© November 2025 | IJIRT | Volume 12 Issue 6 | ISSN: 2349-6002

Time Series Cointegration Analysis of Major Indian Stocks

Ms. Mildred Lemos
Assistant Professor, Department of Computer Science, Rosary College, Goa

***Abstract*—Time Series Cointegration provides a powerful statistical framework for understanding the long-term interconnectedness of stock prices. This understanding is crucial for developing sophisticated trading strategies, managing risk effectively, and gaining deeper insights into the underlying dynamics of the stock market. This paper investigates the long-run equilibrium relationships among the major Indian stocks from diverse sectors listed on the Bombay Stock Exchange over a five-year period - April 2020 to March 2025. Employing the Engle-Granger two-step cointegration test on daily closing prices, we identify statistically significant cointegrated pairs, with HDFCBANK emerging as a central stock strongly linked with multiple banking and financial sector stocks. These findings indicate that deviations from equilibrium between these stocks tend to revert over time, reflecting stable long-term relationships. The results have important implications for portfolio management, pairs trading, risk mitigation, and strategic investment decisions within the Indian equity market.**

***Index Terms*—Cointegration, Time series, Indian stocks**

I. INTRODUCTION

The stock market serves as a complex tapestry woven from intricate interactions among various economic variables. Understanding the dynamic relationships between asset prices is crucial for investors, portfolio managers, and financial analysts. One important technique employed in this domain is cointegration analysis, a statistical approach that identifies whether a group of non-stationary time series variables, such as stock prices, share a long-term equilibrium relationship.

Cointegration, a statistical property rooted in econometrics, indicates a long-run equilibrium between two or more time series variables, even if those variables themselves are individually non-stationary (i.e., they have trends). This means that despite their individual fluctuations, the variables move together in the long run, anchored by an underlying equilibrium relationship.:

Stock prices typically exhibit non-stationary behaviour, which means their statistical properties change over time. Therefore, using classical regression techniques on non-stationary data may yield spurious results, misleading researchers, and investors alike. The concept of cointegration, first introduced by Engle and Granger (1987), has significantly shaped modern time series econometrics, especially in financial modelling. Cointegration suggests that while individual time series may be non-stationary, certain linear combinations of them may be stationary, indicating a long-term equilibrium relationship. In financial markets, this concept has been widely used to detect stable relationships between asset prices, providing a foundation for strategies like pairs trading (Gatev, Goetzmann, and Rouwenhorst, 2006).

The Indian equity market, characterized by its growing sophistication and diverse sectoral representation, presents an interesting landscape for exploring cointegration relationships. Identifying cointegrated stock pairs can offer valuable insights for

- Pairs Trading: Constructing market-neutral trading strategies that capitalize on temporary divergences from the long-run equilibrium.
- Portfolio Diversification: Understanding which stocks move together in the long run can help in building more effective diversification strategies by avoiding highly correlated assets.
- Risk Management: Recognizing cointegrated relationships can aid in assessing and managing systemic risk within a portfolio.
- Market Efficiency: The presence or absence of strong cointegration can provide evidence

regarding the efficiency of price discovery in the market.

At its core, cointegration analysis relies on several key statistical tests, including the Augmented Dickey-Fuller (ADF) test for stationarity and the Johansen test for determining cointegration ranks among multiple time series. The ADF test assesses the presence of a unit root in a univariate time series, thus implying a non-stationary process. If one or more variables in a set are found to be cointegrated, it signifies that despite their individual non-stationary nature, there exists a stationary linear combination among them. This linear combination represents the long-term relationship, lending credence to the notion that stock prices may not drift apart indefinitely but tend to revert to a mean or a common path.

Alexander (1999) applied cointegration techniques to hedging strategies, showing that cointegrated pairs provide more efficient risk management solutions than correlation-based approaches. In the Indian context, research by Sharma and Bodla (2011) explored cointegration between Indian and US markets, establishing that external linkages could affect domestic stock movements. Patel and Patel (2018) specifically examined cointegration within Indian banking stocks, finding significant long-run associations driven by macroeconomic factors like interest rates and credit growth.

Despite these contributions, a study of major stocks within India remains limited. Most prior studies focus either on international linkages or on isolated sectors like banking and IT. This study aims to bridge that gap by systematically examining cointegration across multiple sectors banking, energy, FMCG, IT, and infrastructure thus offering a holistic view of the Indian stock market interdependencies.

II. METHODOLOGY

This study employs the Engle–Granger two-step cointegration test to investigate the existence of long-term equilibrium relationships between pairs of ten major Indian stocks. The analysis was carried out using Python programming, leveraging statistical and econometric libraries such as pandas, numpy, and statsmodels.

2.1 Data Collection:

The study uses data on the daily closing prices of the ten major stocks from diverse sectors of the Indian stock market. The stocks selected represent key sectors, including Banking, Energy, IT, FMCG, Telecom and Infrastructure. These stocks were chosen for their prominence and market capitalization in India.

The data spans from April 2020 to March 2025, covering a period of five years to account for different market conditions.

Daily closing price data for the ten selected stocks listed on the Bombay Stock Exchange (BSE) were obtained using the Alpha Vantage API. The stock tickers and their respective sectors are:

- Reliance Industries Ltd. (RELIANCE.BSE) – Energy, Retail, Telecom
- HDFC Bank Ltd. (HDFCBANK.BSE) – Private Banking
- Tata Consultancy Services Ltd. (TCS.BSE) - IT
- Bharti Airtel (BHARTIARTL.BSE) – Telecom
- ICICI Bank Ltd. (ICICI.BSE) – Private Banking
- SBI (SBIN.BSE) – Public sector Banking
- Infosys Ltd. (INFY.BSE) – IT
- Bajaj Finance Ltd. (BAJFINANCE.BSE) - Financial Services
- Hindustan Unilever Ltd. (HINDUNILVR.BSE) - FMCG
- Larsen & Toubro Ltd. (LT.BSE) - Infrastructure

2.2 Data Preprocessing:

The raw daily price data were cleaned and organized into a Pandas DataFrame. The closing prices were extracted, and any missing data points were removed through pairwise deletion to ensure robust cointegration testing for each pair.

2.3 Cointegration Test:

For each unique pair of stocks, the Engle-Granger two-step test was performed:

Step 1 – Estimation of the Long-Run Relationship

For each stock pair, an Ordinary Least Squares (OLS) regression was conducted, where the logarithm of the closing price of one stock was regressed on that of the other:

$\ln(P_{A,t}) = \alpha + \beta \ln(P_{B,t}) + \varepsilon_t$

Here, ε_t represents the residuals capturing short-term deviations from the long-run equilibrium. This step was implemented in Python using the OLS () function from the statsmodels library, and the resulting residuals were extracted for further stationarity testing.

Step 2 – Testing for Stationarity of Residuals

In the second step, the residuals obtained from the OLS regression were tested for stationarity using the Augmented Dickey–Fuller (ADF) test in Python's statsmodels library.

If the residual series was found to be stationary i.e., integrated of order I (0), the null hypothesis of a unit root was rejected, indicating that the two stock prices are cointegrated and share a long-term equilibrium relationship.

2.4 Interpretation of Test Results

The pairs with statistically significant ADF test results (p-value < 0.05) were identified as cointegrated pairs, implying the existence of a long-term equilibrium relationship between their price movements.

III. RESULT ANALYSIS

STOCK 1	STOCK 2	ADF P-value (Residuals)
HDFCBANK	SBIN	0.0014
HDFCBANK	ICICIBANK	0.0022
HDFCBANK	BAJFINANCE	0.0032
TCS	SBIN	0.0033
HDFCBANK	HINDUNILVR	0.0063
HDFCBANK	LT	0.0068
TCS	BAJFINANCE	0.0098
HDFCBANK	BHARTIARTL	0.0101

Table 3.1: Cointegration Test Results of Stock Pairs Using Engle–Granger Method (ADF Test on Residuals)

These results suggest that the identified stock pairs maintain a stable long-term relationship, meaning that any short-term deviations between the prices tend to revert to a long-term equilibrium.

The observations that can be made based on the results are: -Prominent role of HDFC Bank HDFCBANK appears in five of the eight cointegrated pairs, indicating its strong co-movement with other banking and financial stocks such as SBIN, ICICIBANK, and BAJFINANCE. This suggests that HDFCBANK's stock price dynamics are closely linked with the overall performance of the Indian banking sector, and deviations from equilibrium with these stocks are likely to revert over time.

Banking Sector Co-Movements:

HDFCBANK–SBIN and HDFCBANK–ICICIBANK are the most strongly cointegrated pairs, reflecting a very stable long-term relationship.

The TCS–SBIN and TCS–BAJFINANCE pairs indicate that IT sector stock TCS also exhibits a long-term equilibrium relationship with select banking and financial stocks, suggesting cross-sector interactions driven by broader market factors.

Diversified Sector Links:

HDFCBANK also shows significant long-term relationships with HINDUNILVR (FMCG), LT (engineering), and BHARTIARTL (telecom).

This implies that HDFCBANK's stock movements are correlated with broader market trends, making it a potentially central stock for constructing market-neutral or pairs trading strategies.

IV. CONCLUSION

The study confirms the presence of long-term equilibrium relationships among major Indian stocks, particularly within the banking sector. The strongest cointegration was observed between HDFCBANK–SBIN, HDFCBANK–ICICIBANK, and HDFCBANK–BAJFINANCE, indicating that these stocks move together over time despite short-term market fluctuations.

These findings have practical implications for portfolio management and trading strategies, such as pairs trading, where exploiting cointegrated stock pairs can help in hedging risk or capturing long-term mean-reversion opportunities.

Further research could explore the underlying economic and financial factors driving the observed cointegration relationships. Investigating the stability of these relationships over different time periods and incorporating other multivariate cointegration techniques could also provide a more comprehensive understanding of the long-term dynamics within the Indian equity market. Additionally, the practical implications of these findings for constructing profitable pairs trading strategies and optimizing portfolio diversification could be explored in future studies.

REFERENCES

[1] Engle, R. F., & Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. Econometrica.

[2] Alexander, C. (1999). Optimal Hedging Using Cointegration. Philosophical Transactions: Mathematical, Physical and Engineering Sciences.

[3] Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of a Relative-Value Arbitrage Rule. Review of Financial Studies.