Quantile Regression to Understand Interrelationship amongst Stock Exchanges Globally

Authors: Shraddha Kodavade (M.Sc. Applied Statistics and Analytics, NMIMS) Dhanashri Kanitkar (M.Sc. Business Analytics University of Texas), Hrishita Bapuram (MSc. Data and Computational Science, University College Dublin, Ireland)

Mentored by Dr. Hemant Kulkarni

(Assistant Professor, Statistics and Mathematics Department, MIT WPU, Pune)

Contact: shraddha.kodavade039@nmims.edu.in +91 9594908119
hrishita.bapuram04@ucdconnect.ie +353 89 948 1475
dpk210002@utdallas.edu +1 24 878-1038
hemant.kulkarni@mitwpu.edu.in +91 9831137926

ABSTRACT:

Investments and Stock markets are popular sources of growing money. People have become more conscious and are looking for ways to invest and reduce the risk associated with stocks. Any knowledge on how the stock indices move makes for informed decisions. This paper studies the claim that stock markets worldwide have a mathematical effect on each other and seeks to validate the same through incorporating Quantile Regression. Its scope and use have been discussed and the trends of NSE, when compared to 5 global stock indices, has been explored.

Keywords: Cross-sectional, Heteroscedasticity, Heteroscedastic, Quantile Regression, Error-Component Models.

JEL Classification: C310

INTRODUCTION:

Studies and various research have hypothesized that the stock market is not completely news driven. The preconceived notions and irrational biases also affect the stock prices. This phenomenon is termed as investor sentiment.

A strong financial market with broad participation is pivotal for an economy. India is a developing economy with high potential for growth. Capital markets acts as a barometer which is used to measure the performance of the Indian economy with its development Reform of the Indian stock market began with the founding of the Securities and Exchange Board of India (SEBI), following the global pattern, though it became more effective after the stock market scandal in 1991.

The Indian stock market mainly functions on two major stock exchanges, the BSE (Bombay Stock Exchange) and NSE (National Stock Exchange). In terms of market capitalization, BSE and NSE have a place in the top five stock exchanges of developing economies of the world. The Bombay Stock Exchange is one of the oldest exchanges across the world, while the National Stock Exchange is among the best in terms of sophistication and advancement of technology.

After the process of economic liberalization, Privatization and Globalization, the Indian capital market has been assigned a very dominating place in the financing and loaning industry. The strategy of industrialization, which protected domestic industries from foreign competition, was also responsible for high cost and low growth in the economy. [3]This rate has now been stunted due to low corporate and government plugins. The socio-economic disparity can be undone by trickling the monetary benefits of the economy to everyone. This can be achieved by ensuring a constant and steady stream of flow in the capital markets. Since Secondary and Tertiary sectors are what almost most of the economy is composed of, a fund route via these markets ensures a harmony amongst the producers of goods in the economy with the consumers, who in short supply the deficit funds.

Another phenomenon worth noticing is the dilemma of investors blurring in context to choosing traditional low return routes. Investors have begun a journey in educating themselves about the way a stock market operates to hedge risks and dodge low return bank and other related institutional incomes.

So, to apprehend the actions in inventory markets in India it would be good if the trends in this marketplace is compared to the trends in the different global markets and study the connection among numerous markets.

[3] Comparative Analysis of Indian Stock Market with International Markets Retrieved from

STATISTICAL CONCEPTS:

Regression Analysis:

Regression analysis is used to establish the relationship between two kinds of variables – the dependent or response variables and the independent variables or regressors. Using the regressors, a relationship is developed which enables the estimation of the dependent variable which is why the said variate also known as response variable. It is a very powerful statistical tool with increasing complexity across far ranging applications in different disciplines.

A Brief Introduction to Linear Regression:

- Model

$$Response\ Variable = Model\ Function + Random\ Error$$

- Linear Regression Model $Y = \beta_0 + \beta_1 X + \varepsilon$
- $\beta = (\beta_0, \beta_1)$: Model parameters
  $E(\varepsilon) = 0, V(\varepsilon) = \sigma^2$

Least Squares Estimation image

- Objective function:

$$\hat{\beta} = \text{argmin} \sum^{N} (y_i - \beta_0 - \beta_1 x)^2$$

- Predicted value of the response

$$\hat{Y} = \hat{E}(Y|X = x) = \hat{\beta}_0 + \hat{\beta}_1 x$$

Drawbacks of Linear Regression and the case for Quantile Regression:

If either of the assumptions of linear regression (Linearity, Homoscedasticity, Normality and Independence) are not met, the linear model is not suitable.

One alternative to combat the issue would be to estimate the median value of the response variable using the Least Absolute Deviation, which is similar to the least square procedure. Median regression may seem like a fitting solution but even that is not useful if extremities in variables are to be studied. And that, is where the significance of Quantile Regression is realized.

Quantile: The τ^{th} quantile denotes the value of random variable bellow which proportion of population is tau: $Q_\tau(Y) = y_\tau = F_F^{-1}(\tau)$,

where $F_F(y_\tau) = P(Y \leq y_\tau) = \tau$

Quantile regression enables the estimation of the non-central summary statistics of the response variables, which is so much powerful and useful than one may realize. It is a method which models quantiles instead of mean for a given set of regressors upon which the model is built.

Quantile Regression can serve as an efficient choice over simple linear regression in the following instances:

- In the presence of many outliers (Refer fig. 1.a.)
- When the data being modelled is non-normal in nature (Refer fig. 1.b.)
- When the data does not follow the pattern of homoscedasticity (constant variance in response for all regressors)

Model and Definition of Quantile Regression:

- Model (Koenker and Bassett, 1978) for (x_i, y_i) is $Q_\tau(Y|X = x) = \beta_{0\tau} + \beta_{1\tau}X$
  $Q_\tau(\varepsilon) = 0$

Parameters are estimated using simplex algorithm.

$$\hat{\beta_\tau} = \text{argmin} \sum_{i}^{N} \rho_\tau(y_i - \beta_{0\tau} - \beta_{1\tau}x_i)$$

The fitted conditional quantile is: $\hat{Q}_\tau(Y|X) = \hat{\beta}_{0\tau} + \hat{\beta}_{1\tau}X$

LITERATURE REVIEW:

In recent times, Stock Market and monetary trends have seen considerable attention.

Debijan Mukherjee observed in his research paper that the Indian Stock Exchanges, BSE and NSE, shows similarities in trends with its international counterparts. The stock markets chosen were analyzed both quantitatively and qualitatively. Especially after 2003, there has been a high correlation among the stock exchanges and are found to be more integrated with each other and that socio, economic and political trends affect the stock markets.

- **Manjula K A and Karthikeyan** found that NSE and Asian Stock Exchanges draw an integrated relationship among each other and that data from other stock exchanges can be used to predict the future trends of NSE. They concluded the correlation and regression of Nifty with other Asian Stock indices. They also predicted that linear regression model failed to show the correct trends and that gradient boosting regression model showed better results and were useful to predict the future prices.

- **Fredrik Bergling** used quantile regression to affectively study portfolio management. He divided the value into various quantiles tried to get more exact value of the predicted return. He concluded that quantile regression is more

advantageous and gave better prediction. They concluded that two out of three times the quantile regression does outperform the Mean-Variance and OLS while calculating risk management

METHODOLOGY - EXPLORATORY DATA ANALYSIS:

The line chart for each index displays the yearly change in the index. The 2020 pandemic effect can be observed in the line chart of all indices. (Refer figure2.a to 2.f)

The boxplot of each index shows that there are too many outliers in the indices data and therefore it is not favorable to apply the linear regression model on the data. (Refer figure3.a to 3.f)

Analyzing the relationship of different stock indices:

- The image represents the correlation between the dependent and independent variables. (Refer figure4)
- We can observe that Dow Jones has a negative correlation of (-0.21) with Nifty50.
- Whereas Nikkei225 and LSE have a positive correlation of 0.33 and 0.23 with Nifty50, respectively.
- NASDAQ have a positive relation of 0.26 with Nifty50.
- SSE has a negligible correlation with Nifty50
- The image is a scatterplot of Nifty50 (Refer figure5) which is the dependent variable with the independent variables like Dow Jones, Nasdaq, LSE, and Nikkei
- We can see that the data is not following a linear relationship.

CONCLUSION:

1] Quantile Regression Fit - Results and Analysis

- The Quantile Regression model has been fitted on the data set for tau values. (Ranging from 0.05 to 0.95).

- Here, for each quantile we get a table of **estimated coefficients** with 95% Confidence Interval. These are the **parameter values**, which are based on the following assumptions:

1. It is the same as ordinary regression.

Tau Values	Intercept	Nasdaq	SSE	LSE	Nikkei225	DJI
Linear	0.00000	0.22000	-0.02000	0.11000	0.20000	-0.16000
0.05	-1.40900	0.28800	-0.09700	0.08100	0.15200	-0.22600
0.10	-0.99700	0.19600	-0.05300	0.08900	0.16100	-0.18400
0.15	-0.75881	0.17934	-0.02732	0.09647	0.17835	-0.15820
0.20	-0.60098	0.19537	-0.04221	0.10094	0.17770	-0.15879
0.25	-0.47356	0.17552	-0.03178	0.10074	0.17914	-0.13562
0.30	-0.36302	0.17338	-0.01847	0.10265	0.18751	-0.15425
0.35	-0.26982	0.16528	-0.00870	0.09622	0.18720	-0.14722
0.40	-0.16778	0.15949	-0.00708	0.09652	0.17444	-0.14087
0.45	-0.09387	0.15366	-0.00111	0.10127	0.18099	-0.13433
0.50	-0.02345	0.14986	-0.01052	0.10568	0.17410	-0.13841
0.55	0.05400	0.14653	-0.00275	0.10698	0.18291	-0.13126
0.60	0.14448	0.15026	-0.00546	0.10083	0.19751	-0.12513
0.65	0.24304	0.15246	-0.00711	0.11425	0.18196	-0.12654
0.70	0.3471	0.17706	-0.01270	0.11178	0.18869	-0.13486
0.75	0.47387	0.16843	-0.00189	0.09972	0.21566	-0.12922
0.80	0.59671	0.16246	-0.00482	0.12766	0.21634	-0.11804
0.85	0.74845	0.18410	-0.01456	0.11876	0.21896	-0.14120
0.90	0.99245	0.17148	0.00932	0.12029	0.23341	-0.14801
0.95	1.48503	0.29525	-0.01502	0.13947	0.23068	-0.14860

Table 1: Estimated Coefficients for each quantile with 95% Confidence Interval.

2. Marginal effect is also interpreted similarly.

   The equation: $Y = \beta_0 + \beta_1 X + \beta_2 X + \beta_3 X + \beta_4 X + \varepsilon$

- here, has Nasdaq, SSE, LSE, Nikei225 & DJI as its variables. The intercepts gives us β_0 .
- For the row depicting linear relationship, a difference of 0.22 occurs with the presence and absence of Nasdaq. In a similar fashion, for SSE, LSE, Nikei225 and DJI we notice the linear coefficients of parameters taking values of -0.02, 0.11, 0.200 and -
  0.12 respectively.
- For understanding each quantile (tau value), the following relationship should be considered: -

1. When Linear Parameter > Quantile Parameter – Underestimation takes place
2. When Linear Parameter < Quantile Parameter – Overestimation takes place

Tau Values	Nasdaq	SSE	LSE	Nikkei225	Dow Jones
0.05	0.00000	0.02055	0.10750	0.00000	0.00000
0.10	0.00000	0.01147	0.03652	0.00000	0.00000
0.15	0.00000	0.24971	0.00000	0.00000	0.00000
0.20	0.00000	0.03700	0.00001	0.00000	0.00000
0.25	0.00000	0.06487	0.00000	0.00000	0.00000
0.30	0.00000	0.24757	0.00000	0.00000	0.00000
0.35	0.00000	0.57808	0.00000	0.00000	0.00000
0.40	0.00000	0.62753	0.00000	0.00000	0.00000
0.45	0.00000	0.93643	0.00000	0.00000	0.00000
0.50	0.00000	0.43655	0.00000	0.00000	0.00000
0.55	0.00000	0.85391	0.00000	0.00000	0.00000
0.60	0.00000	0.74130	0.00000	0.00000	0.00000
0.65	0.00000	0.67611	0.00000	0.00000	0.00000
0.70	0.00000	0.48437	0.00000	0.00000	0.00000
0.75	0.00000	0.92247	0.00000	0.00000	0.00000
0.80	0.00000	0.80769	0.00000	0.00000	0.00000
0.85	0.00000	0.54197	0.00000	0.00000	0.00000
0.90	0.00000	0.76188	0.00000	0.00000	0.00003
0.95	0.00000	0.75476	0.05784	0.00000	0.00233

Table 2. p-values

- H_0: There is no linear significance of τ^{th} quantile of Nifty 50
- H_1: There is a linear significance of τ^{th} quantile of Nifty 50
- SSE: The p-value > 0.05, hence we accept the H_0.
- To understand this phenomenon, we need to look into the basket of major portfolios that high correlation impact on the behavior of these indices.
- Nifty50 is dependent on IT companies listed in the US exchanges.
- Dow Jones Industrial Average is a major container for secondary sector than the tertiary sector.
- The analysis of QR plots hence after holds these relations in consideration.

The QR plots:

- This Graph is a function of Quantiles. (Refer figure6)
- The Grey region highlights the confidence interval of 95%.
- The Horizontal Line is the Ordinary Least Squares estimate with 95% limits as horizontal lines.

Mid-50% Analysis:
This is used to build analysis on the bandwidth of Q1 to Q3. (Refer figure7).

Key Insights:

- Nasdaq tends to have a larger impact on net change in Nifty 50 at the extreme 5% of the changes.
- Overall, Shanghai seems to be irrelevant in its impact on Nifty 50.
- Nasdaq and DJI Paradox.
- LSE seems consistent and fits well within the linear model's fit in its impact on Nifty 50
- The middle 50% of the net changes seem consistent around the linear model fit, apart from Nasdaq

In general, stock exchange fluctuations affect top and bottom 10% changes of Nifty50 for a major part. Hence, a linear model would have not solved the problem, and hence justifies the suitability of a Quantile Regression model.

The box plot has been built up on predicted values for each tau values. The blue dots being the actual values. They lie in the range of the box-plots bandwidth.
Therefore, we can establish accuracy in the interpretation of the data. (Refer figure 8)

REFERENCES:

- Financial Integration for Indian Stock Market Sadhan Kumar Chattopadhyay & Samir Ranjan Behera
- Comparative analysis of Indian stock market. (n.d.). Retrieved November 12, 2021, from Anon, Comparative analysis of Indian stock market with ... Available at: https://www.greatlakes.edu.in/pdf/ DebijanMukherjee.pdf [Accessed January 21, 2022].
- Koenker, R., & Hallock, K. F. (n.d.). Quantile regression. Journal of Economic Perspectives. Retrieved November 12, 2021, from Koenker, R. & Hallock, K.F., Quantile regression. *Journal of Economic Perspectives*. Available at: https://www.aeaweb.org/articles?id =10.1257%2Fjep.15.4.143&Article Search%5Bwithin%5D%5Barticlet itle%5D=1&ArticleSearch%5Bwit hin%5D%5Barticleabstract%5D=1 &ArticleSearch%5Bwithin%5D%5 Bauthorlast%5D=1&JelClass%5Bv alue%5D=C&page=32 [Accessed January 21, 2022].
- Anon, *Simple linear regression*. Available at: https://sphweb.bumc.bu.edu/otlt/M PH-odules/BS/R/R5_Correlation-Regression/R5_Correlation- Regression4.html#:~:text=There% 20are%20four%20assumptions%2 0associated,are%20independent%2 0of%20each%20other.[Accessed January 21, 2022].
- Codes https://github.com/shraddhacodes/QuantileRegressionforStockMarketAnalysis

FIGURES

figure 1.a. Outliers

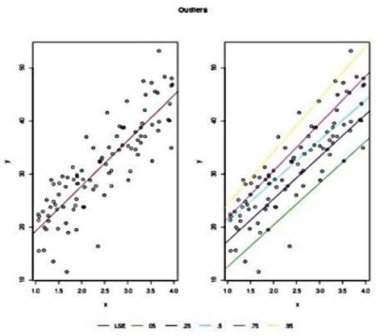

figure 1.b. Non-Normal

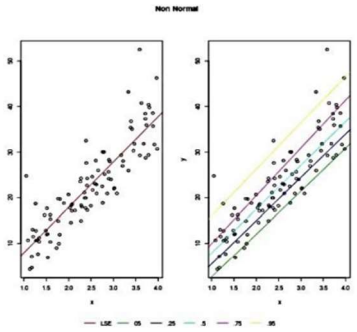

figure 1.c. Heteroscedasticity

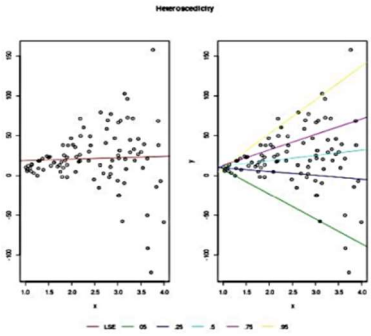

figure 2.a. Line Chart of yearly change of Nifty50

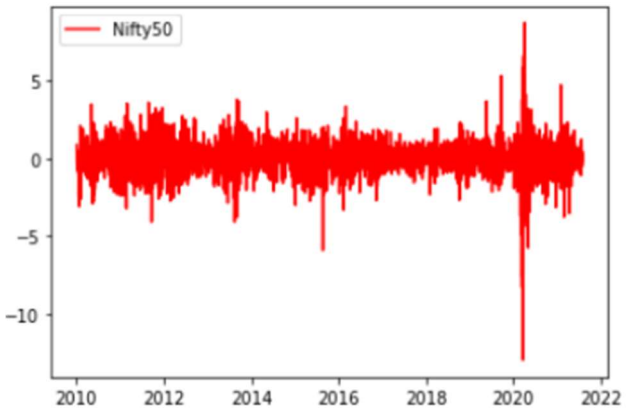

figure 2.b. Line Chart of yearly change of Nasdaq

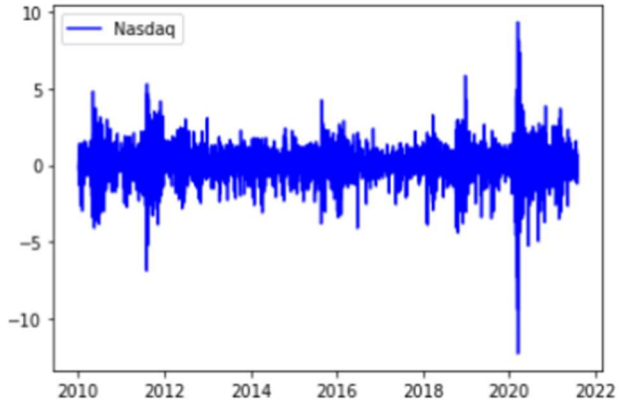

figure 2.c. Line chart of yearly changes of Nikkei 225

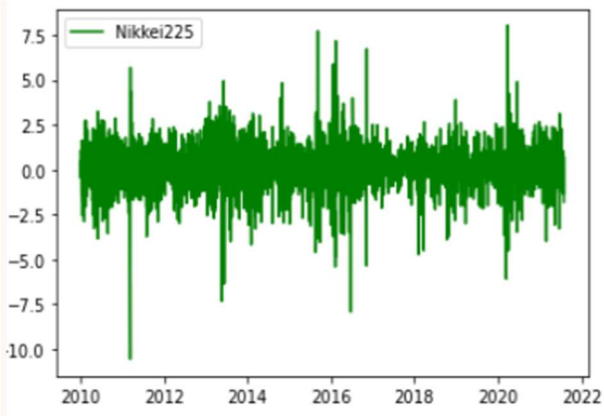

figure 2.d. Line chart of yearly changes of Dow Jones

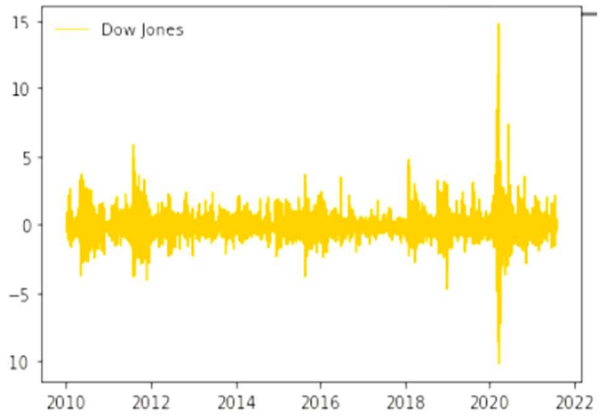

figure 2.e. Line chart of yearly changes of London Stock Exchange

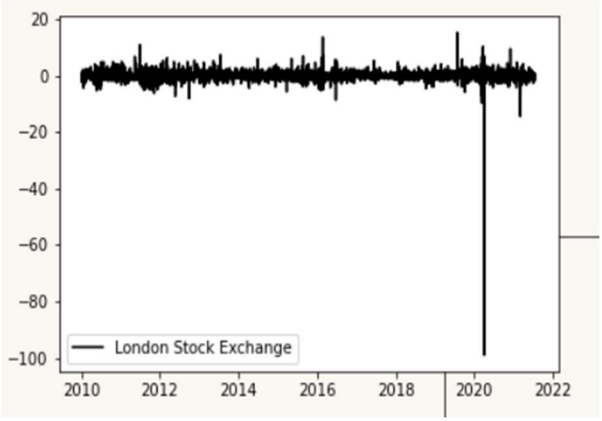

figure 2.f. Line chart of yearly changes of Shanghai Stock Exchange

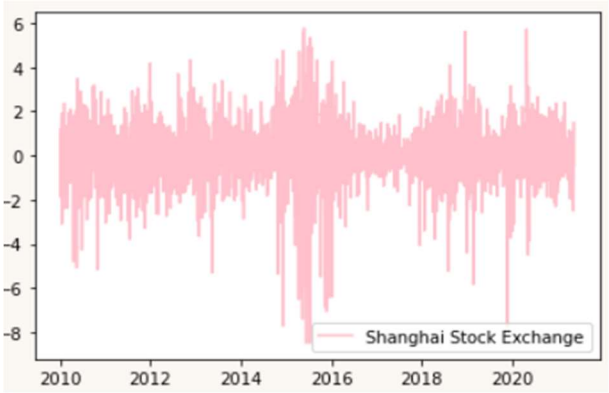

figure 3.a. Box plot of Nifty50

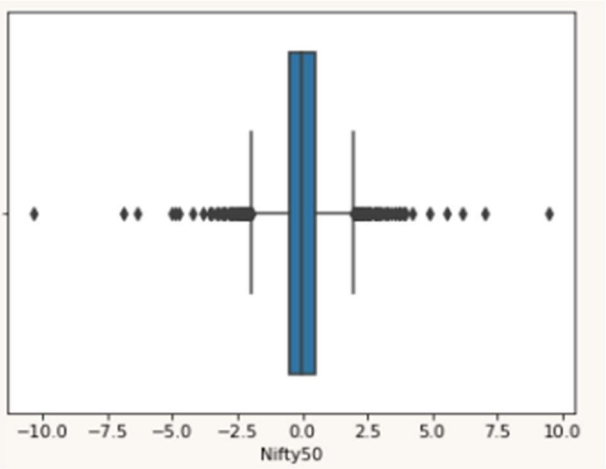

figure 3.b. Box plot of Nasdaq

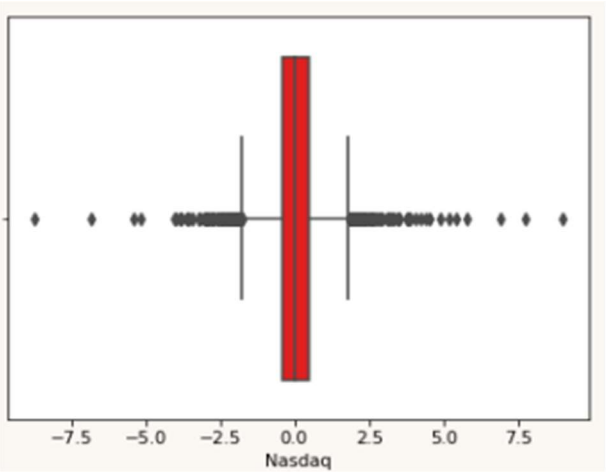

figure 3.c. Box plot of Nikkei225

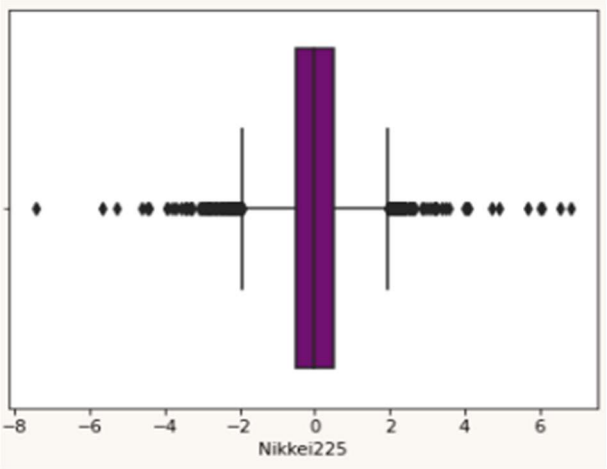

figure 3.d. Box plot of Dow Jones

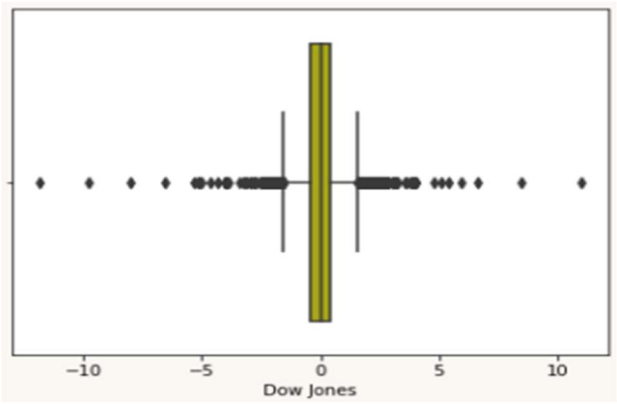

figure 3.e. Box plot of London Stock Exchange

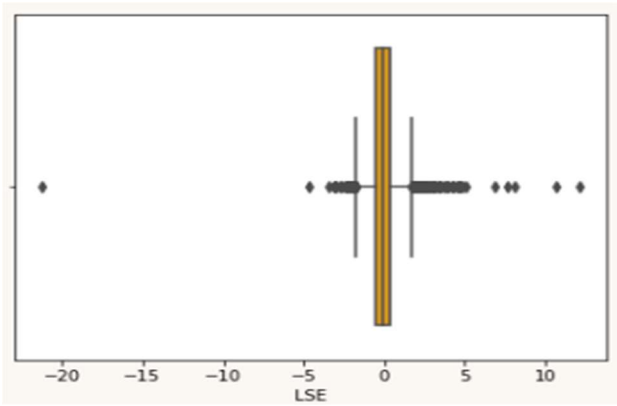

figure 3.f. Box plot of Shanghai Stock Exchange

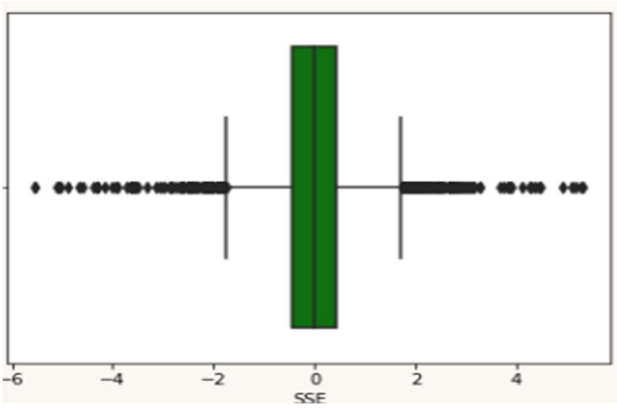

figure 4. Correlation plot of different indices

figure 5. Scatter plot of Nifty50 with other indices

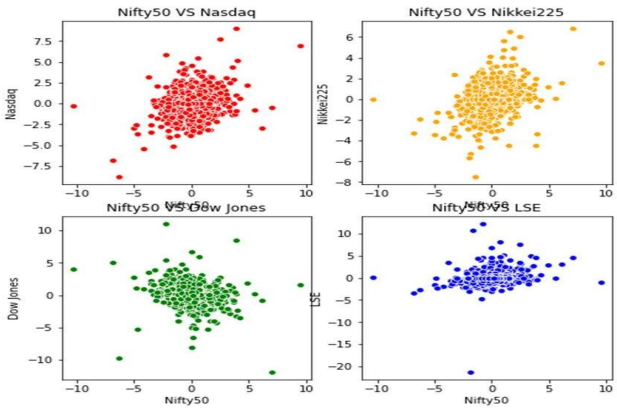

figure 6. The Quantile Regression Plots

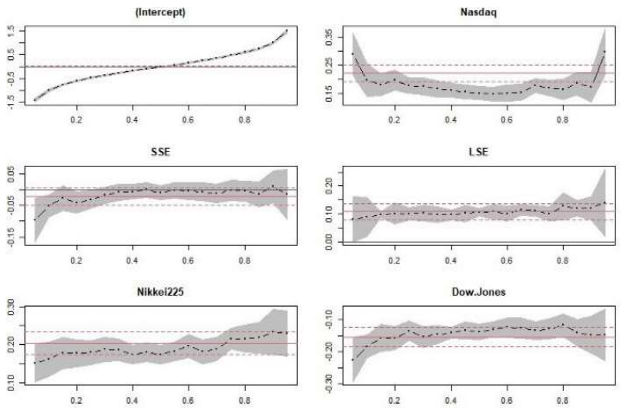

figure 7. The Quantile Regression Mid-50% Analysis Plots

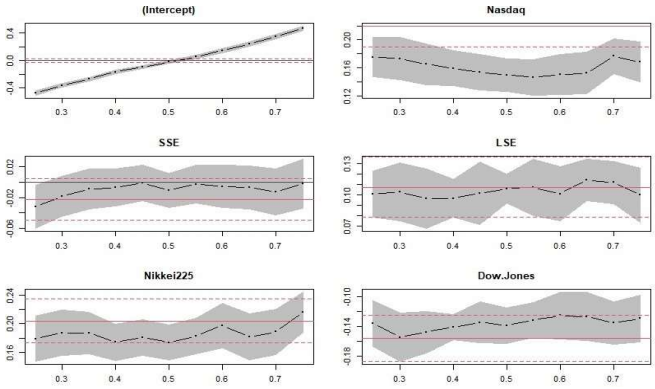

figure 8.Boxplots