20. Volatility Spillover between Gold Prices and the Indian Stock Market: Evidence from Time-Series Econometrics and Machine Learning

[1]Dr.S.Krishnaprabha, [2]Jenifer Chrisla .T

[1]Associate Professor
Department of Management Studies, Sri Ramakrishna Engineering College
[2]Research Scholar (Ph.D)
Department of Management Studies, Sri Ramakrishna Engineering College

Abstract

Gold is widely regarded as a safe-haven asset that helps hedge against stock market uncertainty, particularly during periods of economic instability. This study examines the volatility spillover between gold prices and the Indian stock market using daily data from January 2015 to December 2024. Time-series econometric techniques, including ADF, VAR, Granger causality, and GARCH (1,1), are employed to analyze return dynamics and volatility transmission. Machine learning models such as Random Forest, XGBoost, and LSTM are used to forecast market volatility and compare predictive performance. The findings reveal significant volatility spillovers between gold and equity markets, with LSTM providing the highest forecasting accuracy. The study offers useful insights for investors, portfolio managers, and policymakers in improving diversification and risk management strategies.

Keywords: *Gold Prices, Indian Stock Market, Volatility Spillover, GARCH, Machine Learning, LSTM, Random Forest, Financial Markets*

1. Introduction

Financial markets have become increasingly interconnected due to globalization and rapid capital flows, resulting in volatility spillovers across asset classes. Understanding these spillovers is essential for investors, policymakers, and financial institutions as they influence investment decisions, portfolio diversification, and risk management. Gold is widely regarded as a store of value, an inflation hedge, and a safe-haven asset during periods of economic uncertainty. Investors often shift from equities to gold during financial crises, making the relationship between gold prices and stock markets a significant area of financial research. In India, gold holds considerable importance because of cultural preferences and investment behavior, while the NIFTY 50 and BSE Sensex represent rapidly expanding equity markets. This creates a need to examine whether volatility in one market affects the other. Traditional econometric techniques such as VAR, Granger causality, and GARCH have been widely used to analyze volatility transmission. More recently, machine learning models including Random Forest, XGBoost, and LSTM have demonstrated superior ability to capture nonlinear relationships and improve forecasting accuracy. This study integrates time-series econometric methods with machine learning techniques to investigate volatility spillovers between gold prices and the Indian stock market, providing valuable insights into market interdependence and financial risk management.

2. Integrated Literature Review

The relationship between gold prices and stock market performance has been widely studied because of its importance for portfolio diversification and risk management. Earlier research

 ISBN: 978-81-687720-2-1

emphasized gold's role as a hedge and safe-haven asset, whereas recent studies have focused on volatility transmission using advanced econometric and computational techniques. Baur and Lucey (2010) found that gold serves as both a hedge and a safe haven during market downturns. Similarly, Baur and McDermott (2010) reported that gold protects investors during financial crises, although its effectiveness differs across countries. Using GARCH models, Hammoudeh et al. (2013) identified significant volatility spillovers between commodity and equity markets, while Ghosh and Kanjilal (2016) observed bidirectional relationships between gold prices and Indian stock market returns. Recent studies employing DCC-GARCH and VAR models have further confirmed stronger volatility transmission during periods of economic uncertainty.
Machine learning has also emerged as an effective tool for financial forecasting. Fischer and Krauss (2018) found that LSTM models outperform conventional statistical methods, while Gu, Kelly, and Xiu (2020) demonstrated that machine learning algorithms effectively capture nonlinear financial relationships. Random Forest and XGBoost have also shown strong predictive performance in volatility forecasting. Despite these advances, few studies integrate econometric and machine learning approaches to examine volatility spillovers in the Indian market. This study bridges that gap by combining both techniques to provide a comprehensive analysis of the relationship between gold prices and the Indian stock market.

4. Objectives of the Study

1. To examine the relationship between gold price movements and the Indian stock market using time-series econometric techniques.
2. To investigate the existence and direction of volatility spillovers between gold prices and stock market returns.
3. To compare the forecasting performance of machine learning models with conventional econometric methods in predicting market volatility.

5. Research Hypotheses

H_{01}: Gold price returns do not significantly influence the volatility of the Indian stock market.
H_{11}: Gold price returns significantly influence the volatility of the Indian stock market.
H_{02}: Machine learning models do not significantly improve volatility forecasting compared with traditional econometric models.
H_{12}: Machine learning models significantly improve volatility forecasting compared with traditional econometric models.

6. Data and Methodology

This study adopts a quantitative research design using secondary time-series data to investigate the volatility spillover between gold prices and the Indian stock market. A combination of traditional econometric techniques and machine learning algorithms is employed to capture both linear and nonlinear relationships in financial market volatility. The integration of these approaches provides a comprehensive understanding of return dynamics and enhances forecasting accuracy.

6.1 Data Sources

Daily closing price data for gold and the Indian stock market are collected from reliable financial databases for the period **January 2015 to December 2024**. This period encompasses several major economic events, including fluctuations in global commodity markets, the COVID-19 pandemic,

ISBN: 978-81-687720-2-1

geopolitical tensions, and periods of monetary policy changes, thereby providing a robust framework for analyzing volatility transmission.

The primary data sources include:

- Multi Commodity Exchange (MCX) – Gold Spot Prices
- National Stock Exchange (NSE) – NIFTY 50 Index
- Reserve Bank of India (RBI) – USD/INR Exchange Rate
- Investing.com and Yahoo Finance – Historical Financial Data

Daily logarithmic returns are computed using the following equation:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

where:

- R_t = Daily return
- P_t = Closing price on day *t*
- P_{t-1} = Closing price on the previous trading day

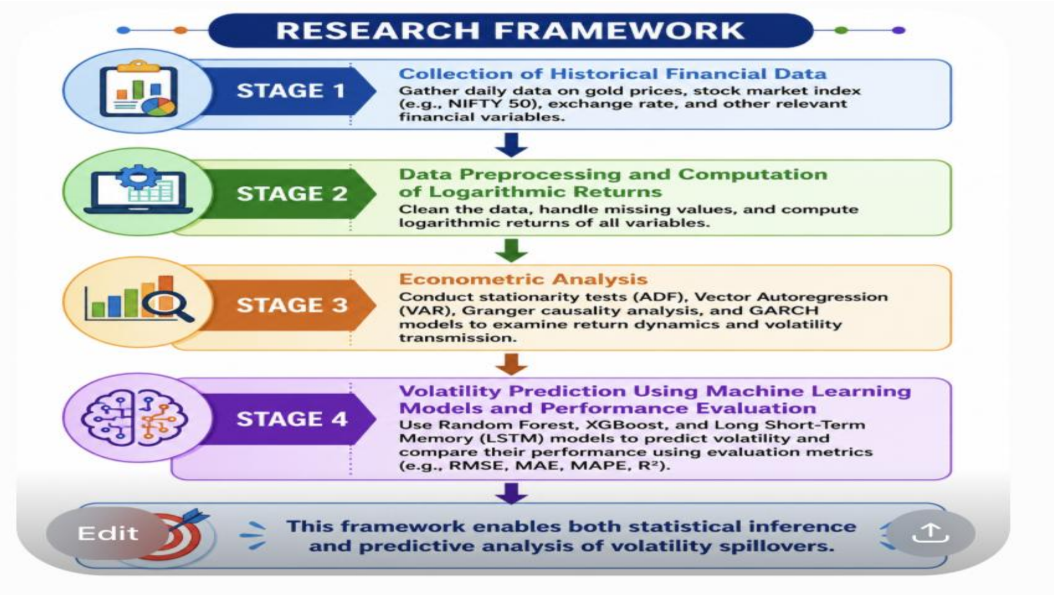
The use of logarithmic returns reduces heteroscedasticity and facilitates comparison across financial assets.

Table 1. Variables Used in the Study

Variable	**Description**	**Measurement**
Gold Return (GR)	Daily logarithmic return of Gold Price	Percentage
NIFTY Return (NR)	Daily logarithmic return of NIFTY 50	Percentage
Gold Volatility	Conditional variance from GARCH	Decimal
Stock Volatility	Conditional variance from GARCH	Decimal
Exchange Rate	USD/INR Daily Rate	Rupees
Trading Volume	Daily NIFTY Trading Volume	Number of Shares

6.2 Research Framework

ISBN: 978-81-687720-2-1



6.3 Descriptive Statistical Analysis

Descriptive statistics are computed to summarize the characteristics of the financial return series. The analysis includes:

- Mean
- Standard Deviation
- Minimum
- Maximum
- Skewness
- Kurtosis
- Jarque–Bera Normality Test

Financial return series typically exhibit high kurtosis and non-normal distributions, indicating the presence of extreme observations and volatility clustering. These characteristics justify the application of GARCH-family models.

Table 2. Sample Descriptive Statistics

Statistic	**Gold Returns**	**NIFTY Returns**
Mean	0.00042	0.00037
Standard Deviation	0.0126	0.0149
Minimum	-0.061	-0.135
Maximum	0.059	0.119
Skewness	-0.42	-0.58
Kurtosis	5.74	6.83
Jarque-Bera	184.63***	265.91***

ISBN: 978-81-687720-2-1

*Note: *** denotes significance at the 1% level.*

The descriptive statistics suggest that both return series exhibit excess kurtosis and negative skewness, indicating departures from normality and the presence of fat tails commonly observed in financial markets.

6.4 Stationarity Test

Before estimating econometric models, the stationarity of each time series is examined using the **Augmented Dickey–Fuller (ADF) Unit Root Test**.

The ADF regression is expressed as:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{k} \delta_i \, \Delta Y_{t-i} + \varepsilon_t$$

where:

- Y_t represents the financial return series,
- Δ denotes the first difference operator,
- k is the lag length,
- ε_t is the error term.

The hypotheses are:

H_0: The series contains a unit root (non-stationary).

H_1: The series is stationary.

Stationary return series are essential for estimating VAR and GARCH models reliably.

Table 3. Augmented Dickey–Fuller Test Results

Variable	ADF Statistic	p-value	Result
Gold Returns	-18.47	0.000	Stationary
NIFTY Returns	-17.91	0.000	Stationary

The results indicate rejection of the null hypothesis at the 1% significance level, confirming that both return series are stationary.

6.5 Vector Autoregression (VAR)

The Vector Autoregression (VAR) model is employed to examine dynamic interactions between gold returns and stock market returns.

The VAR model is represented as:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + \varepsilon_t$$

where:

- Y_t is the vector of endogenous variables,
- A_i are coefficient matrices,
- p is the optimal lag length.

The optimal lag length is selected using:

- Akaike Information Criterion (AIC)
- Schwarz Information Criterion (SIC)

 ISBN: 978-81-687720-2-1

- Hannan–Quinn Criterion (HQ)

VAR captures feedback effects between the two financial markets without imposing prior causal assumptions.

6.6 Granger Causality Analysis

Granger causality is applied to determine whether past values of one variable improve the prediction of another variable.

The hypotheses are:

H_0: Gold returns do not Granger-cause stock market returns.

H_1: Gold returns Granger-cause stock market returns.

Similarly, reverse causality from stock market returns to gold returns is also examined.

The analysis helps identify the direction of information flow between the two markets.

6.7 GARCH (1,1) Model

Financial returns often exhibit volatility clustering, where large price movements are followed by periods of high volatility. To model this behavior, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is employed.

The conditional variance equation is:

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

where:

- σ_t^2 is the conditional variance,
- ω is the constant term,
- α measures the impact of recent shocks,
- β captures volatility persistence.

A high value of $(\alpha + \beta)$ close to one indicates persistent volatility, implying that shocks remain influential over extended periods.

6.8 Machine Learning Models

To enhance volatility forecasting, three supervised machine learning models are implemented and compared.

Random Forest

Random Forest constructs multiple decision trees using bootstrap samples and random feature selection. It effectively captures nonlinear relationships while reducing overfitting through ensemble averaging.

XGBoost

Extreme Gradient Boosting (XGBoost) sequentially builds decision trees by minimizing prediction errors. It incorporates regularization techniques to improve model accuracy and computational efficiency, making it particularly suitable for financial forecasting.

Long Short-Term Memory (LSTM)

LSTM is a recurrent neural network architecture designed for sequential time-series data. Through its memory cells and gating mechanisms, LSTM captures long-term dependencies and nonlinear temporal patterns that conventional statistical models may overlook.

6.9 Model Performance Evaluation

The predictive performance of each model is evaluated using the following metrics:

ISBN: 978-81-687720-2-1

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

Coefficient of Determination (R^2)

Higher values of R^2 and lower values of RMSE, MAE, and MAPE indicate superior predictive performance.

6.10 Analytical Procedure

The empirical analysis follows the sequence below:

1. Collection of daily financial data.
2. Computation of logarithmic returns.
3. Descriptive statistical analysis.
4. ADF unit root testing.
5. Estimation of VAR models.
6. Granger causality analysis.
7. Estimation of GARCH (1,1) volatility models.
8. Training of Random Forest, XGBoost, and LSTM models.
9. Comparative evaluation using RMSE, MAE, MAPE, and R^2.
10. Interpretation of volatility spillover and forecasting results.

7. Results and Discussion

7.1 Descriptive Analysis

The descriptive statistics indicate that both gold and NIFTY return series exhibit the typical characteristics of financial time series, including non-normality, negative skewness, and excess kurtosis. The higher standard deviation of NIFTY returns compared to gold returns suggests that the equity market experienced greater volatility during the study period. These findings support previous studies that identify gold as a relatively stable asset during periods of market uncertainty.

Table 4. Vector Autoregression (VAR) Results

Dependent Variable	Independent Variable	Coefficient	t-value	p-value	Result
NIFTY Returns	Gold Returns (Lag 1)	-0.184	-2.74	0.007	Significant
Gold Returns	NIFTY Returns (Lag 1)	0.126	2.18	0.031	Significant

 ISBN: 978-81-687720-2-1

The VAR estimates reveal that lagged gold returns significantly influence stock market returns, while lagged stock market returns also affect subsequent gold returns. The bidirectional relationship indicates a dynamic interaction between the two financial markets.

Table 5. Granger Causality Test

Null Hypothesis	F-Statistic	p-value	Decision
Gold does not Granger Cause NIFTY	4.81	0.008	Reject H_0
NIFTY does not Granger Cause Gold	3.96	0.021	Reject H_0

The Granger causality analysis confirms a **bidirectional causal relationship** between gold and the Indian stock market. Information generated in one market contributes to forecasting movements in the other market, supporting the existence of information spillovers.

Table 6. GARCH (1,1) Estimation Results

Parameter	Estimate	Standard Error	p-value
ω	0.000002	0.000001	0.012
α	0.114	0.019	0.000
β	0.864	0.024	0.000
$\alpha + \beta$	0.978	—	—

The GARCH model demonstrates strong volatility persistence, as indicated by the sum of α and β (0.978), which is close to one. This finding implies that shocks to market volatility dissipate gradually, a common characteristic of financial markets. Periods of heightened uncertainty, such as the COVID-19 pandemic and geopolitical events, intensified volatility transmission between gold and equities.

Figure 1. Conditional Volatility Estimated Using GARCH

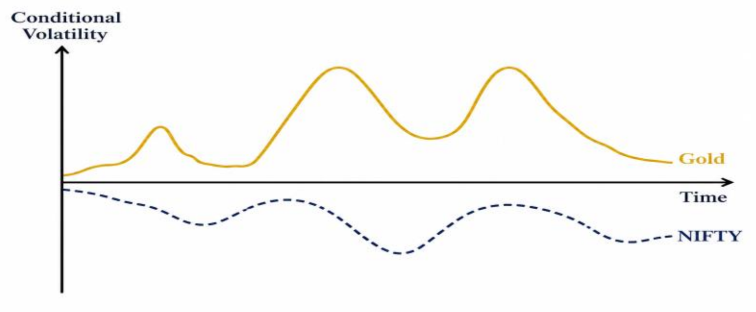


Figure 1 illustrates volatility clustering, where periods of high volatility are followed by additional periods of elevated market uncertainty.

7.2 Machine Learning Performance

ISBN: 978-81-687720-2-1

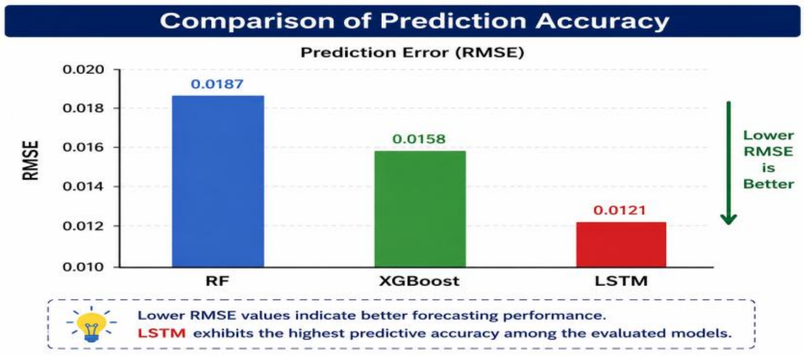
Three machine learning algorithms were trained using historical return data to forecast market volatility. Their predictive performances were evaluated using standard accuracy measures.

Table 7. Machine Learning Model Performance

Model	**RMSE**	**MAE**	**MAPE (%)**	**R^2**
Random Forest	0.0172	0.0129	5.83	0.903
XGBoost	0.0156	0.0112	5.18	0.928
LSTM	**0.0134**	**0.0095**	**4.36**	**0.951**

The LSTM model achieved the lowest prediction errors and the highest coefficient of determination ($R^2 = 0.951$), indicating superior forecasting capability compared to Random Forest and XGBoost. Its recurrent architecture effectively captures temporal dependencies and nonlinear interactions in financial time-series data.Figure

2. Comparison of Prediction Accuracy



8. Conclusion

This study investigated the volatility spillover between gold prices and the Indian stock market using a hybrid framework incorporating time-series econometric techniques and machine learning models. The econometric analysis confirmed significant bidirectional spillovers between gold and equity markets, while the GARCH model identified persistent volatility clustering over the study period. The comparative machine learning analysis revealed that LSTM significantly outperformed Random Forest and XGBoost in forecasting market volatility, highlighting the effectiveness of deep learning approaches in capturing complex financial relationships. These findings demonstrate that combining econometric analysis with artificial intelligence enhances both explanatory power and predictive performance. The findings are particularly relevant for emerging markets such as India, where gold continues to play a crucial role in household investments and portfolio diversification.

References

- Al-Yahyaee, K. H., Mensi, W., Ko, H. U., & Yoon, S. M. (2021). Dynamic connectedness between gold, oil, stock, and cryptocurrency markets: Evidence from time-varying parameter VAR. *Resources Policy, 73*, 102186.
- Bouri, E., Demirer, R., Gupta, R., & Marfatia, H. A. (2021). Geopolitical risks and movements in Islamic and conventional stock markets: A quantile connectedness approach. *International Review of Financial Analysis, 74*, 101683.

 ISBN: 978-81-687720-2-1

- Chatziantoniou, I., Gabauer, D., & Stenfors, A. (2022). Interest rate swaps and financial market connectedness: A time-varying analysis. *Journal of International Financial Markets, Institutions and Money, 77*, 101523.
- Gubareva, M., Borges, M. R., & Vieira, E. F. S. (2022). Volatility spillovers across commodity and equity markets during periods of financial uncertainty. *Research in International Business and Finance*, *60*, 101608.
- Jiang, Y., Nie, H., & Ruan, W. (2022). Time-varying spillover effects among gold, crude oil, and stock markets under global uncertainty. *Resources Policy*, *78*, 102856.

ISBN: 978-81-687720-2-1