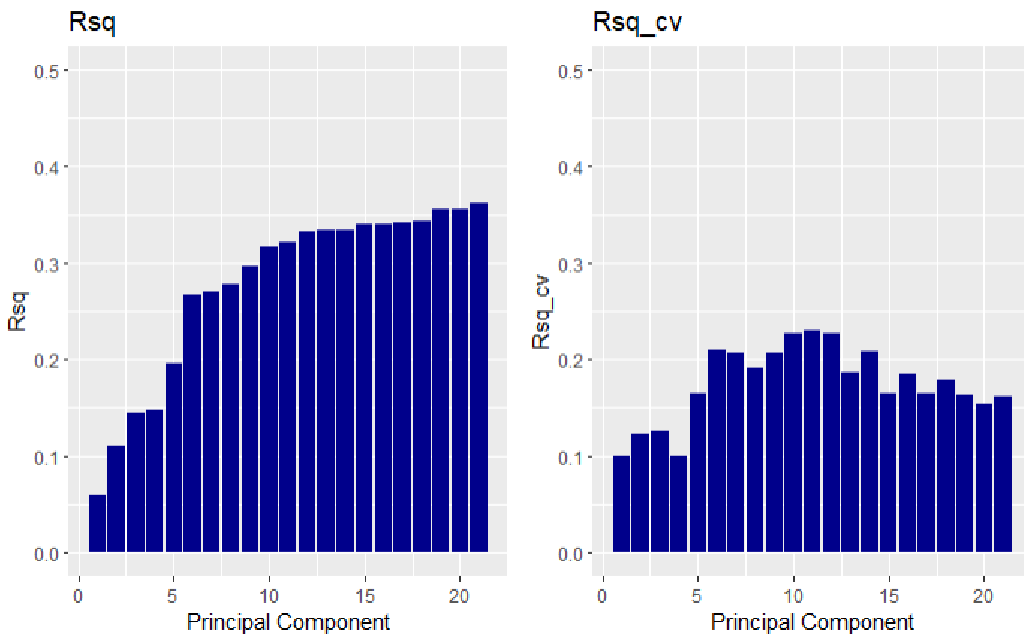


Figure 12: R-sq vs. principal components and cumulative proportion of variance explained

We achieve the highest r-squared value of 0.23 with 11 components, but in general results don't improve after the first 7 principal components.

We also tested PCA together with Lasso Regression, Elastic Net and Stepwise Regression to see if the model could further be improved.The highest r-squared value after cross-validation remained at 0.23, and was achieved with 11 components and with Stepwise Regression.

Method	PCs	p-value < 0.05	R2	R2 - CV
PCA	1, 2, 3, 4, 5, 6, 7,8,9,10,11	1, 2, 3, 5, 6	0.32	0.23
PCA + Lasso Regression	1, 2, 3, 5, 6	All	0.27	0.20
PCA + Elastic Net	1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 15, 19, 21	1, 2, 3, 5, 6, 19	0.35	0.22
PCA + Stepwise Regression	1, 2, 3, 5, 6, 8, 9, 10, 12, 15, 19	1, 2, 3, 5, 6, 19	0.34	0.23

Table 3: Quarter over Quarter (QoQ) PCA Model Performance

5.2.3 Quarterly Analysis Summary

As discussed in our methodology overview section, the Quarter-over-Quarter full model was the best performing model. Similar to our month-over-month modeling, PCA and variable selection did not improve model performance for the Quarter-over-Quarter data set.

5.3 Does market price lead or lag in reaction to macroeconomic predictors?

As part of this study, we decided to look at whether or not the market lags or leads macroeconomic data. In the market, there are investors who price in their expectation - for instance, if they think the economy is doing badly, they may decide to sell their assets early, anticipating a selloff or negative market returns. In other words, they bet on the direction of the market index based on their macroeconomic expectations beforehand. As a result, the S&P 500 index may show the expectations of investors. Alternatively, it may be possible instead that it takes time for markets to interpret and react to macroeconomic data. To study whether the market is more likely to price in expectations or have a delayed reaction to macroeconomic data, we shift S&P 500 index by a range of [-12,12] months relative to the independent variables to examine the change of variance explained. We then trained a full model and a model with five of the most significant predictors.

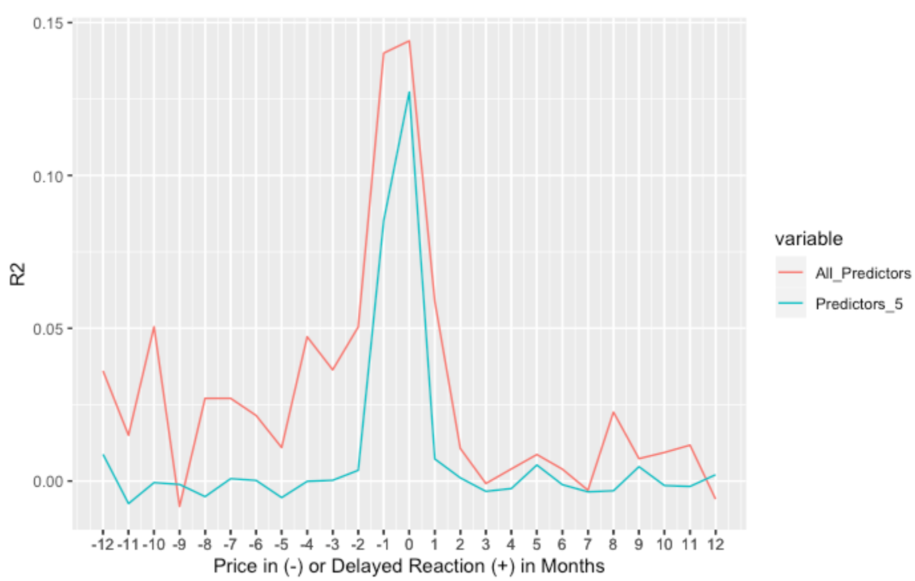


Figure 13: R squares of leading and lagging models

Examining the model performance, the market does not seem to lag because of the poor R2 of lagging models. As the best R-Squared is 0 months, the market seems to be efficient and reflects current value. It is possible that the market may price in changes in the macroeconomic environment early because the R-Squared of -1 month is close to the best R2 value.

5.4 Is there a better model fit when smaller time frames are used?

Earlier modeling used the full range of the data set between 1967-2022 to examine model performance and fit. Our team wanted to evaluate the effectiveness of building models for select timeframes in the dataset. To start, we developed models for 11 five-year bins of the dataset using all parameters. These models were inconsistent. Some had strong correlations, but most of the parameters in any given model were insignificant as demonstrated by the p-values. Additionally, there were no parameters which were significant across all models. The predictors with the lowest average p-value across the five-year bins were DFF, DGS10, DXY, PAYEMS, and UMCSENT. The below plot depicts the parameter values as they changed across the five-year bins. Each of the lines corresponds to a parameter in the model.

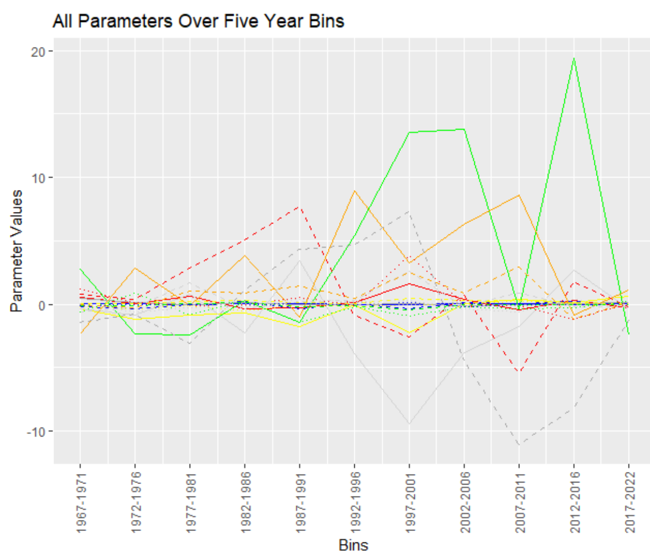


Figure 14: Parameter Values for Five-Year Bin Models (All Parameters)

	Blue	Red	Yellow	Green	Dark Gray	Light Gray	Orange
Solid Line	Intercept	PPIACO	INDPRO	PAYEMS	WTISPLC	GDP	GDPC1
Dashed Line	UNRATE	PRS85006023	UMCSENT	DFF	CPIFESL	DSPIC96	PCE
Dotted Line	PERMIT	W068RCQ27SBEA	DGS10	GFDEBTN	ICSA	DXY	

Table 4: Color and Line types for all Parameter Values in Figures 14-17

Next, we evaluated the dataset based on models fitted to 6 ten-year bins, with similar results. The predictors with the lowest average p-value across the ten-year bins were DSG10, DXY, PCE, GDPC1, and UNRATE. The below plot depicts the parameter values in each of the bins. Each line corresponds to a parameter in the model.

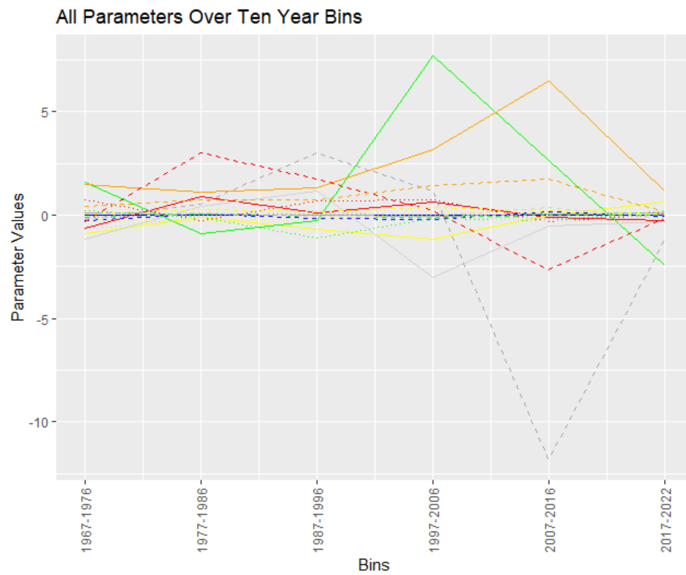


Figure 15: Parameter Values for Ten-Year Bin Models (All Parameters)

Both of the initial binned models were at risk for overfitting and struggled with large numbers of insignificant parameters. We tried to counteract this by testing models with only the most significant parameters and models with variables chosen by the stepwise regression procedure which improved model fit on the unbinneded dataset.

In testing only the most significant parameters, our team kept at least one parameter in each model, but no more than four. The resulting models had weaker correlations (see below table) with the provided data and many of the parameters became insignificant. The chosen parameters were also inconsistent across the models since there were no parameters that were significant in all of the five-year or ten-year bins. These models were scrapped as they were poor representations of the data.

Given the poor model quality with the models built on the most significant parameters, we decided to test the parameters selected by stepwise regression (run on the entire dataset) on the binned datasets. The parameters chosen were DFF, DXY, ICSA, PCE, and UMCSENT. In both the five-year and ten-year bin models, the correlation was weaker than from the binned models with all parameters. Also, the parameters were frequently insignificant. Below are similarly styled plots from previously, depicting the change in parameter values across the bins in the five-year and ten-year models.

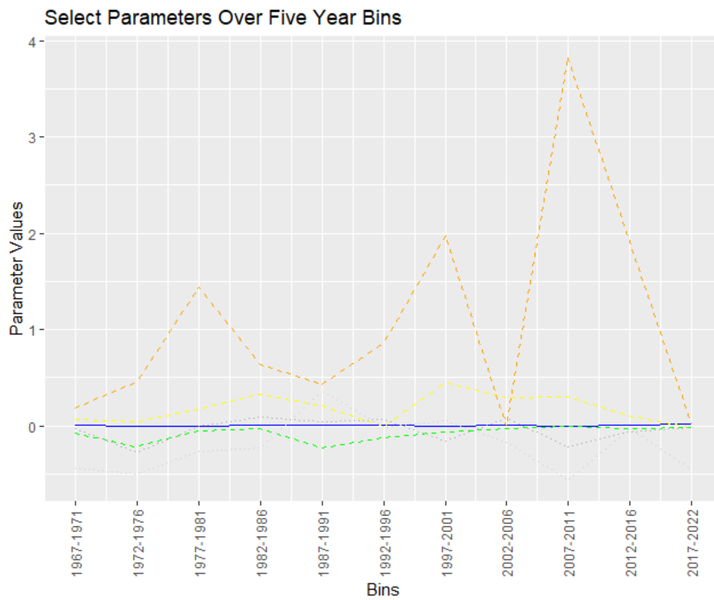


Figure 16: Parameter Values for Five-Year Bin Models (Select Parameters)

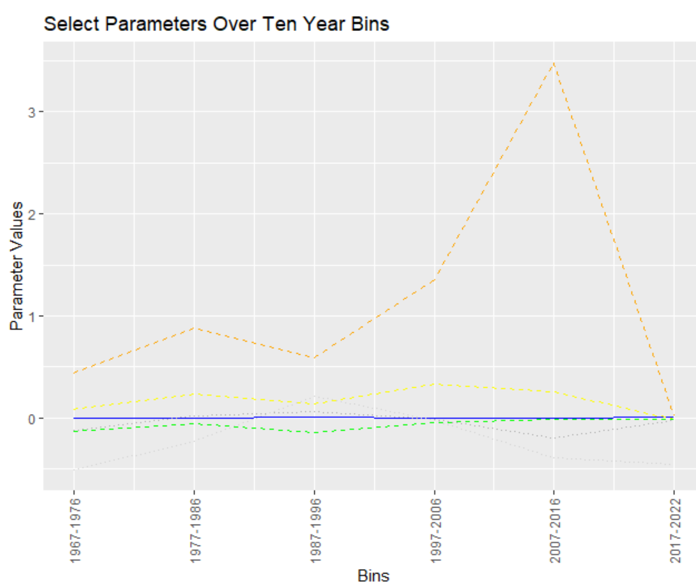


Figure 17: Parameter Values for Ten-Year Bin Models (Select Parameters)

Lastly, since the models with all the parameters had the best fits, but were potentially over fit. We ran a cross validation procedure on the binned models with all parameters. The resulting r-squared values for the five-year bins were very low and indicated that any correlations in the initial models were due to over fitting. The r-squared values for the ten-year bin models were roughly the same as in the initial analysis indicating that the models were not over fit and that the models were fair representations of the data. Below are tables depicting the r-squared values of all the binned models.

	Bin 1 (1967 -1971)	Bin 2 (1972 -1976)	Bin 3 (1977 -1981)	Bin 4 (1982 -1986)	Bin 5 (1987 -1991)	Bin 6 (1992 -1996)	Bin 7 (1997 -2002)	Bin 8 (2003 -2006)	Bin 9 (2007 -2011)	Bin 10 (2012 -2016)	Bin 11 (2017 -2022)
Initial Model	0.306	0.595	0.402	0.645	0.423	0.590	0.470	0.473	0.726	0.567	0.649
Significant Parameter Model	0.041	0.205	0.165	0.487	0.005	0.428	0.228	0.249	0.461	0.184	0.516
Select Parameter Model	0.041	0.316	0.267	0.173	0.199	0.105	0.252	0.234	0.483	0.121	0.549
Initial Model (Cross Validation)	0.005	0.057	0.168	0.419	0.130	0.142	0.294	0.018	0.108	0.000	0.005

Table 5: Five-Year Bin Models Performance (R-Squared)

	Bin 1 (1967-1976)	Bin 2 (1977-1986)	Bin 3 (1987-1996)	Bin 4 (1997-2006)	Bin 5 (2007-2016)	Bin 6 (2017-2022)
Initial Model	0.348	0.406	0.344	0.306	0.576	0.649
Significant Parameter Model	0.170	0.292	0.201	0.197	0.417	0.516
Select Parameter Model	0.160	0.200	0.125	0.188	0.362	0.549
Initial Model (Cross Validation)	0.310	0.436	0.351	0.358	0.527	0.753

Table 6: Ten-Year Bin Models Performance (R-Squared)

6. Conclusion

Model fit for both month-over-month and quarter-over-quarter linear regression models provide no or weak support for our hypothesis that there is a causal relationship between macroeconomic predictors and the performance of the stock market. The relationship is complex and it is not possible to rule out other factors that may cause over or under-performance.

We found that the models built on quarter over quarter performed significantly better than the counterpart of month on month: using a high time frame interval for macroeconomic data was more predictive of the market return. Our best performing model had a cross-validated R-Squared value of 0.38; in other words, macroeconomic factors were able to explain nearly 38% of variation in the S&P 500 market's returns between 1967-2022. This suggests that firms and traders should not wholly discount the importance of the economy in the stock market's performance.

One limitation of this study is that we focused on modeling the data on only the S&P 500 stock market index return. Future studies could model macroeconomic factors against stock market return for other indices, especially international stock markets (such as developed or emerging markets). Another limitation is that it is also possible that our study did not incorporate important macroeconomic factors; we were limited to data that we could find publicly available.

6.1 Future Areas of Study:

Due to the limited duration of this study, we were not able to look at fitting the data against other types of models. For future research, we would be interested in modeling the data using the following models:

- ARIMA regression model. Incorporating an auto-regressive component into the regression model may improve overall fit given that the financial data is time series data.
- GARCH model This model would look at total variability instead of total return.
- Classification model for market direction. We could use this model to classify the next forecasted period as either positive or negative return.
- Regression spline modeling. Binning the data resulted in some models with extremely low R-squared values and extremely high R-squared values. This suggests that a non-linear modeling approach could possibly improve model fit.

One area of study that we were not able to investigate thoroughly in our study due to our models' poor performance was which predictors were most important or significant. Importantly, we note that the importance or statistical significance of a predictor depended on which time interval we were investigating. We would be interested in investigating what predictors may be most important or significant during different market conditions. For example, Merrill Lynch's Investment Clock model is built on the hypothesis that market returns depend on what phase of the business cycle we are in (assuming that market returns are cyclical), especially by factors such as inflation and economic growth.

Appendix 1: Data Points

Predictors	start	end	Interval	API	quarters	FRED Code	Var_name
PPI	1/1/1913	5/1/2022	Monthly	FRED	444	PPIACO	PPIACO
Industrial output	1/1/1919	5/1/2022	Monthly	FRED	419	INDPRO	INDPRO
S&P 500	12/30/1927	6/22/2022	daily	CSV WSJ+kaggle	383		SP500
Total Nonfarm Payrolls	1/1/1939	5/1/2022	Monthly	FRED	338	PAYEMS	PAYEMS
Oil Prices ($/barrel crude oil)	1/1/1946	5/2/2022	Monthly	FRED	310	WTISPLC	WTISPLC
GDP	1/1/1946	1/1/2022	Quarterly	FRED	308	GDP	GDP
Consumer Price Index for All Urban Consumers: All Items	1/1/1947	5/1/2022	Monthly	FRED	306	CPIAUCSL	CPIAUCSL
Capital Expenditure	10/1/1946	1/1/2022	Quarterly	FRED	305	BOGZ1FA895050005Q	BOGZ1FA895050005Q
Unemployment rate	1/1/1948	5/1/2022	Monthly	FRED	302	UNRATE	UNRATE
Nonfarm Business Sector: Average Weekly Hours Worked for All Employed Persons	1/1/1948	1/1/2022	Quarterly	FRED	300	PRS85006023	PRS85006023
Consumer Sentiment (Index)	11/1/1952	5/1/2022	weekly	FRED	282	UMCSENT	UMCSENT
Effective Federal Funds Rate	7/1/1954	6/24/2022	daily	FRED	276	DFF	DFF
Consumer Price Index for All Urban Consumers: All Items Less Food & Energy	1/1/1957	5/1/2022	Monthly	FRED	265	CPILFESL	CPILFESL
Real Disposable Personal Income	1/1/1959	4/1/2022	Monthly	FRED	257	DSPIC96	DSPIC96
Personal Consumption Expenditures	1/1/1959	4/1/2022	Monthly	FRED	257	PCE	PCE
New Privately-Owned Housing Units Authorized in Permit-Issuing Places: Total Units	1/1/1960	5/1/2022	Monthly	FRED	253	PERMIT	PERMIT
Government Spending	1/1/1960	1/1/2022	Quarterly	FRED	252	W068RCQ027SBEA	W068RCQ027SBEA
10-Year Treasury Constant Maturity Rate	1/2/1962	6/24/2022	daily	FRED	245	DGS10	DGS10
Total Public Debt	1/1/1966	1/1/2022	Quarterly	FRED	227	GFDEBTN	GFDEBTN
Initial claim	1/7/1967	6/18/2022	weekly	FRED	225	ICSA	ICSA
US Dollar/USDX - Index	1/31/1967	6/21/2022	daily	CSV yahoo finance	225		DXY

Appendix 2: Works Cited

Alshogeathri, Mofleh (2011). “Macroeconomic Determinants of the Stock Market Movements: Empirical Evidence from the Saudi Stock Market”. Ph.D., Kansas State University.

Comincioli, Brad (1996) "The Stock Market As A Leading Indicator: An Application Of Granger Causality," University Avenue Undergraduate Journal of Economics: Vol. 1 : Iss. 1, Article 1. https://digitalcommons.iwu.edu/uauje/vol1/iss1/1

Fama, Eugene (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". Journal of Finance. 25 (2): 383–417. doi:10.2307/2325486. JSTOR 2325486.

Gordon, M.J and Eli Shapiro (1956) "Capital Equipment Analysis: The Required Rate of Profit," Management Science, 3,(1) (October 1956) 102-110. Reprinted in Management of Corporate Capital, Glencoe, Ill.: Free Press of, 1959.

Markowitz, H.M. (March 1952). "Portfolio Selection". The Journal of Finance. 7 (1): 77–91. doi:10.2307/2975974. JSTOR 2975974.

Ross, Stephen A (1976-12-01). "The arbitrage theory of capital asset pricing". Journal of Economic Theory. 13 (3): 341–360. doi:10.1016/0022-0531(76)90046-6. ISSN 0022-0531.

Sharpe, William F. (1964). "Capital asset prices: A theory of market equilibrium under conditions of risk". Journal of Finance. 19 (3): 425–442. doi:10.1111/j.1540-6261.1964.tb02865.x. hdl:10.1111/j.1540-6261.1964.tb02865.x.