SDSC4107 Group Project Report

Exploring the relationship between macroeconomic indicators and the stock market

Group 10
Jin yuheng 55966786
CHAN Yuk Yee 56230549
WANG Yixiao 57126744

Table of contents

1. Introduction

1.1. Background

The impact of the economic crisis on the national stock market is very significant, and we observed huge fluctuations in stock markets and macroeconomic data worldwide during the financial crisis in 2008. This inspired our curiosity to explore the relationship between macroeconomics and the stock market. And CPI and GDP are both important macroeconomic data, and we find that fluctuations in their data may be related to the stock market. So we chose to study the relationship between GDP, CPI, and the stock market.

1.2. Dataset description

We selected the three largest economies in the world, the United States, China, and Japan for our study. And selected the most representative stock market indices of these three countries, the S&P 500 Index, Shanghai Stock Exchange Index, and Nikkei Index. For the data set, We collected the quarterly data of GDP, CPI, and stock indices for the US, China, and Japan from 2001 to 2022.

2. Methodology

2.1. Augmented Dickey-Fuller test

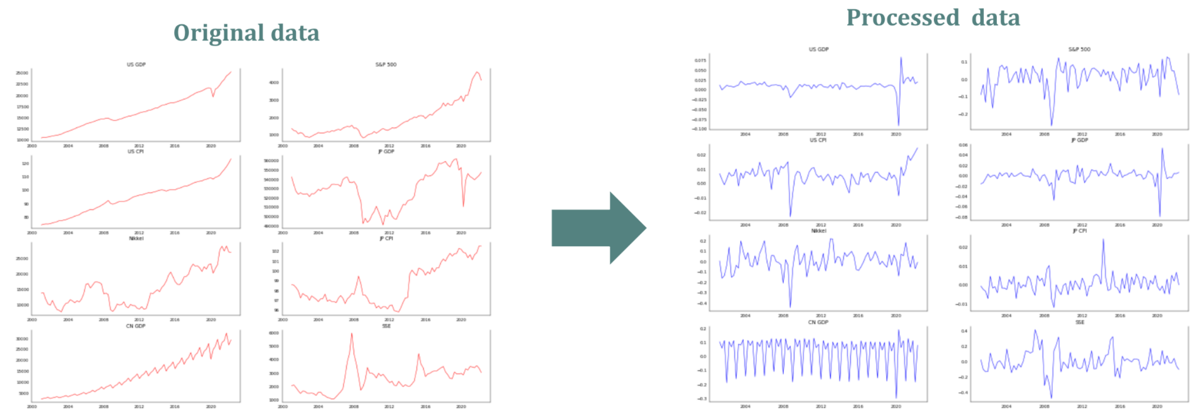
The ADF test is a statistical test that determines whether or not a time series is stationary. It is an extension of the Dickey-fuller test, which looks for a unit root in a times series. If the series has a unit root, it is not stationary. To account for potential autocorrelation in the data, the ADF test includes additional lag terms in the test regression.

The ADF tests work by estimating the following regression equation:

$$\Delta yt = \alpha + \beta t + \gamma yt-1 + \delta 1\Delta yt-1 + \ldots + \delta p\Delta yt-p + \varepsilon t$$

The Δyt is the first difference of the time series yt, α is the intercept term, β is the coefficient of a linear trend, and εt is the error term. The lagged differences $\Delta yt-1$, $\Delta yt-2$, ..., $\Delta yt-p$ are included to control for potential autocorrelation in the data.

If the p-value is less than a chosen significance level, like 0.05 usually means the series is stationary(doesn't have a unit root)



2.2. Granger Causality Test

The Granger Causality Test A statistical concept that uses the hypothesis to determine whether one time series can be used to forecast another. Designed to determine whether one time series can be used to predict another. Clive Granger designed the Autoregressive model in 1969 to give us a type of idea that uses the autoregressive model can be available on the prediction of different variables can be used to detect time series X can "determine" the time series Y. In other words, if one time series X "Granger-causes" another time series Y.

The autoregression model includes two stationary time series, X and Y. We must first determine whether time series X Granger-cause Y by allowing the appropriate lag value of y to be included in the univariate autoregression of Y:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_m y_{t-m} + residual_t.$$

The lag value of yt that may influence the yt value is represented by the Y range from y1 to yt-m, and we will import the X value into the time series in the bivariate auto-regression courses. The Granger Causality Test does not demonstrate a true cause and effect chain.

Figure 2. Granger Cause Test

US CPI and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:21
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
S_P does not Granger Cause USCPI	83	3.69264	0.0154
USCPI does not Granger Cause S_P		4.62711	0.0050

JP CPI and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:22
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
JPCPI does not Granger Cause NIKKEI	83	2.29571	0.0845
NIKKEI does not Granger Cause JPCPI		4.96039	0.0034

CN CPI and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:23
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
CNCPI does not Granger Cause SSE	83	2.35304	0.0788
SSE does not Granger Cause CNCPI		3.55028	0.0183

US GDP and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:18
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
S_P does not Granger Cause USGDP	83	3.53511	0.0186
USGDP does not Granger Cause S_P		1.23216	0.3039

JP GDP and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:22
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
JPGDP does not Granger Cause NIKKEI	83	2.55999	0.0612
NIKKEI does not Granger Cause JPGDP		1.59980	0.1964

CN CPI and stock

Pairwise Granger Causality Tests
Date: 11/20/22 Time: 17:23
Sample: 2001Q1 2022Q2
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Prob.
SSE does not Granger Cause CNGDP	83	0.53603	0.6590
CNGDP does not Granger Cause SSE		1.56122	0.2057

On each variable, we ran Granger causality tests. We can conclude that a variable is Granger Cause with another series when the P-value is less than 5%. Figure 2 depicts the situation. The P-values of the key results (for example, SSE does not Granger Cause CNCPI is 0.0183) are all less than 5%. As a result, we were able to determine which time series pass the Granger Causality Test.

2.3. Cointegration tests

The cointegration test is used to establish a long-term correlation between several time series. Clive Granger and Robert Engel pioneered the concept in 1987.

The cointegration test identifies scenarios in which at least two non-stationary time series are integrated together in such a way that they cannot achieve long-term equilibrium. The tests are used to identify two or more variables that are sensitive to one another over a specific time period. This becomes a very useful nonlinear regression test for determining the relationships between X and Y over time.

The cointegration test includes 2 main test methods:

1. The Two-Step Engle-Granger Method

The Engle-Granger two-step method begins with the creation of errors based on the static regression and then tests the errors. To examine stationarity units X in a time series, we will typically use the Augmented Dickey-Fuller Test (ADF) or other tests. If the time series are cointegrated, the Engle-Granger method will yield residual stationarity.

The Engle-Granger method has the disadvantage of revealing more than two cointegration relationships if there are more than two variables. Another limitation of the Engle-Granger method is that it is a single equation model. However, some of the disadvantages have been addressed in recent cointegration tests such as Johansen. STAT or MATLAB software can be used to calculate the Engle-Granger test.

.

2. Johansen Test

The Johansen test was created to address a limitation of the Engle-Granger method; it can cointegrate the relationships between multiple non-stationary time series data. The Johansen test, unlike the Engle-granger test, can detect multiple cointegration relationships. However, it is subject to asymptotic properties because a small sample size may result in unreliable results. Using the test to determine the cointegration of multiple time series can help to avoid the problems that arise when errors are carried forward to the next step.

.

```
Name   ::  Test Stat > C(95%)    =>   Signif
  ----------------------------------------
US GDP ::  41.57     > 24.2761   =>   True
US CPI ::  14.23     > 12.3212   =>   True
S&P 500 ::  0.04      > 4.1296    =>   False

Name   ::  Test Stat > C(95%)    =>   Signif
  ----------------------------------------
JP GDP ::  50.39     > 24.2761   =>   True
JP CPI ::  26.86     > 12.3212   =>   True
Nikkei ::  8.5       > 4.1296    =>   True

Name   ::  Test Stat > C(95%)    =>   Signif
  ----------------------------------------
CN GDP ::  45.4      > 24.2761   =>   True
CN CPI ::  14.52     > 12.3212   =>   True
SSE    ::  1.33      > 4.1296    =>   False
```

Here we performed Cointegration tests on each variable. There may not be a long-run equilibrium relationship between the S&P 500 and the other two series - CPI and GDP. There is a long-term equilibrium relationship among the three series of Japan's GDP, CPI, and Nikkei. There may not be a long-term equilibrium relationship between the Shanghai Composite Index and the other two series - CPI and GDP.

2.4. Vector autoregression Model

VAR is a popular time series forecasting model for analyzing multiple time series data that are interdependent. The model captures interactions between variables over time.

The dataset used in this study is divided into two parts: training data and testing data. This is done to train the model using the training data and evaluate its performance on unseen data using the test data.

A VAR model is used to predict the last 3 observations or 3 seasons from the training data. The actual values in the test data were compared to these predicted values.

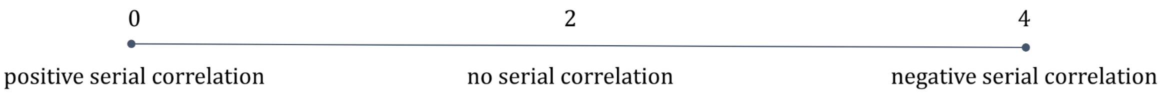
To evaluate the accuracy of the forecasts, several forecast accuracy metrics are used. These metrics can evaluate the performance of a VAR model by comparing predicted values with actual values. Use multiple metrics to get a more complete picture of your model's strengths and weaknesses.

3. Result

3.1. Durbin Watson Test

To validate the ability of our model to explain variance and patterns in the time series, we performed a Durbin-Watson test to examine the serial correlation of the residuals. A positive correlation indicates that our model cannot explain some patterns in the time series, thus indicating the need to enhance the order of the model or introduce more predictors, or use a different modeling algorithm.

The Durbin-Watson statistic is between 0 and 4, with a value of 2 indicating no significant serial correlation. If the value is close to 0, there is a positive correlation, while a value close to 4 means a negative correlation.



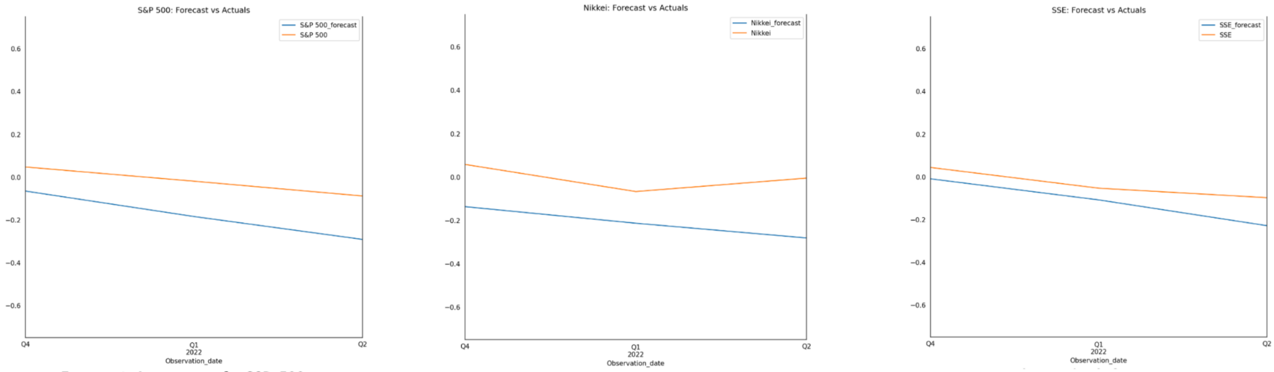

```
US GDP : 2.01
US CPI : 1.89
S&P 500 : 2.1
JP GDP : 2.02
JP CPI : 2.05
Nikkei : 2.02
CN GDP : 2.13
CN CPI : 1.99
SSE : 1.85
```

Our results show that the serial correlation is within an acceptable range, allowing us to proceed with the forecasting.

3.2. Forecasting accuracy

We provide the VAR model with lagged observations of past data to predict the future, since the terms of the model are the lags of the various time series in the dataset.

We analyzed the predicted versus actual plots and evaluated the mean absolute percent error (MAPE), root mean square error (RMSE), and correlation coefficient.




```
Forecast Accuracy of: S&P 500
mape :  4.3579
me :  -0.1601
mae :  0.1601
mpe :  2.7512
rmse :  0.1644
corr :  0.9991
minmax :  -2.7512
```

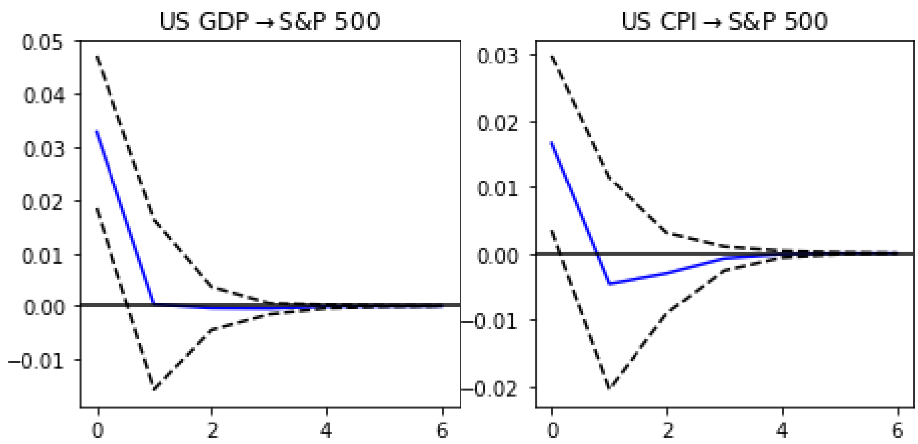
```
Forecast Accuracy of: Nikkei
mape :  17.9264
me :  -0.2056
mae :  0.2056
mpe :  15.6617
rmse :  0.2124
corr :  0.533
minmax :  -15.6617
```

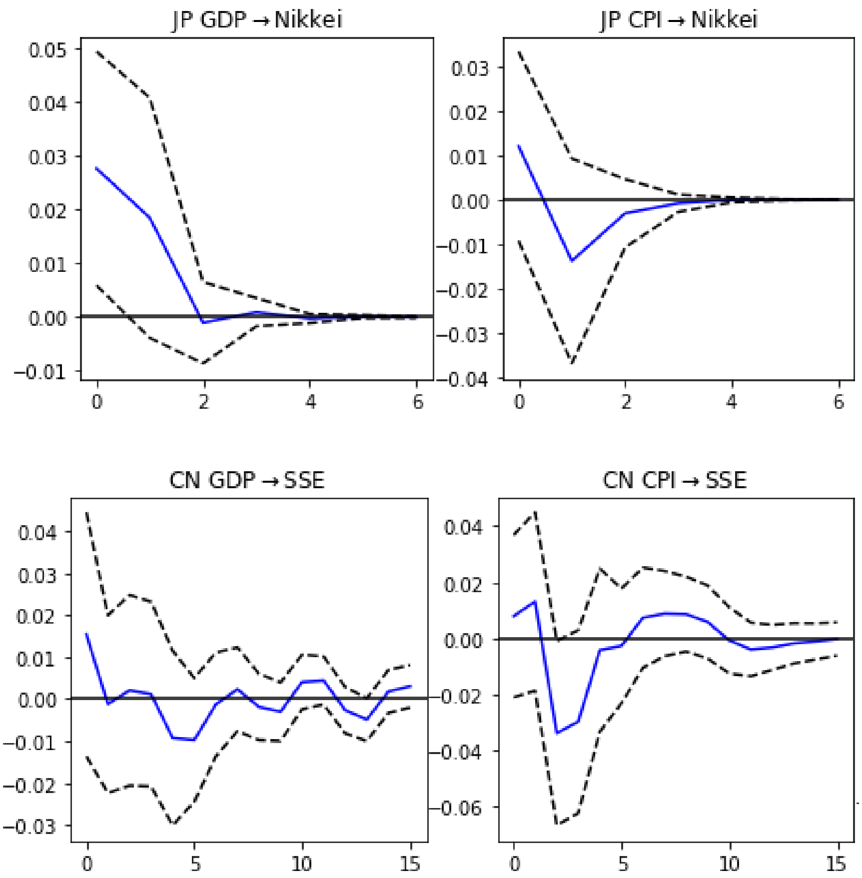
```
Forecast Accuracy of: SSE
mape :  1.1941
me :  -0.0797
mae :  0.0797
mpe :  0.3737
rmse :  0.0875
corr :  0.9638
minmax :  -0.3737
```

A low RMSE between 0.0875 and 0.2124 indicates that the model's predictions are generally accurate. Also, the high correlation coefficients (0.96-0.99) of S&P500 and SSE imply a strong positive relationship between predicted and actual values.

However, the high MAPE values from 1.19 to 17.92 indicate that the performance of the model can be improved.

3.3. Impulse response Analysis





We performed an impulse response analysis to investigate the effect of different variables on stock prices. Our findings show that CPI has a greater impact on stock prices than GDP, and that Japan and China are more responsive than the United States.

4. Discussion

4.1. Comparison of different market

We compared the performance of VAR models across multiple markets, including the S&P 500, Shanghai Composite, and Nikkei 225. Our results show that model performance varies across markets, with the S&P 500 performing best and the Nikkei 25 the worst. Nonetheless, our overall assessment of model performance is positive.

5. Conclusion

To forecast the stock market, we found that GDP and CPI have Granger causality and cointegration, indicating a common long-term trend. However, this is not always the case with the stock market. The VAR model proved effective in modeling the relationship between the S&P 500, Shanghai Composite Index, and Nikkei 225 and GDP and CPI. The Durbin-Watson test indicated that the model could satisfactorily explain variance and patterns in the time series. The model's accuracy was good, with low RMSE and high correlation coefficients, although there is still scope for improvement based on high MAPE values. An impulse response analysis indicated that CPI had a greater impact on stock prices than GDP, and Japan and China responded more strongly to variable changes than the United States. We also found that the model performed well across all markets, with some outperforming others.

Reference：

1. Schwert, G. W. (1990). Stock returns and real activity: A century of evidence. Journal of Finance, 45(4), 1237-1257. doi:10.1111/j.1540-6261.1990.tb02430.x
2. Cai, F., & Zhu, E. (2016). The impact of macroeconomic indicators on stock market indices: the case of China and its trading partners. Applied Economics, 48(55), 5395-5411. doi:10.1080/00036846.2016.1178849
3. Nishat, M., & Shaheen, R. (2004). Macroeconomic factors and Pakistani equity market. The Pakistan Development Review, 43(4), 619-637.

Data set sources：

https://fred.stlouisfed.org/ Federal Reserve Economic Data (FRED)
http://data.stats.gov.cn/ National Bureau of Statistics of China
https://www.esri.cao.go.jp/en/stat/di/historical/historical.html Cabinet Office, Government of Japan
https://finance.yahoo.com/ Yahoo Finance