



ANALYSING FORECASTING ACCURACY OF GARCH MODELS

RACHH KRISHNA MUKESH

Registration number : **2311129**

Supervisor's name : **Chronopoulos, Ilias C**

Word count : **8800**

Date submitted : **10th September 2024**

Submitted in partial fulfilment of the requirements for the award of the degree of
MSc in Financial Engineering and Risk Management

CONTENTS

1. ABSTRACT.....	3
2. INTRODUCTION	4
3. LITERATURE REVIEW	5
3.1 Return and Volatility characteristics:	6
3.2 Evolution of GARCH family of volatility models:.....	7
3.3 Forecasting:.....	9
4. DATA and METHODOLOGY.....	10
4.1 Modeling Conditional Variance:	11
4.1.1 GARCH(1,1) process:.....	11
4.1.2 GJR-GARCH	11
4.1.3 TARCH	12
4.1.4 E-GARCH.....	12
4.2 Analysing Forecasting Performance:	13
4.3 Model Diagnostics:	15
4.3.1 ACF of residuals:	15
4.3.2 Ljung-Box Test:	15
4.3.3 Jarque-Bera Test:.....	16
5. EMPIRICAL ANALYSIS	17
5.1 Evaluating Model Performance:	17
5.2 Evaluating Forecasting Performance:	19
6. CONCLUSION:.....	22
7. REFERENCES:	23

1. ABSTRACT

The paper discusses the importance of predicting volatility in financial markets and the challenges of accurately measuring and predicting volatility due to the complexity of the markets and the amount of data needed. Volatility is crucial to financial decision-making, risk management, and policymaking, because increased volatility increases risk. Engle's 1982 ARCH model is extended to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. The GARCH model predicts volatility. Moreover, it covers a sample period of 13 years from 1996-2009, which is relatively long for the volatility forecasting studies. This study offers a diverse selection of parameterization options in the GARCH models, and it incorporates a broad array of performance evaluation criteria. Based on the diagnostics tests and forecasting errors, it is evident that the standard GARCH(1,1) model performs well, but the TAR(1,1) model, which considers threshold effects in volatility, proves to be superior in terms of both estimation and forecasting accuracy thus demonstrating exceptional predictive ability. According to the analysis, it seems that employing more sophisticated models, capable of capturing detailed patterns in volatility, does not necessarily translate into superior forecasting ability when contrasted with simpler models. These results show that the model effectively accounts for important components of market volatility, including asymmetry and leverage effects. However, there is still potential for improvement to increase the accuracy of future forecasts. The findings highlight the significance of model selection in volatility forecasting and propose areas for future research, such as investigating multi-horizon forecasting and applying volatility models to various financial markets. This approach aligns with the broader literature's work on enhancing the practical usefulness of volatility forecasts in financial markets.

2. INTRODUCTION

Volatility prediction is crucial in financial markets due to its wide-ranging impact on various aspects of financial decision-making, risk management, and policy formulation. Risk has become an important part of financial analysis as different people perceive different levels of risk given different level of returns. Though risk and volatility are perceived as two different things, in terms of modelling or forecasting volatility, higher the degree of volatility of an asset's price higher is its risk. Volatility forecasting can prove to enhance decision-making across various facets of financial markets by providing accurate predictions of future market fluctuations. Measuring volatility is a complex task, with numerous approaches available to tackle it. Considering the perspective of an investor who is presented with investment opportunities and has access to returns y_t at time t and information F_s at time $s < t$, it is worth exploring the matrix $\text{Var}(y_t | F_s)$.

Dealing with this situation can be quite challenging as the time horizon $t - s$ may not be known or could be subject to randomness. Additionally, the information set F_s could be quite extensive, encompassing current and past values of numerous variables. Furthermore, the probability distribution $f(y_t | F_s)$ may not be known. The recent focus on continuous time methods of volatility measurement and forecasting addresses all these concerns to some extent. One such fairly advanced is the GARCH model, an extension on the ARCH model introduced by Engle (1982). Many authors have used GARCH family models to forecast volatility, such as Pagan and Schwert (1990), Franses and Dijk (1996), Brailsford and Faff (1996), and Corrado and Miller (2005).

The central aim of this paper is to empirically examine the forecasting accuracy of most widely used GARCH models like GARCH, GJR-GARCH, TARCH, EGARCH, for a set comprising some of the major world stock index DJIA, Nasdaq, SP500, CAC, FTSE just to name some as well as some foreign exchange rates USD/GBP, USD/EUR, etc.

3. LITERATURE REVIEW

Understanding and predicting volatility is of utmost importance in financial markets, as it has far-reaching effects on important aspects such as financial decision-making, risk management, and reserve bank's policy formulation.

Investors must strategically mitigate the risk linked to their investment portfolios. Gaining a comprehensive understanding of volatility prediction is essential for evaluating the possible risk associated with asset returns and for constructing effective portfolios that optimize returns while mitigating risk. Financial institutions depend on Value at Risk (VaR) to evaluate the plausible depreciation in the worth of their portfolios throughout a set period, considering a certain degree of certainty. Ensuring accurate forecasts of volatility significantly enhances the reliability of VaR computations. Regulatory bodies and financial organizations do stress tests to assess the resilience of financial systems under difficult situations. The ability to forecast volatility allows them to model possible unfavourable situations and formulate tactics to minimize their consequences.

Models such as the Black-Scholes model heavily depend on the volatility factor as a crucial input. Having precise volatility forecasts is crucial for ensuring accurate pricing of options, which helps to avoid any mispricing or potential arbitrage opportunities. Using derivatives, traders employ hedging strategies to protect themselves from potential negative shifts in asset prices. Having accurate predictions of volatility allows for the development of stronger hedging strategies, which can help minimize the potential for substantial financial losses.

In order to detect any potential issues that might impact financial stability, regulatory agencies diligently evaluate volatility. Anomalous increases in volatility might suggest fundamental problems in financial markets or in institutions which requires government intervention. Central banks, when formulating their policies, incorporates volatilities to assess what factor is affected by it. Significant fluctuations within the financial sector have the potential to greatly influence economic growth and inflation, subsequently impacting decisions related to interest rates and other monetary policies.

Corporate entities also rely on volatility projections to assess the level of risk associated with new projects and investments, akin to the assessment made by a financial adviser. The precise evaluation of risk is essential for making well-informed financial choices and establishing the most advantageous capital structure. The cost of capital for firms is determined by the level of risk associated with their equity and debt.

Volatility forecasting is of utmost importance in the domain of algorithmic and high-frequency trading (HFTs), as it aids in informing real-time trading choices. A comprehensive understanding of volatility models may significantly enhance trading algorithms, resulting in high profitability and improved risk management. High-frequency traders meticulously analyse volatility patterns to exploit short-term price inefficiencies. Precise volatility forecasts allow traders to fine-tune their execution strategies and minimize transaction expenses.

3.1 Return and Volatility characteristics:

The total return is the aggregate of the increase in value of an investment plus any payments received as dividends. One can modify a price time series by including dividends to create a total return index. If the price were represented as a total return index, the returns created would indicate the total return that an asset holder would get during a specific time, including both capital gains and dividend payments. (Brooks, 2014).

Asset returns have following stylized facts. Firstly, returns do not follow a normal distribution. In most cases the distribution is symmetric (Taylor, 2005), however, in some cases, empirical distribution has an asymmetric shape, normally skewed to the left and high kurtosis (Mandelbrot, 1963) and fat tails (Fama, 1965). Secondly, there is almost no autocorrelation between returns of an asset during consecutive days. There is positive dependence between absolute returns on nearby days and thus for squared returns (Taylor, 2005). Returns on different assets tend to be positively correlated, that is market activity is strongly correlated with price variability. Lastly, the average of short-term (daily) returns in the long term is expected to be zero.

Return variability over a certain length of time is measured by volatility. As the literature reflects, one of the most crucial aspects in finance is volatility. It is a common indicator of an asset's risk. It is measured in several Value-at-Risk models for the market risk assessment. Its essential in deriving the price of traded options and other theoretical asset pricing models including the Sharpe model is also critical for the Black-Scholes formula. Though there is a lot of research on this topic not all of them concentrate on volatility's predicting capability.

Volatility is not constant over time and there seems to be no plausible explanation its variability, yet it can be predicted with some certainty. It is common that large returns (of any sign) are followed by large returns (of any sign) and vice versa. Such phenomena are called volatility clustering and its possible that it occurs because information also arrives in a short period of time (Mandelbrot, 1963). High volatility produces more dispersion in returns than low volatility, thus returns are more spread out if volatility is higher. A high volatility cluster will contain several large positive returns as well as negative returns, similarly during low volatility regime, price changes in returns will be minimal (Taylor, 2005). Asset price fluctuations exhibit periods of low volatility, during which the price changes are minimal and occur irregularly, followed by high volatility regime, during which prices tend to rise (Gauersdorfer and Hommes, 2007). One possible explanation for this phenomenon is that new information that impacts the price tends to arrive in clusters rather than being equally distributed throughout time (Brooks, 2014).

The announcement effect, or in other words, the arrival of new information to the market of some economic factors such as gross national product (GNP), inflation and unemployment rate could also have an impact in volatility. In the literature there is an ongoing discussion Flannery and Protopapadakis (2002) found that macroeconomic factors such as real GNP and money growth, despite that they affect returns, are associated with lower return volatility. In the other hand, Glosten et al. (1993) found that inflation has a strong impact on return volatility.

The leverage effect is that negative returns tend to increase future volatility more than positive returns of the same magnitude. Thus, is also an important feature that impacts

forecasting volatility. The leverage effect was first observed by Black in 1976 in his findings that stock prices and volatility are negatively correlated. Whenever the firm's stock price falls, the debt-to-equity ratio of the firm rises, and higher debt means higher financial leverage; with increases in financial leverage, the firm's risk gets magnified, resulting in higher volatility. Christie refined Black's research with empirical data to demonstrate the presence of the leverage effect in stock markets. He stated that, when the equity value of the company drops, the financial leverage of the firm rises, giving the increased volatility. For example, a company with debt and equity outstanding, the leverage effect is normally higher when the value of the firm falls, which raises equity return's volatility if the whole is constant (Bollerslev et al., 1994). The increased leverage enhances the risk of the firm and thus implies a mean future volatility. The leverage effect can be given a wider explanation when applied in the context of Merton's option pricing theory (Merton 1973). According to this theory, Merton viewed a firm's equity as a call option on its assets with a strike price equal to the value of liabilities. Value decrement is an increment in the volatility of the option, leading to an increase in the volatility of the firm's equity. Geske (1977) was one of the first to develop this idea further by relating equity volatility to firm leverage within the framework of contingent claims.

Bekaert and Wu (2000) suggested that the leverage effect might be an outcome of asymmetric information between various market participants. Negative news or uncertainty can create more information asymmetry, which heightens perceived risk and, therefore, rollover volatility. Studies by Andersen et al. (2007) have examined the presence of leverage effects across different markets and other asset classes as well like commodities, bonds, and foreign exchange.

In periods of crisis volatility is expected to increase and tend to be more stable, thus lower, during periods of economic stability. Schwert (1989) states that the average level of volatility is much higher in periods of recession and that, particularly for financial asset returns, volatility increases during periods of financial market crisis. This conforms with the studies by Bollerslev, Sizova, and Tauchen (2012) wherein they found leverage effect is more prominent in periods of high volatility.

In this paper we include several models which explain this leverage effect and accurately predict the volatility in presence of these asymmetries.

3.2 Evolution of GARCH family of volatility models:

Prior to the 20th century, there was a need for understanding and modeling volatility. Quantitative study was scarce, and most of the observations were anecdotal. Louis Bachelier's PhD thesis, titled "The Theory of Speculation," brought about a significant change in the understanding of asset volatility. This was achieved by introducing the notion of modeling by employing the concept of Brownian motion. Harry Markowitz revolutionized investment strategies by creating "Modern Portfolio Theory (MPT)" introducing portfolio optimisation techniques using expected returns and variance (risk measure) as its metric. MPT highlights the significance of comprehending and forecasting volatility.

Robert F. Engle was instrumental in developing the ARCH model for accounting time-varying volatility or famously known as volatility clustering. The ARCH model proposed by Engle (1982) states that the conditional variance of a time series is determined by the squared

residuals from prior time periods. This model has verified the presence of volatility clustering, which means that times of high volatility are followed by periods of high volatility, and periods of low volatility are followed by periods of low volatility. The ARCH model employs lots of parameters making it computationally inefficient and less reliable.

Tim Bollerslev (1986) expanded the ARCH model to the Generalised ARCH (GARCH) model, which incorporated lagged values of both squared returns and past variances. This proved to be more flexible and parsimonious for modeling volatility. Despite this, the limitation of GARCH model was that it couldn't account for the leverage effect nor the non-normal distribution of volatility so models like GJR-GARCH, EGARCH, TGARCH were introduced.

Introduced by Nelson (1991), the EGARCH model captures the asymmetric effects of positive and negative shocks on volatility, addressing the leverage effect observed in financial markets. Zakoian (1994) and Glosten, Jagannathan, and Runkle (1993) developed the GJR-GARCH (or TGARCH) model to allow for different responses to positive and negative shocks. Engle and Bollerslev (1986) introduced the IGARCH model, which imposes a unit root in the GARCH process, implying that shocks have a permanent effect on volatility.

GARCH models are the standard tools for volatility forecasting. GARCH models predict volatility better than simple and exponentially weighted moving averages, according to several research. In their comprehensive analysis, Poon and Granger (2003) examined several models for predicting volatility and determined that GARCH models, namely those that incorporate asymmetries such as EGARCH and TGARCH, offer more precise forecasts compared to non-GARCH models. Another research by Andersen, Bollerslev, Diebold, and Labys (2003) found that GARCH models produce superior volatility forecasts compared to stochastic volatility models when applied to high-frequency data.

Many econometricians have tested also tested more sophisticated GARCH-family models and concluded an unimpressive forecasting performance. Studies conducted by Pagan and Schwert (1990) showed EGARCH outperforming GARCH model in predicting monthly volatility of US stocks. Bali and Demirtas (2008) concluded that EGARCH is the best model for predicting SP500 index futures volatility. Cao and Tsay (1992) that when it comes to long-horizon forecasts for small-cap stocks, EGARCH reigns supreme. However, for large stocks, other models perform well. Ederington and Guan (2005) found that there were many similarities between the GARCH and EGARCH models when analysing the volatility of multiple asset classes. Lee (1991) showed that the accuracy assessment criteria had an impact on the predictive ability of GARCH models in forecasting future outcomes. In the study by Brailsford and Faff (1996), they discovered that the GJR(1,1) model demonstrated superior performance compared to other models. This conclusion was based on the evaluation of various loss criteria, including root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). When it comes to forecasting exchange rate volatility, Brooks and Burke (1998) discovered that GARCH(1,1) is favoured based on the MSE loss criterion, although it may not be the best choice according to the MAE criterion. The EGARCH (GJR) model is the most accurate (least accurate) for predicting fluctuations in exchange rate volatility (Balaban 2004). However, the performance rankings were sensitive to the choice of loss criteria. In this paper we will compare different criteria and find out superior models for our indices for the period of 1996-2009.

In most cases, a heavily parameterized model is better able to capture the multiple dimensions of volatility dynamics, it would generally have a better in-sample fit than a relatively parsimonious model. However, a better in-sample fit may not necessarily translate into a better out-of-sample forecasting performance. On the out-of-sample forecasting ability, the simpler models often outperform the more complex models.

3.3 Forecasting:

Realized volatility quantifies historical market fluctuations. The realized variance (RV) is an accurate measure of volatility when prices are watched continuously and without any measurement mistake. It is calculated by summing the squared intraday returns. According to Andersen and Benzoni (2008), the realized variance and realized return may be accurately evaluated using constantly observed quotation data, assuming there are no transaction costs. According to Taylor (2005), increasing the frequency of returns can lead to a more accurate estimation of volatility within a certain time, as long as the returns within that period are not linked and certain other requirements are met.

The evaluation of forecasting performance is a non-trivial task due to following two reasons. Firstly, the true conditional variance σ_t^2 is not directly observable in the market at a particular instant and secondly there is no unique criterion for selecting the best forecasting model and therefore depends on users' discretion and market conditions. For the first issue Anderson et. al. (2003), in their paper concluded that high-frequency data based realized variance is a good approximator for the true conditional variance. But it is not practical to obtain high-frequency intra-day data for all the stock indices as it is computationally complex owing to large number of observations, asynchronous and irregularly spaced timestamps and the microstructure noise. To build upon this theory Yang and Zhang, (2000). Alizadeh *et al.* (2002) found that empirically range-based volatility proxies are highly efficient and robust to the microstructure effects. According to Shu and Zhang (2006), the range-based estimators offer highly accurate predictions of the daily integrated volatility. In their study, Jacob and Vipul (2008) discovered that the volatility predictions derived from range-based estimators are almost as effective as those derived from the realized variance estimator. So, in the data by 'Oxford Mans Institute the daily realized variance is based on the sum of 5-minute intra-day squared returns.

The second issue is identifying a suitable performance evaluation criterion, as described earlier many economists have found different loss functions for selecting best forecasting model. We use MSFE metric to assess and quantify the accuracy of our forecasts. The general approach is to evaluate forecast on multiple loss functions like Root mean squared forecast error (RMSFE) and will discuss in coming sections.

The remaining paper is organized as follows. A section describing the data set used in this analysis and the methodology used to compare the forecasting performance of the GARCH models along with definition of different models. Next section discusses the empirical results, and one which concludes the article.

4. DATA and METHODOLOGY

The dataset primarily consists of the daily closing price and realized variance for the major global indices like DJIA, Nasdaq, S&P500, CAC, FTSE including many others; as well as some major world currency pairs like USDEUR, USDGBP, USDYEN, etc. taken from the database “Oxford-Man Institute’s realised library,” Version 0.1, produced by Lunde and K. Sheppard over the period from January 3, 1999, to December 23, 2007. The daily realized variance is estimated based on the sum of 5-minute intra-day squared returns.

For modelling purpose, we consider full sample period from January 3, 1996, to February 27, 2009. For forecasting, the in-sample period is from January 3, 1996, to January 2, 2005, and the remaining period will be used for the out-of-sample evaluation that is from January 2, 2005, to February 27, 2009. As volatility is latent, the realized variance is used as proxy for the true volatility.

For simplicity, the code and methodology we develop for the DJIA can be easily adapted to other indices since the structure of the data and the analytical approach remain consistent across index. Table 1 presents the descriptive statistics and Figure 1 below plots the returns of DJIA.

Table 1 Descriptive statistics for DJIA returns series

Mean	9.66E-05
Median	0.000495
Maximum	0.105083
Minimum	-0.082005
Std. Dev.	0.012279
Skewness	-0.103638
Kurtosis	11.19676
Jarque-Bera	9134.854
Probability	0
Sum	0.314928
Sum Sq. Dev.	0.491489
Observations	3261

Table 2 ADF test for DJIA returns series.

Augmented Dickey-Fuller test statistic		t-stat	p-value
		-57.68548	0.0001
Test critical values:	1% level	-3.432608	
	5% level	-2.862423	
	10% level	-2.567285	

Figure 1: DJIA Return series.

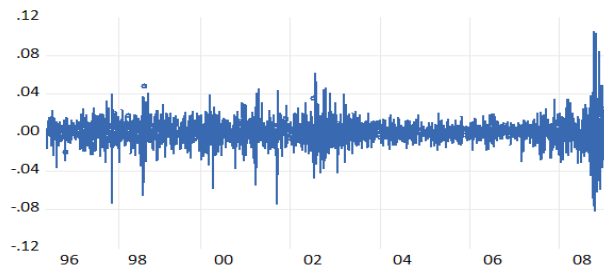


Table 1 shows that the mean and the median are consistently close to zero. As far as the values of skewness and kurtosis are concerned, for a normal distribution, they should be zero and three, respectively. Here, the negative skewness indicates asymmetric distributions skewed to the left, while the kurtosis indicates returns show leptokurtic characteristic. The evidence of non-normality is further supported by the Jarque–Bera test statistic which rejects the null hypothesis of normal distribution at the 1% level. Similarly, Table 2 shows the summary statistics of the IV indices along with Augmented Dickey–Fuller (ADF) test for unit roots. The t-stat of the ADF tests show that DJIA return series is stationary at 1%, 5% as well as 10% significance levels.

4.1 Modeling Conditional Variance:

4.1.1 GARCH(1,1) process:

We define a GARCH model for y_t as,

$$y_t - \mu_t = \epsilon_t = \sigma_t z_t \quad (\text{eq. 1})$$

where,

z_t : an i.i.d with mean equal to 0 and variance equal to 1 i.e. $E(z_t) = 0$; $\text{Var}(z_t) = 1$
 μ_t, σ_t^2 : conditional mean and conditional variance of y_t i.e. $\mu_t = E(y_t | F_{t-1}) = E_{t-1}(y_t)$
and $\sigma_t^2 = \text{Var}(y_t | F_{t-1}) = \text{Var}_{t-1}(y_t)$

The GARCH(1,1) equation,

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \epsilon_{t-1}^2 \quad (\text{eq. 2})$$

where ω, β , and α are variable parameters.

The positivity of σ_t^2 is ensured by the following sufficient restrictions: $\omega > 0$, $\alpha \geq 0$, and $\beta \geq 0$

To understand why the GARCH(1,1) equation together with (Eq. 1.1) and the assumptions stated above is able to account for volatility clustering and leptokurtosis, let us note that the autocorrelation coefficients of ϵ_t^2 , denoted by ρ_j , are equal to

$\rho_1 = \alpha(1 - \beta^2 - \alpha\beta)/(1 - \beta^2 - 2\alpha\beta)$, which is larger than α , and $\rho_j = (\alpha + \beta)\rho_{j-1}$ for $j \geq 2$, if $\alpha + \beta < 1$.

The last inequality ensures that $\text{Var}(\epsilon_t) = \omega/(1 - \alpha - \beta)$ (denoted by σ^2) exists and that ϵ_t is covariance stationary. Moreover, the autocorrelations of ϵ_t^2 are positive and decaying at the rate $\alpha + \beta$.

The sum $\alpha + \beta$ is referred to as the persistence of the conditional variance process. For financial return series, estimates of α and β are often in the ranges $[0.02, 0.25]$ and $[0.75, 0.98]$, respectively, with α often in the lower part of the interval and β in the upper part for daily series, such that the persistence is close to but rarely exceeding 1. Hence, ρ_1 is typically small, and the autocorrelations decay slowly.

The general form of GARCH(p,q) model can be written as,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (\text{eq. 3})$$

where, there are p MA lags and q AR lags.

4.1.2 GJR-GARCH

The GJR-GARCH model named after Glosten, Jagannathan and Runkle, is an extension of the traditional GARCH model designed to capture the asymmetry in financial time series data, that is, it accounts for modeling leverage effect.

The GJR model is an extension of GARCH model wherein it incorporates an additional term that captures asymmetry. The conditional variance for the GJR-GARCH(p,q) model is given by:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p (\alpha_i + \gamma_i I(\epsilon_{t-i} < 0)) \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (\text{eq. 4})$$

In this model, α_0 is the constant term, α_i is the coefficient for the lagged squared residual, γ_i is the coefficient for the asymmetric effect, I_{t-i} is the indicator function, β_j is the coefficient for the lagged conditional variance.

If there was a negative shock at time t-1 (i.e. $\epsilon_{t-1} < 0$), the conditional variance σ_t^2 will be increased by $\gamma_i \epsilon_{t-i}^2$, reflecting the higher impact of negative news on volatility.

4.1.3 TARCH

It is similar to GJR-GARCH in terms of capturing asymmetry in the response of volatility to positive and negative shocks. TARCH adds a term that multiplies the squared residuals by an indicator function given as follows,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p (\alpha_i \epsilon_{t-i}^2 + \gamma_i \epsilon_{t-i}^2 I(\epsilon_{t-i} < 0)) + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (\text{eq. 5})$$

However, in the GJR-GARCH model, the asymmetry is captured by modifying the impact of past squared residuals based on whether the residuals were positive or negative.

Within the TARCH model, the parameter γ directly quantifies the incremental effect of negative shocks on volatility in comparison to positive shocks. If the value of γ is greater than 0, it suggests that negative shocks exhibit a greater impact on volatility compared to positive shocks (leverage effect).

The GJR-GARCH model adapts the sensitivity of the conditional variance to previous squared residuals by modifying the parameter γ , depending on the sign of the residuals. A positive coefficient γ implies that negative shocks exert a relatively more significant influence on future volatility compared to positive shocks, therefore suggesting the presence of a leverage effect.

4.1.4 E-GARCH

Nelson (in 1991) introduced the Exponential GARCH model, which is given as,

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \alpha_i \left(\frac{\epsilon_{t-i}}{\sigma_{t-i}} \right) + \sum_{i=1}^p \gamma_i \left(\left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| - E \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| \right) + \sum_{j=1}^q \beta_j \ln(\sigma_{t-j}^2) \quad (\text{eq.6})$$

The E-GARCH model uses the logarithm of the conditional variance at time t, which ensures that the variance is always positive without imposing non-negativity constraints on the parameters. The term $\left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| - E \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right|$ captures the asymmetry in the impact of shocks on

volatility so that positive and negative shocks have different reactions on predictive power of future volatility.

For each index, we generate one-step-ahead forecasts, using fixed-size windows and Recursive window of the last observations, re-estimating the parameters for each period. The parameters are estimated using the method of maximum likelihood. The selection of the estimation window is primarily subjective, as there is no specific theoretical prescription for the optimal size of the estimation window. Generally, using a large estimation window is statistically desirable, as it increases the efficiency of the parameter estimates. A practical consideration is to use a sufficiently long estimation window to avoid the non-convergence problem in the likelihood-based estimation, particularly when estimating the more richly parameterized GARCH models. Using a long estimation window may be detrimental under an equal-weighting scheme, like the moving average model, as the forecasts would get influenced by the past information that may no longer be relevant. However, this consideration would not apply to the GARCH models that have exponentially decaying weights, which limit the influence of the data that is too old.

4.2 Analysing Forecasting Performance:

We use MSFE metric to assess and quantify the accuracy of our forecasts. The general approach is to evaluate forecast on multiple loss functions like Root mean squared forecast error (RMSFE).

RMSE penalizes greater errors more strongly since the errors are squared before averaging. This indicates that when the spread or variation of the mistakes is of concern, RMSE is a particularly relevant indicator. If one desires to assign greater significance to larger errors (which may have a more detrimental effect in certain applications), the Root Mean Square Error (RMSE) proves to be efficient as it magnifies the influence of larger deviations from the true values by squaring them. For example, if one forecast has minor mistakes generally but a few large errors, the RMSE will be higher than if the forecast has consistently modest errors. This is of utmost importance in our case since outliers or extreme deviations are significant.

The Mean Squared Error (MSE), while useful, measures the average squared error, which makes its units different from the real values. Discrepancy in scale complicates the interpretation of MSE as it squares the units of the original data. For example, if the data is in dollars, MSE would be expressed in square dollars.

RMSE, on the other hand, returns the error metric back to the same scale as the original projected numbers. Since RMSE is the square root of MSE, it has the same units as the target variable. This makes it easy to understand how far, on average, the predictions are from the actual numbers. For instance, if you're projecting stock prices, an RMSE of \$5 would suggest that, on average, the anticipated values depart from the actual prices by \$5, which is more interpretable than an MSE of 25 (square dollars).

RMSFE can be calculated using the formula:

$$MSFE(\widehat{\sigma}_t^2, \sigma_t^2) = \sqrt{\frac{1}{T - T_0} \sum_{t=T_0+1}^T (\widehat{\sigma}_t^2 - \sigma_t^2)^2} \quad (\text{eq. 7})$$

over $t \in (T_0, T]$ and where, $\widehat{\sigma}_t^2$ is the forecasted variance for day t ; σ_t^2 is the realized variance for day t .

To make comparisons across different forecasting methods, we use the benchmark MSE_{GARCH} of the parametric GARCH (1, 1) volatility model and calculate the relative root quadratic loss,

$$RMSFE_j = \frac{(MSFE_j)^{1/2}}{(MSFE_{GARCH})^{1/2}} \quad (\text{eq. 8})$$

The best performing model will be the one with lowest RMSFE for that given underlying index.

It is commonly believed that out-of-sample tests are highly regarded as the benchmark for evaluating forecasts. The accuracy of estimated models should be assessed by conducting out-of-sample tests rather than using the same data that was used to estimate the model's parameters, which is known as a "in-sample" forecast. Studies by Bartolomei and Sweet (1989) and Pant and Starbuck (1990) demonstrate that even the most accurate predictions made within a given sample may not accurately forecast data outside of that sample. In addition, empirical studies have shown that in-sample forecasting performance is often less reliable than out-of-sample tests. This could be attributed to the susceptibility to outliers and data mining, as noted by White (2000). Thus, the out-of-sample forecast is considered the ultimate test of a forecasting model by econometricians and forecasters (Stock and Watson 2015, p. 571).

Forecasting models for each index are estimated using fixed length and recursive or expanding) window methods and evaluated based on their out-of-sample performance.

Fixed-windows forecasting uses data up to a specified date to generate all forecasts after that date. In this approach, a predetermined initial sample is utilized to estimate the model parameters, and the model generates forecasts based on this fixed sample size. This approach is more efficient in terms of computation as the parameters are estimated once and then applied to all future forecasts. However, it might face difficulties when it comes to handling structural changes or shifts in the underlying data over time, as it doesn't adjust to newer information once the initial sample is taken.

The model in recursive window forecasting continuously updates its sample size by incorporating all available new data points and re-estimating the parameters after each new observation. This approach is more flexible and responsive to changes in the data, allowing it to incorporate new information and potentially improve the accuracy of forecasts, especially

when dealing with evolving data patterns. Recursive methods can be more computationally demanding, but they provide a more reliable approach when there are trends or shifts in the time-series.

4.3 Model Diagnostics:

Model diagnostics play a crucial role in time series analysis and econometric modeling by evaluating the suitability and dependability of a model. Once a model has been fitted, diagnostics are used to verify that the assumptions of the model have not been violated and that the model accurately represents the data. In this discussion, we will explore the fundamental diagnostic tools and tests that are frequently employed in practical settings. Residuals refer to the discrepancies between the actual values and the values that are anticipated by the model. Examining residuals aids in detecting any discernible trends that the model has not accounted for, indicating potential model misspecification or the omission of a crucial variable.

4.3.1 ACF of residuals:

Introduced by Engle and Bollerslev, ACF of residuals is used to test for any remaining ARCH effects that might be present after fitting our models. It uses ACF of standardized residuals and squared standardized residuals. Significant autocorrelations in the squared residuals suggests the presence of conditional heteroscedasticity, indicating that higher order models of GARCH may be appropriate.

4.3.2 Ljung-Box Test:

The Ljung-Box test introduced by Ljung and Box in 1978, commonly referred to as the Ljung-Box Q-test, is a statistical test employed to detect autocorrelation at various lags in a dataset of time series. It is a portmanteau test that evaluates the null hypothesis of no autocorrelation among a set of lags.

The Ljung-Box test is commonly employed in time series analysis to ascertain if the residuals of a model exhibit independent distribution, which is a fundamental assumption in time series modeling.

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k} \quad (\text{eq. 9})$$

Where, n is the number of observations, h is the lags being tested, in our case we will test lags 10 and 20, $\hat{\rho}_k^2$ is the sample autocorrelation at lag k .

The Ljung-Box statistic Q follows a chi-square distribution with h degrees of freedom. If the p-value is below a chosen significance level (e.g., 0.05), we reject the null hypothesis, it indicates that the data exhibits autocorrelation at the tested lags.

After fitting our models, the Ljung-Box test is applied to the residuals to check for remaining autocorrelation. A significant result suggests that the model has not fully captured the time series structure.

4.3.3 Jarque-Bera Test:

The Jarque-Bera (JB) test is a statistical test used to assess whether a particular sample of data exhibits the same skewness and kurtosis as a normal distribution. The goodness-of-fit test examines whether a series follows a normal distribution, which is a widely used assumption in several statistical models.

The JB test is valuable for evaluating the normal distribution of residuals in regression or time series models, which is a crucial assumption in econometrics and time series analysis.

The JB test statistic is calculated as follows:

$$JB = \frac{6}{n} (S^2 + \frac{1}{4} (K - 3)^2) \quad (\text{eq. 10})$$

Where, n is the number of observations, S is the sample skewness, K is the sample kurtosis.

The JB statistic follows a chi-square distribution with 2 degrees of freedom. If the computed JB statistic is significantly large (i.e., the p-value is below a certain threshold like 0.05), we reject the null hypothesis, indicating that the data does not follow a normal distribution. Normality of residuals is an important assumption in time series analysis. The JB test will help verify our assumption.

In the next section upon fitting the models, we will use these 2 diagnostic tools to assess the fit of the model.

5. EMPIRICAL ANALYSIS

In this section the GARCH models namely GARCH(1,1), GJR-GARCH(1,1), TARCH(1,1), E-GARCH(1,1) has been fitted. To balance between model complexity and predictive accuracy, parsimonious models are favoured more often as they that avoid overfitting (Econometric Analysis, William Greene).

5.1 Evaluating Model Performance:

Upon fitting different GARCH models on DJIA return series, we get following output of optimal parameters and diagnostic tests in tabular format as follows:

Table 3 : Model Parameters

GARCH(1,1)		
	Normal	Students-T
Mean Equation		
mu	0.0005	0.0139
	<i>0.0000</i>	<i>0.0000</i>
Variance Equation		
omega	0.0000	0.0001
	<i>0.0000</i>	<i>0.0000</i>
alpha[1]	0.1000	0.9970
	<i>0.0000</i>	<i>0.0000</i>
beta[1]	0.8800	0.0029
	<i>0.0000</i>	<i>0.7330</i>
nu	-	25.1945
	-	<i>0.3700</i>

E-GARCH(1,1)		
	Normal	Students-T
Mean Equation		
mu	0.0005	-21.7580
	<i>0.0002</i>	<i>0.0000</i>
Variance Equation		
omega	-0.1023	0.4102
	<i>0.0060</i>	<i>1.0000</i>
alpha[1]	0.1744	334.9618
	<i>0.0000</i>	<i>0.2780</i>
beta[1]	0.9881	1.0000
	<i>0.0000</i>	<i>0.9970</i>
nu	-	3.7300
	-	<i>0.2900</i>

GJR-GARCH(1,1)		
	Normal	Students-T
Mean Equation		
mu	0.0002	-0.0201
	<i>0.0000</i>	<i>0.0000</i>
Variance Equation		
omega	0.0000	0.0000
	<i>0.0000</i>	<i>1.0000</i>
alpha[1]	0.0100	0.1618
	<i>0.2930</i>	<i>0.0000</i>
gamma[1]	0.2000	0.6007
	<i>0.0000</i>	<i>0.0000</i>
beta[1]	0.8700	0.5770
	<i>0.0000</i>	<i>0.0000</i>
nu	-	3.1200
	-	<i>0.0000</i>

TARCH(1,1)		
	Normal	Students-T
Mean Equation		
mu	0.0009	0.0003
	<i>0.0000</i>	<i>0.0149</i>
Variance Equation		
omega	0.0002	0.0002
	<i>0.0768</i>	<i>0.0000</i>
alpha[1]	0.0100	0.0071
	<i>0.4300</i>	<i>0.3420</i>
gamma[1]	0.1000	0.1121
	<i>0.0000</i>	<i>0.0000</i>
beta[1]	0.9200	0.9363
	<i>0.0000</i>	<i>0.0000</i>
nu	-	3.8600
	-	<i>0.0000</i>

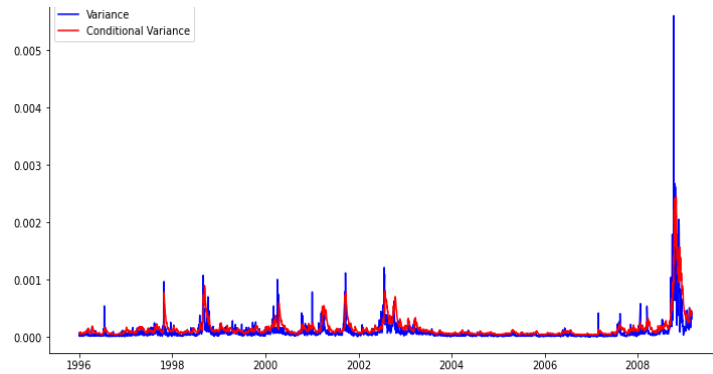
Note : parameter values and their corresponding p-values (in italics)

* indicates p-value at 1% level of significance and ** indicates 5% level

This visualization in Figure 2 below presents a comparison between realized variance (blue line) and computed conditional variance (red line) over time, using GARCH(1,1) model. The timeline from 1996 to around 2008 is represented on the x-axis, while the variance values are displayed on the y-axis. The blue line shows periodical sharp spikes, which coincide with times of heightened market volatility. The red line, which represents conditional variance, tends to closely follow the blue line, albeit with fewer pronounced peaks. The plot reveals a notable surge in both variances around 2008, suggesting a possible indication of the financial

crisis. The red line tends to smooth the realized variance, reflecting predicted behaviour over time.

Figure 2 : DJIA Realized Variance vs. Conditional Variance



By critically examining the Log-Likelihood estimation (in table 4) and significance of model parameters (in table 3) we can easily say that for GARCH(1,1) , E-GARCH(1,1) and GJR-GARCH(1,1) with normal distributions have outperformed those with t-distributions.

Table 4: Diagnostics tests

GARCH(1,1)		
	Normal	Students-T
Log-Likelihood	10322.30	8079.22
AIC	-20636.60	-16148.40
BIC	-20612.20	-16118.00
Ljung-Box Residuals (10)	9.92 <i>0.45</i>	240.70 <i>0.00</i>
Ljung-Box Residuals (20)	19.67 <i>0.48</i>	273.01 <i>0.00</i>
Ljung-Box Squared Residuals (10)	11.26 <i>0.34</i>	281.23 <i>0.00</i>
Ljung-Box Squared Residuals (20)	21.35 <i>0.38</i>	392.96 <i>0.00</i>
Jarque-Bera Test Statistic	389.74 <i>0.00</i>	7853.04 <i>0.00</i>

E-GARCH(1,1)		
	Normal	Students-T
Log-Likelihood	10335.40	-18873.70
AIC	-20662.80	37757.30
BIC	-20638.50	37787.80
Ljung-Box Residuals (10)	10.40 <i>0.41</i>	13726.33 <i>0.00</i>
Ljung-Box Residuals (20)	20.15 <i>0.45</i>	23893.08 <i>0.00</i>
Ljung-Box Squared Residuals (10)	7.77 <i>0.65</i>	13080.62 <i>0.00</i>
Ljung-Box Squared Residuals (20)	16.30 <i>0.70</i>	21416.23 <i>0.00</i>
Jarque-Bera Test Statistic	462.70 <i>0.00</i>	851106.71 <i>0.00</i>

GJR-GARCH(1,1)		
	Normal	Students-T
Log-Likelihood	10360.20	4939.35
AIC	-20710.40	-9866.71
BIC	-20679.90	-9830.17
Ljung-Box Residuals (10)	9.65 <i>0.47</i>	71.29 <i>0.00</i>
Ljung-Box Residuals (20)	17.83 <i>0.60</i>	110.04 <i>0.00</i>
Ljung-Box Squared Residuals (10)	17.97 <i>0.06</i>	123.10 <i>0.00</i>
Ljung-Box Squared Residuals (20)	32.33 <i>0.04</i>	134.27 <i>0.00</i>
Jarque-Bera Test Statistic	280.37 <i>0.00</i>	4239.15 <i>0.00</i>

TARCH(1,1)		
	Normal	Students-T
Log-Likelihood	10347.50	10438.20
AIC	-20685.00	-20864.40
BIC	-20654.60	-20827.80
Ljung-Box Residuals (10)	11.72 <i>0.30</i>	9.94 <i>0.45</i>
Ljung-Box Residuals (20)	21.58 <i>0.36</i>	18.51 <i>0.55</i>
Ljung-Box Squared Residuals (10)	14.12 <i>0.17</i>	10.20 <i>0.42</i>
Ljung-Box Squared Residuals (20)	26.87 <i>0.14</i>	22.68 <i>0.30</i>
Jarque-Bera Test Statistic	250.93 <i>0.00</i>	269.16 <i>0.00</i>

Note : parameter values and their corresponding p-values (in italics)

* indicates p-value at 1% level of significance and ** indicates 5% level

By critically examining the Log-Likelihood estimation and significance of model parameters

The GARCH(1,1) model, when using a normal distribution for residuals, demonstrates superior performance compared to when a student's t-distribution is employed. Parameters : μ , ω , α , β are all significant at 5% significance level. Also, log-likelihood under normal distribution is higher. Ljung-box test's p-values at lags 10 and 20 are all greater than 0.05. It passes the diagnostics test, and we can confidently conclude that the model is well-fitted assuming normally distributed residuals.

From the above table 4, it is observed that all the estimated results for the two distributions are quite distinct therefore the selection of the error distribution in any GARCH model is of utmost importance in accurately modeling the volatility of financial time series data. The choice of distribution has a substantial impact on the model's capacity to accurately represent the data's characteristics, especially with respect to the tails and kurtosis of the distribution. The two most used distributions are the Normal (Gaussian), Student's t. To briefly discuss about the distribution, normal distribution assumes the residuals(errors) are symmetrically distributed around mean with constant variance. The students-t distribution is similar to normal but with heavier tails. It is characterized by degrees-of-freedom parameter (ν). This parameter ' ν ' controls the thickness of the distribution at its tails: lower the degree, thicker is the tail.

For GJR-GARCH(1,1) model except one of the parameters ' α ' is insignificant and Ljung-box test for squared residual at lag 20 is insignificant at 5% level but is significant at 1%.

In case of TARARCH(1,1), both the distributions provide significant parameters but one could argue that ω and α parameters are insignificant but, since they are close to zero which conforms with the previous findings in literature, we can say that based on γ and β alone, models are significant. It passes the diagnostics test as well, suggesting that the TARARCH(1,1) model is successful in capturing threshold effects in volatility. The students-t distribution enhances the model's capability to accommodate heavy tails which can be seen as parameter ν is significant at 5% level with an estimate of 3 degrees-of-freedom. And, evaluating diagnostic test, all the tests favour students-t distribution to be a better fit for modeling conditional variance.

E-GARCH(1,1) with Normal distribution is a better model compared to students-t distribution. We can also see that log-likelihood is negative which indicates that the optimizer did not reach successful convergence in estimating parameters and has given garbage output in this case.

Overall, TARARCH(1,1) has the highest log-likelihood value among other models suggesting it to be the best fitting model. We can verify this based on RMSE values in the next section.

5.2 Evaluating Forecasting Performance:

At first let's only consider only DJIA returns series, we compute out-of-sample forecast using models specified earlier. The following forecast is conducted using all four models with normal distributions of error and we can see from the table below that TARARCH(1,1) model with Fixed length dominates as it has least Relative RMSFE conforming from the earlier section that TARARCH(1,1) model was a better fit.

Table 5: RMSFE and Relative RMSFE with Fixed Limit (FL) and Recursive(RECC) length windows for DJIA

RMSFE	Dow Jones	Relative RMSFE wrt FL_GARCH	
FL_Garch	0.00024		
FL_gjrgarch	0.00022	FL_gjrgarch	0.93
FL_tarch	0.00019	FL_tarch	0.81
FL_egarch	0.00022	FL_egarch	0.94

RMSFE	Dow Jones	Relative RMSFE wrt RECC_GARCH	
Recc_Garch	0.00024		
Recc_gjrgarch	####	Recc_gjrgarch	####
Recc_tarch	0.00022	Recc_tarch	0.92
Recc_egarch	0.00022	Recc_egarch	0.95

Note: #### indicates number is too large thus can ignore that model;
Box filled with grey has the least value and the best model among others.

Conducting the forecasting analysis on all indices using our assumption of normally distributed errors, we get following output for out-sampled fixed limit window forecasting:

Table 6: RMSFE and Relative RMSFE wrt FL_GARCH for all indices

RMSFE	Dow Jones	CAC 40	FTSE 100	Spanish ibex	Nasdaq 100	Italian mibtel	S&P 400	Nikkei 250	Russell 3000	Russell 1000	Russell 2000	Milan MIB 30	German DAX	S & P TSE	S&P 500	USD/GBP	USD/EUR	USD/Swiss	USD/JPY
FL_Garch	0.24	0.27	0.23	0.25	0.25	0.22	0.27	0.41	0.26	0.26	0.35	0.96	0.25	0.25	0.27	0.05	####	0.03	0.07
FL_gjrgarch	0.22	0.31	0.17	0.27	0.26	68.7	0.36	0.42	0.28	0.28	0.44	0.27	0.24	0.21	0.3	####	0.03	0.03	####
FL_tarch	0.19	0.19	0.14	0.18	0.17	0.21	0.17	0.18	0.18	0.18	0.23	0.15	0.2	0.17	0.2	0.07	0.05	0.04	0.07
FL_egarch	0.22	0.26	0.21	0.25	0.23	0.22	0.22	0.32	0.23	0.23	0.29	0.23	0.23	0.25	0.25	0.06	0.04	0.04	0.07

Relative RMSFE wrt FL_GARCH																			
FL_gjrgarch	0.93	1.13	0.76	1.07	1.05	####	1.31	1.01	1.07	1.06	1.26	0.28	0.97	0.87	1.11	####	0.00	0.98	####
FL_tarch	0.81	0.68	0.61	0.70	0.68	0.92	0.61	0.43	0.67	0.68	0.66	0.16	0.78	0.68	0.72	1.46	0.00	1.35	1.13
FL_egarch	0.94	0.95	0.93	0.98	0.94	0.97	0.82	0.77	0.86	0.87	0.84	0.24	0.91	1.02	0.93	1.29	0.00	1.19	1.07

Note: RMSFE values presented in the table have been scaled by a factor of 1000;
indicates number is too large thus can ignore that model;
Box filled with grey has the least value and the best model among others.

And, we get following output for out-sampled recursive window forecasting:

Table 7: RMSFE and Relative RMSFE wrt Recc_GARCH for all indices

RMSFE	Dow Jones	CAC 40	FTSE 100	Spanish ibex	Nasdaq 100	Italian mibtel	S&P 400	Nikkei 250	Russell 3000	Russell 1000	Russell 2000	Milan MIB 30	German DAX	S & P TSE	S&P 500	USD/GBP	USD/EUR	USD/Swiss	USD/JPY
Recc_Garch	0.24	0.28	54.5	0.81	1.32	1.18	0.27	0.41	0.26	0.26	0.35	84.6	0.27	0.24	0.27	7388	####	####	####
Recc_gjrgarch	####	####	####	2096	1483	####	9674	1.21	####	####	4366	2306	####	####	####	####	####	####	####
Recc_tarch	0.22	0.21	146	0.18	0.17	0.15	0.21	0.33	0.2	0.2	0.29	0.16	0.21	0.2	0.22	0.07	0.05	0.04	0.07
Recc_egarch	0.22	0.26	0.21	0.25	0.23	0.21	0.22	0.32	0.22	0.23	0.28	0.23	0.23	0.27	0.25	-	-	-	-

Relative RMSFE wrt RECC_GARCH																			
Recc_gjrgarch	####	####	####	####	####	####	####	2.93	####	####	####	####	####	####	####	####	####	3.19	0.42
Recc_tarch	0.92	0.77	2.68	0.22	0.13	0.12	0.77	0.80	0.78	0.78	0.83	0.00	0.78	0.83	0.82	0.00	0.00	0.00	0.00
Recc_egarch	0.95	0.96	0.00	0.30	0.18	0.18	0.80	0.77	0.86	0.87	0.81	0.00	0.83	1.08	0.94	-	-	-	-

Note: RMSFE values presented in the table have been scaled by a factor of 1000;
indicates number is too large thus can ignore that model;
- indicates model did not converge to optimal value;
Box filled with grey is the least and the best model among others.

From tables 6 and 7, from the values of relative RMSFE of TARCH(1,1) we can say that it performed better than our benchmark model GARCH(1,1). Value less than 1 indicates that GARCH(1,1) is worse than the other, vice versa. So, we can see that 12 out of 15 indices have followed TARCH(1,1) for any of Fixed limit and Recursive window forecast.

Upon examining the Root Mean Square Forecast Error (RMSFE) significant trends become apparent. For prominent US stock indices such as the Nasdaq 100, S&P 500, and Dow Jones, the models often exhibit strong performance, with low RMSFE values. This observation suggests that the predictability of volatility in these markets is somewhat higher, maybe attributed to their abundant liquidity and extensive market data. In contrast, European indices such as the CAC 40, FTSE 100, and German DAX show slightly higher RMSFE values, suggesting that forecasting volatility in these markets is somewhat more challenging. This could be due to regional market dynamics or external influences, like varying economic conditions across European countries. In the currency markets, models like gjr-garch and garch have higher RMSFE values, indicating difficulties in predicting conditional variance. Consistently higher RMSFE values of gjr-garch across indices and pairs suggests that GJR-GARCH(1,1) may be less robust when applied to data with high volatility clustering or structural breaks, which is crucial in high-frequency financial data. For both fixed limit and recursive window forecasting, TARCH(1,1) model consistently produces relative RMSFE values below 1 for most indices, indicating that it performs better than the benchmark GARCH(1,1) model. These results show that the TARCH(1,1) model effectively accounts for important components of market volatility, including asymmetry and leverage effects.

But, low RMSFE suggests that GARCH(1,1) is still better than GJR-GARCH(1,1) or E-GARCH(1,1) models. So it seems that often employing a more sophisticated models, capable of capturing detailed patterns in volatility, does not necessarily translate into superior forecasting ability when contrasted with simpler models.

6. CONCLUSION:

This paper investigates the ability of GARCH-type econometric models to forecast volatility in major world indices and currency pairs. The study aims to bridge the gap between theoretical accuracy and practical evidence for the selected markets and models. Therefore, good volatility model must be able to capture and reflect commonly held stylised facts about conditional volatility. Fixed and Recursive window methods are employed to provide a more comprehensive analysis.

There are five GARCH models being considered: GARCH, GJR-GARCH, TGARCH, and EGARCH. The first model is symmetric, while the remaining three are asymmetric. The evaluation of the forecasts is determined by using RMSFE error statistics.

Empirical analyses have shown that employing the TARCH(1,1) is recommended since it outperformed other models both in estimation and forecasting the volatility of the major indices.

In fact, other than TARCH in our case, we find that even the advanced GARCH models do not provide better out-of-sample forecasts than with the standard GARCH(1,1) model. We recommend that the standard GARCH(1,1) model is good enough, if not better than the more advanced GARCH models for the volatility forecasting of stock indices.

For future research, the key implication is that additional investigation may be required to test the persistence of the parameters that incorporate various stylized facts. For instance, if the estimates for the leverage effect parameter are highly unstable over time, excluding it from the forecasting model might improve the out-of-sample performance of the model. Volatility modeling has undergone significant evolution, progressing from its early theoretical foundations to the era of advanced modern techniques. It evolved for its increasing complexity in financial markets and the ongoing pursuit of improved tools for comprehending and forecasting market. This evolution reflects the growing complexity of financial markets and the continuous search for better tools to understand and predict market behaviour. The development of ARCH and GARCH models marked a significant breakthrough, while subsequent advancements in stochastic volatility, multivariate models, high-frequency data analysis, and machine learning have continued to refine and expand the capabilities of volatility forecasting.

To find a better fitting model than used in this paper, we could utilise multi-horizon forecasting instead of one-step-ahead forecasting or use a different volatility model which allow for volatility breaks, such as the Markov-switching models of conditional heteroscedasticity (see e.g., Lange and Rahbek (2009)).

To gauge the performance of the models in live market VaR is used as one of the metrics. So, accurate volatility forecasting is critical for reliable VaR estimates. Jorion (2000) highlighted the importance of using GARCH-based VaR estimates, showing that they are more reliable than those assuming constant volatility. After Jorion, Brooks and Persaud (2003) compared various VaR models and found that GARCH-based models outperform those assuming constant volatility in capturing tail risk and extreme market movements.

7. REFERENCES:

1. Jorion, P. (2000). *Value at Risk: The New Benchmark for Managing Financial Risk*. 2nd ed. New York: McGraw-Hill.
2. Brooks, C. and Persaud, G. (2003). The effect of asymmetries on stock index return value at risk estimates. *Journal of Risk Finance*, 4(2), pp.29-42.
3. Lange, T. and Rahbek, A. (2009). Markov-Switching GARCH models. *Econometric Theory*, 25(2), pp.1-30.
4. Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), pp.987-1007.
5. Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31(3), pp.307-327.
6. Glosten, L.R., Jagannathan, R., and Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48(5), pp.1779-1801.
7. Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), pp.347-370.
8. Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), pp.579-625.
9. Zakoian, J.M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5), pp.931-955.
10. Bali, T.G. and Demirtas, K.O. (2008). Testing conditional volatility models of stock returns: A reality check based on realized volatility. *Journal of Futures Markets*, 28(5), pp.919-944.
11. Ederington, L.H. and Guan, W. (2005). Forecasting volatility. *Journal of Financial and Quantitative Analysis*, 40(1), pp.85-110.
12. Pagan, A. and Schwert, G.W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, 45(1-2), pp.267-290.
13. Poon, S.H. and Granger, C.W.J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2), pp.478-539.
14. Andersen, T.G. and Benzoni, L. (2008). Realized volatility. In: S.N. Durlauf and L.E. Blume, eds. *The New Palgrave Dictionary of Economics*. 2nd ed. London: Palgrave Macmillan, pp.529-533.
15. Lee, J.H. (1991). Prediction evaluation of GARCH models in stock market volatility. *Journal of Econometrics*, 54(1-2), pp.239-261
16. Brooks, C., and Burke, S.P. (1998). Forecasting exchange rate volatility using conditional variance models selected by information criteria. *Economics Letters*, 61(3), pp.273-278.

17. Engle, R.F., and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1), pp.1-50.
18. Alizadeh, S., Brandt, M.W., and Diebold, F.X. (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3), pp.1047-1091.
19. Brailsford, T.J., and Faff, R.W. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking & Finance*, 20(3), pp.419-438.
20. Corrado, C.J., and Miller, T.W. (2005). The forecast quality of CBOE implied volatility indexes. *Journal of Futures Markets*, 25(4), pp.339-373.
21. Taylor, S.J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton, NJ: Princeton University Press.
22. Todorova, N. (2012). Range-based volatility models and their applications in finance. *Journal of Financial Econometrics*, 10(3), pp.525-550.
23. Merton, R.C. (1973). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science*, 4(1), p.141. [online]. Available from: <https://doi.org/10.2307/3003143>.
24. Hansen, P.R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), pp.873–889. [online]. Available from: <https://doi.org/10.1002/jae.800>.
25. Sharma, P. and Vipul. (2015). Forecasting stock index volatility with GARCH models: international evidence. *Studies in Economics and Finance*, 32(4), pp.445–463. [online]. Available from: <https://doi.org/10.1108/sef-11-2014-0212>.
26. Lin, Z. (2018). Modelling and forecasting the stock market volatility of SSE Composite Index using GARCH models. *Future Generation Computer Systems*, 79, pp.960–972. [online]. Available from: <https://doi.org/10.1016/j.future.2017.08.033>.
27. Costa, F.J.M., 2017. *Forecasting volatility using GARCH models*. Master's thesis, University of Minho. Available at: <https://core.ac.uk/download/pdf/132797589.pdf>