Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/140208

A Comparative Modelling of Essential Characteristics of Volatility

Richard T. A. Samuel[1,*], *Charles Chimedza*[1] *and Caston Sigauke*[2]

[1]School of Statistics and Actuarial Science, University of the Witwatersrand. 1 Jan Smuts Ave, Braamfontein, Johannesburg, 2017, South Africa
[2]Department of Mathematical and Computational Sciences, University of Venda, Private Bag X5050, Thohoyandou 0950, Limpopo, South Africa

Received: 20 Nov. 2024, Revised: 11 Jan. 2025, Accepted: 23 Jan. 2025
Published online: 1 Mar. 2025

Abstract: This study utilised the dynamics of three time-varying models to estimate six essential features of financial return volatility that are relevant for robust risk management. These features include pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Both simulation and empirical evidence are provided. Through the applications of these models using the S&P Indian index, the study shows that the market returns are characterised by these volatility features. The study further used a parametric model through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle estimator, and the exact local Whittle estimator to estimate and determine the presence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns. The results of the estimations indicate that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns when compared with the squared returns and absolute returns. Our findings from the long-memory decomposition revealed that although the response to shocks is greater in the short-term component, it is, however, short-lived. On the contrary, despite a high degree of persistence in the long-term component, market information or unexpected news arrival only has a low long-run impact on the market. Based on this, the long-run investment risks within the Indian stock market seem to be under control. Hence, our findings suggest that rational investors should try to stay calm with the arrival of unexpected news in the market because the long-run effect of such news will not be severe, and the market will eventually return to its normal state.

Keywords: heteroscedasticity, persistence, score-driven model, simulation, time-variation

1 Introduction

Stock market indices are prone to be characterised by features of volatility such as pronounced persistence, mean reversion, leverage effect (i.e., volatility asymmetry), conditional skewness, conditional fat-tailedness, and the long-memory behaviour of volatility decomposition into long-run and short-run components [1]. Hence, these characteristics are usually exhibited by financial returns (see [2, 3]). Volatility persistence is a process where the return of today affects the future's forecast variance [3]. The economic implications of the degree of persistence of a shock include its influence on dynamic hedging policies, the valuation of options and the price of securities [4]. Mean reversion implies that there is a normal volatility level to which volatility will eventually return. Asymmetry is a process in which positive and negative shocks have different impacts on volatility. Asymmetry in equity returns is sometimes known as the leverage effect [3]. The leverage effect implies that volatility tends to be higher following negative returns. Conditional fat-tailedness implies that the standardised conditional return is more fat-tailed than the Normal distribution, while conditional skewness denotes that the standardised return is not symmetric. The long-memory decomposition of volatility occurs when volatility is decomposed into one short-term component and one long-term component [2].

Volatility on its own is not directly observable, hence its measurement and the modelling of its evolution rely on some measurement methods. Two well-documented methods in the literature are the Generalised Autoregressive Conditional

[*] Corresponding author e-mail: richieeniks@gmail.com

Heteroscedasticity (GARCH) model and the Generalised Autoregressive Score (GAS) model. These two models use the conditional variance for measuring volatility (see [5–7]). Moreover, it is also well documented that the Beta-Skew-t-EGARCH model can adequately model essential features of volatility [1, 2] for efficient risk management. Hence, these three models can reliably capture some features of asset returns. However, it is well reported that asymmetric GARCH models perform better at modelling volatility than the simple GARCH model (see [7]), since the simple GARCH model is incapable of capturing asymmetries in the volatility process [8]. Therefore, this study applies a robust extension of the GARCH model called the family GARCH (fGARCH) model [9] that incorporates the asymmetric volatility process. Based on this, the fGARCH, GAS and Beta-Skew-t-EGARCH models are used to model the stated essential features of volatility using the S&P Indian daily equity returns.

To be specific, the study comparatively applies the fGARCH and GAS models to estimate the magnitude and dynamics of the persistence (with mean reversion) in the conditional volatility of the returns. The main difference between these models is that the fGARCH model uses the dynamics of the residuals to drive the conditional variance, while the GAS model uses the dynamics of the conditional score to drive the time-varying conditional variance. Furthermore, the study comparably uses the one and two components of the Beta-Skew-t-EGARCH model to estimate the asymmetry (leverage effect), skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components. In addition, the study applies the parametric ARFIMA-FIGARCH[1] models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle estimator, and the exact local Whittle estimator to determine the existence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns.

Engle and Patton [3] used the GARCH(1,1) model to estimate various stylised facts of volatility, namely, pronounced persistence, mean-reversion and asymmetry in the Dow Jones Industrial Index returns from 1988 to 2000, and found a highly persistence volatility estimate with a half-life of about 73 days. Moreover, the tendency of volatility to persist over time is well documented in the literature (see [6, 10, 11] among others). Oh and Patton [12] used the GAS model with a factor copula model to study systemic risk. Other GAS modelling applications are in credit risk analysis [13], spatial econometrics [14, 15], and high-frequency data [16, 17], among others. See [5, 18] for more dynamic applications of the GAS model. The applications of the Beta-Skew-t-EGARCH model for modelling relevant features of volatility are well studied in Sucarrat [2], and Harvey and Sucarrat [1]. Moreover, the applications of the ARFIMA and FIGARCH models for modelling long memory in various time series data have been widely used by authors like Sowell [19], Cheung [20], and Bollerslev and Wright [21], among others. Also, the applications of the log periodogram frequency domain approach of long memory modelling can be found in the works of Shea [22], Bobeica and Bojesteanu [23] and Vera-Valdés [24], among others.

Our study should result in responses to the following questions. First, which assumed innovations are the most adequate from the fGARCH and GAS modelling to estimate the persistence of the volatility of the returns? Second, how persistent is the market returns volatility? Third, how are the models compared in terms of performance? Fourth, what volatility features characterise the returns and what are the implications on investment in the market? This study used the daily S&P Indian index data, obtained from Thomson Reuters [25] from January 4th, 2010, to June 18th, 2021, with a total of 2990 observations. We chose these periods so as to include the periods of both relative calm and the recent volatility spike (market turmoil) caused by the global COVID-19 pandemic. Moreover, with a vast population, a rapidly growing economy, and admirable investment opportunities, India is one of the greatest investment destinations in the world. It is projected to have the largest population globally before 2030 [26], and its financial markets are increasingly attracting many foreign and domestic investors. The rest of the paper is structured as follows. In Section 2, the applied research theories and methodologies are presented. Section 3 presents the empirical and simulation results, and the discussion of the novel findings, while Section 4 concludes.

2 Materials and Methods

2.1 The GARCH Model

The GARCH model [27] is an extension of the Autoregressive Conditional Heteroscedasticity (ARCH) model proposed by Engle [28] for volatility modelling. It is usually described by its conditional mean and variance equations. The mean equation can be stated as:

$$r_t = \mu_t + \varepsilon_t, \tag{1}$$

[1] ARFIMA-FIGARCH is the Autoregressive Fractionally Integrated Moving Average-Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity.

where r_t denotes the returns, $\varepsilon_t = z_t\sigma_t$ is the random or unpredictable residual. The z_t is the standardized residual returns $(z_t = \varepsilon_t/\sigma_t)$ which are independent, identically distributed (i.i.d.) random variables with zero mean and unit variance. The μ_t in Equation (1) represents the mean function and it is usually expressed as an Autoregressive Moving Average (ARMA(m,n)) process,

$$\mu_t = \sum_{j=1}^{m} \varsigma_j r_{t-j} + \sum_{j=1}^{n} \psi_j \varepsilon_{t-j}, \tag{2}$$

where ς_j and ψ_j are unknown parameters. The conditional variance equation of the general GARCH(p,q) model can be defined as:

$$\sigma_t^2 = \underbrace{\underbrace{\omega + \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2}_{\text{ARCH}} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2}_{\text{GARCH}}, \tag{3}$$

where $\alpha_j \geq 0$ and $\beta_j \geq 0$ are the ARCH and GARCH coefficients, respectively, while $\omega > 0$ is the intercept. The first-order GARCH(1,1) is possibly the best candidate and the most widely used GARCH model for modelling volatility [29]. The rate of decay of shocks to volatility in the conditional variance of the GARCH process is measured by summing the coefficients (α, β). This refers to the volatility persistence of the GARCH models and it indicates the speed of the decay of volatility after a shock. If the sum of the coefficients equals one, then shocks to the volatility do not decrease over time, hence the persistence is felt forever, and the unconditional variance of the process does not exist. Such a situation is called integrated GARCH (IGARCH) [6]. A further extension of the IGARCH process known as the fractional IGARCH (FIGARCH) was introduced by Baillie et al. [30], where the volatility persistence is shorter than an IGARCH but longer than the standard GARCH [31–33]. Shocks to volatility show long persistence into the future when the sum is close to one. This produces a mean-reversion system in which the variance process (volatility), be it high or low, eventually returns very slowly to the mean (normal) state. Lastly, the process of shocks to the conditional variance shows high persistence when the sum is greater than one, which implies explosive volatility forecasts.

2.2 Persistence and Mean Reversion in Volatility

Volatility can be described as persistent if today's return produces a large effect on the prediction variance for many periods in the future [3]. Volatility clustering means that small volatility shocks are followed up by small shocks while large volatility shocks are followed by large shocks in turn. Hence, a period of low volatility will be followed by a volatility rise, while a period of high volatility will sooner or later make way for more normal volatility [3]. Mean reversion in volatility implies that there is a normal volatility level to which volatility will return eventually. A familiar classical measure of volatility persistence is called the "half-life" of volatility, denoted as $h2l$ [34]. This can be described as the number of days it will take the volatility to revert or move halfway back towards its unconditional mean after deviating from it (see [3,34]) and is given by

$$h2l = \frac{\log_e \frac{1}{2}}{\log_e \hat{P}}, \tag{4}$$

where $\log_e$ denotes the natural logarithm, and $\hat{P}$ represents the estimate of the persistence parameter.

2.3 The fGARCH Model

The family GARCH abbreviated fGARCH [9] is an omnibus model that subsumes some familiar asymmetric and symmetric GARCH models as sub-classes [34]. These sub-classes include the standard GARCH (sGARCH) model [27], the Threshold GARCH (TGARCH) model [35], the Nonlinear Asymmetric GARCH (NAGARCH) model [36], the Absolute Value GARCH (AVGARCH) model [37, 38], the Nonlinear ARCH model [39], the Exponential GARCH (EGARCH) model [40], the Glosten-Jagannathan-Runkle GARCH (GJRGARCH) model [41], and the Asymmetric Power ARCH (apARCH) model [42]. The fGARCH(p,q) model can be stated as:

$$\sigma_t^{\gamma} = \omega + \sum_{j=1}^{p} \alpha_j \sigma_{t-j}^{\gamma} (|z_{t-j} - \zeta_{2j}| - \zeta_{1j}\{z_{t-j} - \zeta_{2j}\})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\gamma}. \tag{5}$$

As discussed in Ghalanos [34], Equation (5) is related to the Box-Cox transformation of the conditional standard deviation, where the absolute value function can be transformed by the parameter δ, while the shape is determined by γ. This omnibus model allows the decomposition of the residuals to be driven by different powers for z_t and σ_t in the conditional variance equation. The model also allows for both rotations and shifts in the news impact curve, where the rotation drives large volatility shocks while the shift is the main origin of asymmetry for small volatility shocks. The parameter δ is subject to shifts and rotations through the ζ_{2j} and ζ_{1j}, respectively. The full specification of the family GARCH model can be fitted when $\delta = \gamma$ (see [34]). The volatility persistence of the fGARCH model can be obtained through the parameter estimate $\hat{P}$, stated as:

$$\hat{P} = \sum_{j=1}^{q} \beta_j + \sum_{j=1}^{p} \alpha_j \rho_j, \tag{6}$$

where ρ_j, as stated in Equation (7), denotes the expectation of z_t below the Box-Cox transformation that is associated with the absolute value asymmetry term.

$$\rho_j = \mathrm{E}(|z_{t-j} - \zeta_{2j}| - \zeta_{1j}(z_{t-j} - \zeta_{2j}))^{\delta} = \int_{-\infty}^{\infty} (|z - \zeta_{2j}| - \zeta_{1j}(z - \zeta_{2j}))^{\delta} f(z, 0, 1, \ldots) dz \tag{7}$$

This study used the "persistence()" function from the rugarch package [33, 34] of R software to estimate the persistence. The unconditional variance of the fGARCH model, as related to the persistence, is $\hat{\sigma}^2 = \hat{\omega}/(1-\hat{P})^{2/\gamma}$ [34]. Readers can refer to [9, 33, 34] for more information on the nested models and the fGARCH model.

2.4 The GAS Model

An alternative method to the family GARCH model for modelling volatility can be found in a Score Driven (SD) model known as the Generalised Autoregressive Score (GAS) model, introduced by Harvey [29] and Creal et al. [43] (see [5,44]). The model uses the score of the conditional density function to determine the time variation in the parameters. The score functions are robust to outliers, and the model is quite suitable for modelling skewed or fat-tailed time series data like financial returns [16, 18, 29, 45]. Moreover, like the other observation-driven[2] models, extensions to long memory behaviour, asymmetric, and other time series dynamics are possible. Furthermore, likelihood estimation using the GAS model is simple and direct [5].

2.4.1 Model Specification

Let an $N \times 1$ vector $\boldsymbol{r}_t$ imply the dependent variable of interest, $\boldsymbol{\vartheta}_t$ the vector of time-varying parameter, $\boldsymbol{x}_t$ a vector of exogenous variables (i.e., the covariates), all at time t, and $\boldsymbol{\zeta}$ a vector of time-invariant parameters. Define $\boldsymbol{R}^t = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_t\}$, $\boldsymbol{\Theta}^t = \{\boldsymbol{\vartheta}_0, \boldsymbol{\vartheta}_1, \ldots, \boldsymbol{\vartheta}_t\}$, and $\boldsymbol{X}^t = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t\}$. The information set that is available at time t consists of $\{\boldsymbol{\vartheta}_t, \mathscr{F}_t\}$, where

$$\mathscr{F}_t = \{\boldsymbol{R}^{t-1}, \boldsymbol{\Theta}^{t-1}, \boldsymbol{X}^t\}, \quad \text{for } t = 1, \ldots, n. \tag{8}$$

It is assumed that the generation of $\boldsymbol{r}_t$ is through the observation density [43,46]

$$\boldsymbol{r}_t \sim p(\boldsymbol{r}_t | \boldsymbol{\vartheta}_t, \mathscr{F}_t; \boldsymbol{\zeta}). \tag{9}$$

It is further assumed that the mechanism for updating $\boldsymbol{\vartheta}_t$ (i.e., the time-varying parameter) is given by the autoregressive updating equation:

$$\boldsymbol{\vartheta}_{t+1} = \boldsymbol{\kappa} + \sum_{i=1}^{p} \boldsymbol{A}_i \boldsymbol{s}_{t-i+1} + \sum_{j=1}^{q} \boldsymbol{B}_j \boldsymbol{\vartheta}_{t-j+1}. \tag{10}$$

Equation (10) is presented in Ardia et al. [5] as:

$$\boldsymbol{\vartheta}_{t+1} \equiv \boldsymbol{\kappa} + \mathbf{A}\boldsymbol{s}_t + \mathbf{B}\boldsymbol{\vartheta}_t, \tag{11}$$

[2] Observation-driven models are developed to model large changes (which may occur in the form of jumps or shifts) and distributional asymmetries that often exist in financial time series [18].

where $\boldsymbol{\kappa}$, $\mathbf{A}$ and $\mathbf{B}$ are matrices of coefficients with appropriate dimensions and they are functions of the static parameter $\boldsymbol{\zeta}$, while the scaled score $\boldsymbol{s}_t$ is a suitable function of past data, $\boldsymbol{s}_t = \boldsymbol{s}_t(\boldsymbol{r}_t, \boldsymbol{\vartheta}_t, \mathscr{F}_t; \boldsymbol{\zeta})$ (see [43, 46]). Vector $\boldsymbol{\kappa}$ controls the level of the process $\boldsymbol{\vartheta}_t$, the matrix of coefficients $\mathbf{A}$ controls (or determines) the impact of $\boldsymbol{s}_t$ on $\boldsymbol{\vartheta}_{t+1}$, while matrix $\mathbf{B}$ determines the persistence of the process [5, 47]. In particular, $\boldsymbol{s}_t$ denotes the direction for updating the vector of parameters from $\boldsymbol{\vartheta}_t$ to $\boldsymbol{\vartheta}_{t+1}$, hence, $\mathbf{A}$ can be described as the step of the update. In other words, $\boldsymbol{s}_t$ acts as a steepest ascent algorithm to improve the local fit of the model given the current parameter position (see [5]). To implement the GAS models[3], the $\mathbf{A}$ and $\mathbf{B}$ matrices are constrained to exist as diagonals (see [5, 12]), e.g., for a GAS model with Student's t error distribution, $\mathbf{A} \equiv \mathrm{diag}(a_\mu, a_\sigma, a_\nu)$ and $\mathbf{B} \equiv \mathrm{diag}(b_\mu, b_\sigma, b_\nu)$, where μ, σ and ν are location, scale and shape parameters, respectively. Hence, b_μ refers to the persistence of the conditional mean (location), while b_σ is the persistence of the conditional variance (or scale) (see [5]). This persistence parameter b_σ of the GAS model coincides with the persistence parameters $\alpha + \beta$ of the standard GARCH model[4] of Bollerslev (see [5, 12, 43, 48]).

The GAS approach depends on the observation density in Equation (9) for a given parameter $\boldsymbol{\vartheta}_t$. When an observation r_t is realised, the time-varying $\boldsymbol{\vartheta}_t$ to the next period $t+1$ can be updated using Equation (10) with

$$\boldsymbol{s}_t = \boldsymbol{S}_t \cdot \boldsymbol{\nabla}_t, \qquad \boldsymbol{\nabla}_t = \frac{\partial \ln p\left(\boldsymbol{r}_t | \boldsymbol{\vartheta}_t, \mathscr{F}_t; \boldsymbol{\zeta}\right)}{\partial \boldsymbol{\vartheta}_t}, \qquad \boldsymbol{S}_t = S\left(t, \boldsymbol{\vartheta}_t, \mathscr{F}_t; \boldsymbol{\zeta}\right), \tag{12}$$

where $S(\cdot)$ represents a matrix function, ln denotes the natural logarithm, $\boldsymbol{\nabla}_t$ is the score of Equation (9) evaluated at $\boldsymbol{\vartheta}_t$, and $\boldsymbol{S}_t$ is the scaling matrix (see [5, 18, 43, 46]). Given the dependence of the stated driving mechanism in Equation (10) on the scaled score vector in Equation (12), the GAS model with orders p and q can be defined by Equations (9), (10) and (12). The model can be referred to as GAS (p, q) and the orders p and q are typically taken as $p = q = 1$ (see [43, 46]). However, for details on including more lags in the GAS process, see [43, 48].

As reported by Ardia et al. [5], the authors Creal et al. [43] suggested setting $\boldsymbol{S}_t$ to the inverse of the information matrix ($\mathscr{I}$) to a power $\gamma > 0$ of $\boldsymbol{\vartheta}_t$ to account for the variance of $\boldsymbol{\nabla}_t$. To be precise,

$$\boldsymbol{S}_t = \mathscr{I}_{t|t-1}^{-\gamma}, \qquad \mathscr{I}_{t|t-1} = \mathrm{E}_{t-1}\left[\boldsymbol{\nabla}_t \boldsymbol{\nabla}_t^\top\right], \tag{13}$$

where the expectation E_{t-1} is taken with respect to the conditional distribution of $\boldsymbol{r}_t | \boldsymbol{r}_{1:t-1}$. The parameter γ normally takes value in the set $\{0, \frac{1}{2}, 1\}$. However, other choices of S_t are possible as well (see [48]). When $\gamma = 0$, $\boldsymbol{S}_t = \mathbf{I}$ (identity matrix), which means there is no scaling. If $\gamma = \frac{1}{2}$ ($\gamma = 1$), then the conditional score $\boldsymbol{\nabla}_t$ is pre-multiplied by the square root of (the inverse of) its covariance matrix $\mathscr{I}_t$. However, whatever the choice of γ, $\boldsymbol{s}_t$ is a martingale difference with respect to the distribution of $\boldsymbol{r}_t | \boldsymbol{r}_{1:t-1}$, i.e., $\mathrm{E}_{t-1}[\boldsymbol{s}_t] = 0$ for all t (see [5]). The GAS framework embodies many available observation-driven models in the literature for a suitable choice of the scaling matrix $\boldsymbol{S}_t$ [18, 43]. Readers can refer to [5, 18, 29, 43, 44, 46, 47, 49] for more details on the GAS model.

2.5 Long Memory and Short Memory Processes

The long memory behaviour of a time series describes the correlation pattern of that series at distant lags. A series with long memory usually shows persistent temporal dependence among distant observations [50]. Long memory is used to describe the high-order correlation structure of a time series. A long memory time series autocorrelation function (ACF) decays hyperbolically. Such series exhibit low-frequency spectral distributions. On the other hand, the low-order correlation structure of a series is used to characterise short memory. The existence of long memory implies that the market gradually responds to information over a long period of time. Shocks to volatility usually have long-running effects in a long-memory process, and such persistence is a vital component of derivative pricing, investment portfolios, and risk management [51]. The idea of long memory was introduced by Hurst [52], and other earlier contributions to the study of time series long memory include Mandelbrot [53] and Mandelbrot and Van Ness [54]. These contributors formalised the empirical findings of Hurst using the cumulative river flow data (see [51, 55, 56]).

In both empirical and theoretical studies, finance researchers have focused on long memory (persistence) in asset returns. Stock market researchers do not only investigate the existence of long memory in the returns but also in the return volatility using squared returns or absolute values of returns [51]. The presence of long memory can be determined by measuring the fractional order of integration d of a time series [57]. In other words, the differencing parameter d of

[3] The GAS model is implemented in this study with the use of the R package GAS developed by Ardia et al. [5].

[4] The GAS model with assumed Normal distribution coincides with the standard GARCH(1,1) model of Bollerslev [27] (see [5, 12, 43, 48]). Hence, we investigated this by comparing the estimate of the persistence $\hat{b}_\sigma$ from the GAS model fitted with a time-varying scale parameter, and the estimate $\hat{\alpha}_1 + \hat{\beta}_1$ from the GARCH(1,1) model. Both models were fitted to the real return S&P Indian stock data under the Normal error, and their outcomes yielded $\hat{b}_\sigma \equiv \hat{\alpha}_1 + \hat{\beta}_1 \approx 0.97$.

fractionally differenced models characterises the long-memory identity of a time series. This parameter d can be related to the Hurst exponent (H) as $d = H - 1/2$. The Hurst exponent is a popular measure of long-memory, and it can be calculated by various methodologies that include the classical-rescaled range (R/S) analysis introduced by Hurst [52]. Other methodologies for calculating H include the modified R/S analysis and the Rescaled Variance (V/S) analysis. When $0 < H < 0.5$, the autocovariances are negative at all lags, and the time series process is termed anti-persistent. On the other hand, when $0.5 < H < 1$, the autocovariances are positive at all lags, and the process is termed persistent (see [57]).

The applications of the fractionally differenced models have been extensively used in finance, hydrology, econometrics, economics, telecommunication, and geophysics, among others (see [58–60]). This brought to light the ARFIMA-FIGARCH models, which were developed for long memory modelling in the first and second moments [30, 32, 55, 61]. The ARFIMA model is used for modelling long memory in the return time series (first moment), while the FIGARCH model is used to model long memory in volatility [57].

A covariance stationary stochastic process will exhibit long memory with memory parameter d under the condition that its spectral density function $f(\lambda)$ satisfies:

$$f(\lambda) \sim C\lambda^{-2d} \text{ as } \lambda \to 0^+, \tag{14}$$

where "$\sim$" indicates that the ratio of the right and left-hand sides tends to one at the limit, and C is a finite positive constant. As the process satisfies the necessary and sufficient condition in Equation (14) and $d > 0$, its autocorrelation function decays at a hyperbolic rate (see [51, 55, 58, 61]), i.e.,

$$\rho_\kappa \sim C_\rho \kappa^{2d-1} \text{ as } \kappa \to \infty, \tag{15}$$

where C_ρ is a constant. The nature of the process memory is determined by parameter d. If $d = 0$, the spectral density is bounded at zero, and the process is referred to as short memory. However, the spectral density is unbounded near the origin if $d > 0$, hence, the process shows long memory. Lastly, the process is termed antipersistent and shows negative memory when $d < 0$ because the spectral density is zero at the origin [51].

Generally, estimation of the fractional differencing parameter d to determine the existence of long memory can be carried out through several methods (see [58]). These methods include the parametric approach (e.g., through the FIGARCH model), and semi-parametric approach (e.g., the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator). Other methods used to estimate H (where $d = H - 1/2$) include the graphical techniques (like the aggregated variance method, the classic rescaled adjusted range analysis, i.e., analysis based on R/S statistic, etc.). Compared to the parametric and semi-parametric approaches, the graphical methods are sensitive to short-range autocorrelation and are generally inaccurate in their estimation of d. However, they are useful to heuristically determine the presence of long-range dependence in datasets, and can also be used to obtain a first estimate of d [62]. To avoid the drawback of the graphical methods, this study applies the parametric approach through the FIGARCH model, and the semi-parametric methods via the GPH, LW, and ELW estimators to obtain the estimate of the fractional integration parameter d that reveals the existence of long-memory in the series. The semi-parametric approach is particularly attractive to users because it allows fractional differencing parameter d to be estimated without specifying the entire time series model [50].

2.6 The ARIMA and ARFIMA Processes

Time series that exhibit long memory are referred to as fractionally integrated series, or $I(d)$, where d is a non-integer in the interval $-\frac{1}{2} < d < 1$, excluding 0 [63]. Hence, the ARFIMA(p,d,q) model class is developed with the introduction of non-integer fractional integration parameter d. A long memory (or fractionally integrated) series is neither a nonstationary or unit root ($I[1]$) nor a stationary ($I[0]$) process; it is an $I(d)$ process, where d is a real number [64]. One of the two approaches to estimating d involves the classical time-series approach that requires the full specification of the ARFIMA(p,d,q) model, where the parameters are estimated by maximum likelihood [19]. It has been shown by Dahlhaus [65] and Fox and Taqqu [66] that the maximum likelihood (ML) estimates of the ARFIMA(p,d,q) model are asymptotically unbiased. Dahlhaus [65] has also shown that the ML estimator of d in the general ARFIMA(p,d,q) Gaussian processes is strongly consistent, asymptotically efficient, and asymptotically normally distributed in the Fisher sense [67]. The second approach is the frequency domain technique, where a consistent and asymptotically normal estimate of d can be obtained without the full specification of the ARMA components of the model. The estimators of d in this second approach are mainly regression-based [63].

2.7 ARFIMA Model

Following Granger and Joyeux [61] and Hosking [55], the ARFIMA(p,d,q) process can be stated as:

$$\phi(L)(1-L)^d(y_t-\mu)=\theta(L)\varepsilon_t, \tag{16}$$

where y_t is a time-series process for $t = 1,\ldots,T$, L is the backward shift operator, d is the fractional integration (or long memory) parameter, $\phi(L)=1+\phi_1L+\cdots+\phi_pL^p$ is the autoregressive (AR) polynomial, $\theta(L)=1+\theta_1L+\cdots+\theta_pL^p$ is the moving average (MA) polynomial, and $L^iy_t=y_{(t-i)}$ [20, 62, 68]. The AR and MA coefficients are of orders p and q, respectively, and the roots of $\phi(L)$ and $\theta(L)$ are strictly outside the unit circle to ensure stationarity and invertibility, respectively. The ε_t is white noise [31], and μ is the mean of the process.

The fractional integrating operator $(1-L)^d$ for non-integer values of d is defined by the binomial expansion as:

$$(1-L)^d=\sum_{i=0}^{+\infty}\begin{bmatrix}d\\i\end{bmatrix}(-L)^i=1-dL-\frac{1}{2}d(1-d)L^2-\frac{1}{6}d(1-d)(2-d)L^3-\ldots, \tag{17}$$

or can be re-defined as:

$$(1-L)^d=\sum_{i=0}^{+\infty}\frac{\Gamma(i-d)L^i}{\Gamma(-d)\Gamma(i+1)}, \tag{18}$$

where $\Gamma(\cdot)$ is the gamma function. The fractional integration parameter d is used to describe the long-memory behavior of the process, while the $\phi(L)$ and $\theta(L)$ parameters make up the short-memory parameters and they affect only the short-run dynamics of the process.

The ARFIMA model is a general form of the standard linear ARIMA[5](p,d,q) processes of Box and Jenkins because it allows the degree of integration d to assume non-integer values. This generalisation is a more flexible way of studying time series data, and it enables researchers to simultaneously account for long and short-term dynamics [69]. The standard ARIMA model is obtained when d is arbitrarily restricted to integer values [50]. The ARFIMA class of fractional processes can be used to model data dependence that is weaker than allowed by unit root processes, but stronger than implied by stationary ARMA processes. The ARFIMA model is a parametric long-memory time series process [21].

The ARFIMA(p,d,q) process is covariance stationary for $-0.5<d<0.5$, while $d<1$ indicates mean reversion. The simplest possible model is the fractional white noise model, i.e., the ARFIMA$(0,d,0)$, where $p=0=q$ [31]. When $d=1$, the process reduces to ARIMA, which implies infinite memory (random walk) [57, 68]; and when $d=0$, the process becomes the standard ARMA. If $-0.5<d<0$, the process shows anti-persistence. However, the process becomes stationary long memory when $0<d<0.5$, and it possesses shocks that decay hyperbolically; here, the autocorrelations decay to zero and will not be summable [32]. Moreover, if $0.5\le d<1$, the process is a mean-reverting non-stationary, with finite impulse response weights. Hence, a broad range of low frequency behaviour can be studied and modelled when d is not confined to the integer domain [20]. With the assumption that $-0.5<d<0.5$, and $d\neq 0$, it was shown by Hosking [55] that as $\kappa\to\infty$, the correlation function, $\rho(\cdot)$, of an ARFIMA process is κ^{2d-1}. Hence, as opposed to the faster geometric display of a stationary ARMA process, the autocorrelations of the ARFIMA process hyperbolically decay to zero as $\kappa\to\infty$. The ARFIMA model summarily comprises a short memory ARMA part (that describes the short-term behaviour of a series), and a fractional differencing of the right order that accounts for any long-term persistent property in the series [50].

2.8 The FIGARCH Model

The ARFIMA-FIGARCH models are used for modelling persistence in a time series mean and volatility [32]. As a motivation from the developments of ARFIMA model types for long memory estimation, Baillie et al. [30] introduced the FIGARCH model to capture long memory in the volatility of time series financial returns [23, 33, 57]. The parametric formulation of the FIGARCH model was built on the ARFIMA processes [21, 55, 61]. As opposed to the IGARCH model where shocks persist forever, and the standard GARCH where shocks decay at an exponential rate, the effect of shocks for the FIGARCH model is in between these two models, where shocks decay at a slower hyperbolic rate [31–33].

Given the standard GARCH equation:

$$\sigma_t^2=\omega+\alpha(L)\varepsilon_t^2+\beta(L)\sigma_t^2, \tag{19}$$

[5] ARIMA is the “Autoregressive Integrated Moving Average”.

where L is the backward shift lag operator, for which $\alpha(L)=\sum_{i=1}^{q}\alpha_i(L)^i$ and $\beta(L)=\sum_{j=1}^{p}\beta_j(L)^j$. This is re-arranged to give the ARMA in squares representation:

$$[1-\alpha(L)-\beta(L)]\varepsilon_t^2=\omega+[1-\beta(L)]\nu_t, \tag{20}$$

where $\nu_t=\varepsilon_t^2-\sigma_t^2$ and the left-hand side is condensed as:

$$(1-L)\phi(L)\varepsilon_t^2=\omega+[1-\beta(L)]\nu_t \tag{21}$$

with $\phi(L)=\sum_{i=1}^{m-1}\phi_i(L)^i$ and $m=\max\{p,q\}$. In the fractionally integrated model, $(1-L)$ is substituted by

$$(1-L)^d=\sum_{\kappa=0}^{\infty}\frac{\Gamma(d+1)}{\Gamma(\kappa+1)\Gamma(d-\kappa+1)}L^\kappa=1-\sum_{\kappa=1}^{\infty}\pi_\kappa L^\kappa, \tag{22}$$

where $\pi_i=\prod_{1\le\kappa\le i}\frac{\kappa-1-d}{\kappa}$. The hypergeometric function expansion is normally truncated to some large number, like, 1000. The following representation of the FIGARCH model is obtained after rearranging again as:

$$\begin{aligned}\sigma_t^2&=\omega[1-\beta(L)]^{-1}+\left\{1-[1-\beta(L)^{-1}\phi(L)(1-L)^d]\right\}\varepsilon_t^2\\&=\omega^*+\lambda(L)\varepsilon_t^2\\&=\omega^*+\sum_{j=1}^{\infty}\lambda_i L^i\varepsilon_t^2,\end{aligned} \tag{23}$$

where $\lambda_1=\phi_1-\beta_1+d$ and $\lambda_\kappa=\beta_1\lambda_{\kappa-1}+(\frac{\kappa-1-d}{\kappa}-\phi_1)\pi_{d,\kappa-1}$. Sufficient conditions to ensure the positivity of the conditional variance for the FIGARCH$(1,d,1)$ model are $\omega>0,\beta_1-d\le\phi_1\le\left(\frac{2-d}{2}\right)$ and $d\left(\phi_1-\frac{1-d}{2}\le\beta_1(\phi_1-\beta_1+d)\right)$. Equation (23) is re-written as follows, by setting $\phi(L)\equiv(1-\alpha(L))$:

$$\begin{aligned}\phi(L)(1-L)^d\varepsilon_t^2&=\omega+(1-\beta(L))+(\varepsilon_t^2-\sigma_t^2)\\\phi(L)(1-L)^d\varepsilon_t^2&=\omega-\sigma_t^2+\varepsilon_t^2+\beta(L)\sigma_t^2-\beta(L)\varepsilon_t^2\\\sigma_t^2&=\omega+\varepsilon_t^2+\beta(L)\sigma_t^2-\beta(L)\varepsilon_t^2-\phi(L)(1-L)^d\varepsilon_t^2\\\sigma_t^2&=\omega+\left\{1-\beta(L)-\phi(L)(1-L)^d\right\}\varepsilon_t^2+\beta(L)\sigma_t^2\\\sigma_t^2&=\omega+\left\{1-\beta(L)-(1-\alpha(L))(1-L)^d\right\}\varepsilon_t^2+\beta(L)\sigma_t^2.\end{aligned} \tag{24}$$

The expansion is truncated to 1000 lags and set as $(1-L)^d\varepsilon_t^2=\varepsilon_t^2+\left(\sum_{\kappa=1}^{1000}\pi_\kappa L^\kappa\right)\varepsilon_t^2=\varepsilon_t^2+\overline{\varepsilon}_t^2$, Equation (24) can be re-written as:

$$\begin{aligned}\sigma_t^2&=\omega+\left\{\varepsilon_t^2-\beta(L)\varepsilon_t^2-(1-L)^d\varepsilon_t^2+\alpha(L)(1-L)^d\varepsilon_t^2\right\}+\beta(L)\sigma_t^2\\\sigma_t^2&=\omega+\left\{\varepsilon_t^2-\beta(L)\varepsilon_t^2-(\varepsilon_t^2+\overline{\varepsilon}_t^2)+\alpha(L)(\varepsilon_t^2+\overline{\varepsilon}_t^2)\right\}+\beta(L)\sigma_t^2\\\sigma_t^2&=\omega+\varepsilon_t^2-\beta(L)\varepsilon_t^2-(\varepsilon_t^2+\overline{\varepsilon}_t^2)+\alpha(L)(\varepsilon_t^2+\overline{\varepsilon}_t^2)+\beta(L)\sigma_t^2\\\sigma_t^2&=\omega-\overline{\varepsilon}_t^2-\beta(L)\varepsilon_t^2+\alpha(L)(\varepsilon_t^2+\overline{\varepsilon}_t^2)+\beta(L)\sigma_t^2\\\sigma_t^2&=(\omega-\overline{\varepsilon}_t^2)-\sum_{j=1}^{q}\beta_j\varepsilon_{t-j}^2+\sum_{j=1}^{p}\alpha_j\varepsilon_{t-j}^2+\sum_{j=1}^{p}\alpha_j\overline{\varepsilon}_{t-j}^2+\sum_{j=1}^{q}\beta_j\sigma_{t-j}^2\\\sigma_t^2&=(\omega-\overline{\varepsilon}_t^2)+\sum_{j=1}^{p}\alpha_j(\varepsilon_{t-j}^2+\overline{\varepsilon}_{t-j}^2)+\sum_{j=1}^{q}\beta_j(\sigma_{t-j}^2-\varepsilon_{t-j}^2).\end{aligned} \tag{25}$$

When $d=1$, the FIGARCH is reduced to the IGARCH model, and when $d=0$, to the standard GARCH model. Conrad and Haag [70] provided a general set of sufficient conditions for the FIGARCH(p,d,q) process [33]. The long memory operator in the FIGARCH model is applied to the squared errors, i.e., the memory parameter acts on the squared errors. This contrasts the ARFIMA model that applies the long memory operator to the unconditional mean of y_t [33,57].

2.9 The Log-periodogram Regression

Geweke and Porter-Hudak [56] (GPH, henceforth) proposed a method for estimating the fractional differencing parameter d in the ARFIMA(p,d,q) model (see [71]). Although the long-memory dependence is described entirely by the memory parameter d, the traditional approach to estimating an ARFIMA model requires the specification of the AR and MA polynomials. However, a simpler approach to estimating d without specifying the AR and MA parts was introduced by Geweke and Porter-Hudak through a semiparametric technique [56]. The log-periodogram regression estimate of d is one of the main methods used under the semiparametric approach. The proof of the limiting distribution and consistency of the GPH approach was provided by Robinson [72] in the Gaussian case [21]. The GPH estimator is based on the expression of the spectral density function of the ARFIMA process with frequencies around zero and it estimates only the memory parameter d [68,69]. Through evaluation close to the origin, the semiparametric estimators bypass the need to specify the short-term time series dynamics. Hence, the semiparametric estimators are robust to the dynamics of short run such as observational noise [24].

2.10 Geweke and Porter-Hudak Estimator

Geweke and Porter-Hudak (GPH) [56] introduced a semiparametric estimator of the fractional differencing parameter, d, formulated on a regression of the log spectral density's ordinates on trigonometric function [31]. The estimator utilises the theory of linear filters to state the process $(1-L)^d y_t = \mu_t$, where $\mu_t \sim I(0)$, as

$$f(\omega)_y = |1-e^{-i\omega}|^{-2d} f(\omega)_\mu, \tag{26}$$

where $f(\omega)_\mu$ and $f(\omega)_y$ are the spectral densities of μ_t and y_t, respectively. Hence, Equation (26) can be stated as

$$\begin{aligned} \log\{f(\omega)_y\} &= \{4\sin^2(\omega/2)\}^{-d} + \log\{f(\omega)_\mu\}, \\ \log\{f_y(\omega)_j\} &= \log\{f_\mu(0)\} - d\log\{4\sin^2(\omega_j/2)\} + \log[f_\mu(\omega_j)/f_\mu(0)]. \end{aligned} \tag{27}$$

GPH proposed to estimate d from a regression that is based on Equation (27) with the use of spectral ordinates $\omega_1, \omega_2, \ldots, \omega_m$, from the periodogram of y_t, i.e., $I_y(\omega_j)$. Thus, for $j = 1,2,\ldots,m$, with $m = g(T)$, where $g(T)$ is such that $\lim_{T\to\infty} g(T) = \infty, \lim_{T\to g}\{g(T)/T\} = 0, \lim_{T\to g}\{\log(T)^2/g(T)\} = 0$,

$$\log\{I_y(\omega_j)\} = a + b\log\{4\sin^2(\omega_j/2)\} + \nu_j, \tag{28}$$

where

$$\nu_j = \log[f_\mu(\omega_j)/f_\mu(0)], \tag{29}$$

and on the assumption that ν_j is i.i.d with mean zero and variance $\pi^2/6$. A good estimate of d should be provided by the regression in Equation (28) when μ_t is white noise ε_t [31]. When μ_t is autocorrelated, GPH revealed that Equation (28) holds approximately for frequencies in the proximity of (or near) zero. If the proximity shrinks at a suitable rate with sample size, then a consistent estimator of d should be realized by the GPH procedure [31].

The ordinary least squares (OLS) estimator of d in Equation (28) will produce the limiting distribution in Equation (30), if the number of ordinates m is selected.

$$\frac{(\hat{d}_{GPH} - d)}{\{\text{var}(\hat{d}_{GPH})\}^{1/2}} \Rightarrow N(0,1). \tag{30}$$

The var$(\hat{d}_{GPH})$ is derived from the common regression formula, either by setting it as $\pi^2/6$ or with the use of the regression residual variance. Geweke and Porter-Hudak [56] originally suggested the bandwidth choice of m equals $\sqrt{T}$ [51]. However, it is obvious from the outcome that the GPH estimator is not $\sqrt{T}$ consistent and convergence will occur at a slower rate [31]. Consistency and asymptotic normality of the process were proven by Geweke and Porter-Hudak [56] only for $d < 0$, while Robinson [73] gave a proof of consistency for $0 < d < 0.50$.

While the GPH estimator is potentially robust to nonnormality and its application is simple, its potential attractiveness is reduced by the behaviour $\hat{d}_{GPH}$ where substantial autocorrelation of μ_t is present. Specifically, Agiakloglou et al. [74] showed that it gives "serious bias" [31, 68]. Robinson [72] proved the asymptotical normality of the GPH estimator for

$d \in (-\frac{1}{2}, \frac{1}{2})$. However, Velasco [75] and Kim and Philips [76] later showed its consistency for $d \in (-\frac{1}{2}, 1)$ and that it has an asymptotically normal limit distribution for $d \in (-\frac{1}{2}, \frac{3}{4})$:

$$\hat{d}_{GPH} \sim N\left(d, \frac{\pi^2}{24m}\right). \tag{31}$$

The Geweke and Porter-Hudak [56] estimator is theoretically valid for $0 < d < \frac{1}{2}$. If the number of frequencies m[6] that is included in the regression is restricted in a way that $m = O(T^{4/5})$, then the asymptotic normality is obtained (see [68, 77]).

2.11 Whittle Estimation

Estimation of the long memory parameter d of a time series with fractional integration in the frequency domain based on the maximum likelihood method is built on the approximation to a Gaussian likelihood proposed by Whittle [78]. See also Fox and Taqqu [66]. Unlike the time domain estimator, the frequency domain approach has gained much popularity due to the fact that its maximum-likelihood estimator is invariant to the unknown mean of the process [63, 79].

2.12 The Whittle Likelihood

Consider a sample with T observations that consist of a stationary centered process $y_1, \ldots, y_T$ that are spaced uniformly in the time domain and a sequence of m frequencies (see [63]),

$$\omega_j = \frac{2\pi j}{T} \quad \text{for } j = 1, 2, \ldots, m. \tag{32}$$

These frequencies, with $m \ll T$, denote a set of angular frequencies, where each is a multiple of the fundamental frequency $2\pi/T$, since it relates to a single oscillation with period T. Equation (33) states the discrete Fourier transform of y_t as:

$$\hat{C}(\omega_j) = \frac{1}{\sqrt{2\pi T}} \sum_{\kappa=1}^{T} y_\kappa e^{i\omega_j \kappa}. \tag{33}$$

The Whittle likelihood is based on the fact that the coefficients $\hat{C}(\omega_j)$ of y_t are asymptotically independent Gaussian random variables having mean value zero and the variance is given by the spectral density of the process at that frequency. As a consequence, the likelihood function at frequency ω_j is

$$L_j = \frac{1}{\sqrt{2\pi f_y(\omega_j)}} \exp\left\{-\frac{I(\omega_j)}{2 f_y(\omega_j)}\right\}, \tag{34}$$

where $I(\omega_j)$ is the sample periodogram stated as:

$$I(\omega_j) = |C(\omega_j)|^2, \tag{35}$$

and $f_y(\omega_j)$ refers to the spectral density at ω_j:

$$f_y(\omega_j) = \frac{1}{2\pi} \sum_{\kappa=-\infty}^{\infty} \gamma_\kappa e^{-i\omega_j \kappa}, \tag{36}$$

where γ_κ represents the autocovariance at lag κ.

Estimation of the fractional differencing parameter d through the local Whittle begins when it is known that the behaviour of the spectral density of y_t at low frequencies can be defined through the condition

$$\lim_{\omega \to 0^+} \omega^{2d} f_y(\omega) = G, \tag{37}$$

where G is a positive quantity, whose dependence is on the parameter d. The process y_t possesses finite power when $2d < 1$, and based on that, d is used to measure the long-term duration of the memory of process y_t. The series is covariance

[6] m is the bandwidth parameter

stationary if $d \in [0, \frac{1}{2})$, but the autocorrelations decay more slowly than in the $I(0)$ state. They fade away hyperbolically to zero as opposed to the stationary ARMA process which decays faster at a geometric rate. The process is mean reverting for $d \in [\frac{1}{2}, 1)$, but it is not covariance stationary because there is no long-term effect of an innovation on future values of the process. It is shown by Granger and Joyeux [61] that a process is nonstationary for $d \geq \frac{1}{2}$ because it has infinite variance. The process exhibits antipersistence (intermediate memory), or long-range negative dependence for $d \in (-\frac{1}{2}, 0]$. When $f_y(\omega_j)$ is substituted by its asymptotic approximation $G\omega_j^{-2d}$ from Equation (37), the negative of the log-likelihood at ω_j will satisfy

$$-\log L_j(G,d) = \frac{1}{2}\left\{\log 2\pi + \log G - 2d\log\omega_j + \frac{1}{G}\omega_j^{2d}|\hat{C}(\omega_j)|^2\right\}. \tag{38}$$

The local Whittle estimators of d are found based on this expression [63].

2.13 The Local Whittle Estimator

The analysis of the local Whittle estimator of d by Robinson [72], motivated by the findings of K*ü*nsch [80] and based on the m lowest frequencies $\omega_1, \ldots, \omega_m$, can be obtained through the minimisation of the negative log-likelihood function

$$-\log L(G,d) = \frac{m}{2}\left\{\log 2\pi + \log G - \frac{2d}{m}\sum_{j=1}^{m}\log\omega_j + \frac{1}{G}\frac{1}{m}\sum_{j=1}^{m}\omega^{2d}|\hat{C}(\omega_j)|^2\right\}. \tag{39}$$

If Equation (39) is partially differentiated with respect to G, it can be shown that the optimal value of G for any value of d is[7]

$$\hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}|\hat{C}(\omega_j)|^2. \tag{40}$$

Based on this outcome, it is obvious that the local Whittle estimator of d, $\hat{d}$, minimises

$$R(d) = \log\hat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j, \quad \hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}I(\omega_j). \tag{41}$$

Robinson [81] has shown that the local Whittle estimator $\hat{d}$ obtained through the minimisation of Equation (34) is asymptotically normally distributed for $d \in (\frac{1}{2}, \frac{3}{4})$ and consistent if $d \in (\frac{1}{2}, 1)$ so that

$$\sqrt{m}(\hat{d} - d) \xrightarrow{d} N\left(0, \frac{1}{4}\right). \tag{42}$$

2.14 The Exact Local Whittle Estimator

An improvement to the local Whittle method was introduced by Shimotsu and Phillips [82] (see [24, 63]). The authors defined the exact local Whittle (ELW) estimator as the minimiser of the function

$$R(d) = \log\hat{G}(d) - \frac{2d}{m}\sum_{j=1}^{m}\omega_j, \quad \hat{G}(d) = \frac{1}{m}\sum_{j=1}^{m}\omega_j^{2d}I_{\Delta_y^d}(\omega_j), \tag{43}$$

where $I_{\Delta_y^d}(\omega_\kappa)$ denotes the periodogram of the fractionally differenced series $\Delta_{y_t}^d$ and the difference operator is now described by the binomial expansion

[7] Different authors may have varying expressions of the Fourier coefficients in Equation (33) by a constant factor. However, the value of d (the fractional differencing parameter) is independent of this multiplier, but the choice will affect the value of $\hat{G}$ [63].

$$\Delta_{y_t}^d = y_t - \frac{d}{1!}y_{t-1} + \frac{d(d-1)}{2!}y_{t-2} - \frac{d(d-1)(d-2)}{3!}y_{t-3} + \frac{d(d-1)(d-2)(d-3)}{4!}y_{t-4} \ \cdots$$

The presence of $I_{\Delta_y^d}(\omega_j)$ in the optimisable function demands that a fractional difference of y_t be calculated for every alteration of d. Hence, the ELW estimation is computationally more demanding than the estimation of the simple local Whittle. The consistency and asymptotic normality of this estimator are the same as those of the local Whittle estimator [24, 63].

2.15 The Beta-Skew-t-EGARCH Model

To begin with, the Beta-t-EGARCH as a model without skewness was proposed by Harvey and Chakravarty [83], and Harvey [29]. It is an extension made to the EGARCH model where an equation that depends on the conditional score of the last observation drives the variance or scale [49, 84]. The Beta-t-EGARCH model can be regarded as an unrestricted version of the GAS model of Creal et al. [43] (see [2]). In other words, it is a dynamic model driven by a conditional score, which is a martingale difference. A martingale difference has a constant (or zero) conditional expectation [29]. Due to various implications of conditional skewness on asset pricing, conditional score models were extended to skew distributions [1]. Hence, the Beta-Skew-t-EGARCH model is a skewed version of the Beta-t-EGARCH model, where skewness can be brought in through the method introduced by Fernández and Steel [85] (see [1, 2]).

A number of useful and attractive properties are attributed to this skewed model. As in the Beta-t-EGARCH model where observations that could be seen as outliers for a Normal distribution are down-weighted with the use of the conditional score [1], the Beta-Skew-t-EGARCH model also displays robustness to outliers or jumps. In particular, it performs quite well in modelling key stylised facts of financial returns like fat-tails, leverage effect, conditional skewness and the long-memory behaviour of volatility decomposition into long-term and short-term components. The model has two versions, described below as “one-component and two-component models”.

2.15.1 The One-Component Beta-Skew-t-EGARCH model

The three equations below illustrate the first order Beta-Skew-t-EGARCH model’s martingale difference version [1].

$$R_t = \exp(\lambda_t)z_t = \sigma_t z_t, \quad z_t \sim st(0, \sigma_z^2, \nu, \eta), \ \ \nu > 2, \ \ \eta \in (0, \infty), \tag{44}$$

$$\lambda_t = \omega + \lambda_t^\dagger, \tag{45}$$

$$\lambda_t^\dagger = \phi_1 \lambda_{t-1}^\dagger + \kappa_1 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1}+1), \quad |\phi_1| < 1, \tag{46}$$

where R_t[8] is the demeaned return, σ_t represents the volatility or conditional scale, and the conditional error z_t is not standardised i.e., the variance is not one. Equation (44) shows that z_t follows the skew Student’s t distribution with a mean of zero, scale σ_z^2, degrees of freedom ν, and skewness parameter η. The z_t is defined as $z_t = z_t^* - \mu_{z^*}$, where z_t^* is an uncentred[9] skew Student’s t variable with ν degrees of freedom, mean μ_{z^*}, and skewness parameter η. The ω denotes the log-scale intercept that is described as the long-term log-volatility, while ϕ_1 represents the persistence parameter (the bigger the persistence, the more the clustering of volatility). The κ_1 is the ARCH parameter that indicates the response to shocks (the bigger the absolute value of the parameter, the greater the response of volatility to shocks), κ^* is the leverage parameter, u_t denotes the conditional score, that is, the derivative of the log-likelihood of R_t at t with respect to λ_t, and *sgn* is the sign function. Details on this model are discussed in Sucarrat [2].

[8] The demeaned return R_t is the same as the error term ε_t (i.e., the unpredictable part of return) in Equation (1), but with z_t distributed as a skew Student’s t, and it is not standardised to have unit variance (see [2]).
[9] That is, the mean is not necessarily equal to zero.

2.15.2 The Two-Component Beta-Skew-t-EGARCH Model

In order to accommodate the long memory feature of financial return like a two-component GARCH model of Engle and Lee [86], the two-component Beta-Skew-t-EGARCH model was introduced. The model is used to decompose volatility persistence into long-term (or long-run) and short-term (or short-run) components [87]. The first order two-component Beta-Skew-t-EGARCH model's martingale difference version [1] is given as:

$$R_t = \exp(\lambda_t)z_t = \sigma_t z_t, \quad z_t \sim st(0,\sigma_z^2,\nu,\eta), \quad \nu,\eta \in (0,\infty), \tag{47}$$

$$\lambda_t = \omega + \lambda_{1,t}^{\dagger} + \lambda_{2,t}^{\dagger}, \tag{48}$$

$$\lambda_{1,t}^{\dagger} = \phi_1 \lambda_{1,t-1}^{\dagger} + \kappa_1 u_{t-1}, \quad |\phi_1| < 1, \tag{49}$$

$$\lambda_{2,t}^{\dagger} = \phi_2 \lambda_{2,t-1}^{\dagger} + \kappa_2 u_{t-1} + \kappa^* sgn(-R_{t-1})(u_{t-1}+1), \quad |\phi_2| < 1, \quad \phi_1 \neq \phi_2, \tag{50}$$

where $\lambda_{1,t}$ and $\lambda_{2,t}$ represent the time-varying long-run and short-run components of log-volatility, respectively. Both components are driven by the conditional score u_t. The ϕ_1 and ϕ_2 are the long-term and short-term persistence parameters and the model is not identifiable if $\phi_1 = \phi_2$. The κ_1 and κ_2 are parameters that indicate the long-run and short-run responses to shocks, respectively.

2.15.3 The Skew Student's t Distribution

Let an ordinary (i.e., symmetric and centred) Student's t distributed variable with unit scale be denoted by ε^*, and let its density be denoted by $f(\varepsilon^*)$. The density of an uncentred skew Student's t variable can be stated through the skewing technique of Fernández and Steel [85] as:

$$f(z^*|\eta) = \frac{2}{\eta+\eta^{-1}} f\left(\frac{z^*}{\eta^{sgn(z^*)}}\right). \tag{51}$$

The right skewness is produced when $\eta > 1$, and the left skewness is attained when $\eta < 1$, whereas symmetry can be attained when $\eta = 1$ [1,2]. That is, the right (left) hand tail is heavier when $\eta > 1$ ($\eta < 1$).

2.16 The True Parameter Recovery Measure

Monte Carlo simulation (MCS) studies largely focus on the estimator's ability to recover the true data-generating parameter (see [88]). Hence, the True Parameter Recovery (TPR) measure was introduced by Samuel et al. [89] as a way of measuring how the MCS estimator performs at recovering the true parameter. The measure is used as a proxy for the coverage[10] of the MCS experiment to calculate the level of recovery of the true parameter by the MCS estimator. It can be stated (see [89]) as:

$$\text{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K\right]\right)\%, \tag{52}$$

where ϑ is the true data-generating parameter, $\hat{\vartheta}$ denotes the estimator, and $K = 0, 1, 2, \ldots, 100$ represents the nominal recovery level. A TPR outcome of 90% or 95%, for instance, implies that the MCS estimator is able to recover 90% or 95% of the true parameter. The MCS estimator $\hat{\vartheta}$ will fully recover the true parameter $\vartheta > 0$ when $\hat{\vartheta} = \vartheta$, such that the outcome of the TPR is equal to the specified nominal recovery level K (that is, $K\%$ = TPR) (see [89]).

3 Experimental Results and Discussion: Empirical and Simulation

This section presents the outcomes of the estimations through the applications of the fGARCH, GAS and Beta-Skew-t-EGARCH models. The three applied models can be estimated through the Maximum Likelihood Estimation (MLE) once a distribution for the innovations is specified.

[10] The probability that the true parameter is contained within a confidence interval of estimates is known as the coverage probability [90].

3.1 Application of the fGARCH Model

In this section, we present the outcomes of the estimations involving the fGARCH model. It is believed that observation-driven models like the (family) GARCH can yield efficient outcomes when fitted with the true (or appropriate) innovation distribution [8,91]. Hence, among ten selected innovations, we use this study to determine an optimal or the most suitable assumed innovation distribution that is relevant for volatility persistence estimation through the fGARCH model. The ten selected assumed innovations are the Gaussian (or Normal), Student's t, Generalised Error Distribution (GED), skew-Normal, skew-Student's t, skew-GED, Johnson's reparametrised SU (JSU) distribution, Generalised Hyperbolic (GHYP) distribution, Generalised Hyperbolic Skew-Student's t (GHST) distribution, and Normal Inverse Gaussian (NIG). See [34,92,93] for details on the assumed innovations.

The estimation is carried out using the actual return data from the S&P Indian stock index. To transform the daily closing price data to the log returns, we take the log-difference of the value of the index as:

$$r_t = \ln\left(\frac{P_{t+1}}{P_t}\right) \times 100, \tag{53}$$

where P_t is the daily closing equity price at time t, ln the natural logarithm and r_t is the current returns.

3.1.1 Exploratory Data Analysis

We begin with the visual inspections of the price level and the index returns over the sample period using exploratory data analysis (EDA) as revealed in Panels ($A-F$) of Figure 1. The EDA gives relevant insights into the dataset to disclose vital information such as detecting possible outliers. The EDA shows that the market is characterised by time-varying volatility, with a steep plunge in volatility of price and return (see Panels A and B) in 2020 due to the emergence of the global COVID-19 pandemic crisis. Volatility clustering is apparent in the return series, where large (small) changes tend to follow large (small) changes of either sign. The quantile-quantile (QQ) plot in Panel C shows that the returns are non-Normally distributed, while the density plot in Panel D denotes stability (stationarity) in the return series. The return series plot in Panel B also reveals stationarity in the returns. Panel E displays the correlogram of the returns, showing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The plots indicate weak dependence in the mean of the series. Hence, we assume a constant conditional mean (see [3]). As for the squared returns, the correlogram as displayed in Panel F shows some moderate dependence at the initial lags, hence we fit an ARMA model as presented in Section 3.1.4 to capture this.

3.1.2 Descriptive Statistics

Panel A of Table 1 presents some descriptive statistics on the return data. The table shows that the mean of the daily return is close to zero. The daily variance is 1.6383, which implies an average annualised volatility of 20.32% [11]. The skewness coefficient shows that the distribution of returns is significantly negatively skewed, a familiar feature of stock returns. This outcome reveals the impact of information arrival in the market, and it shows that investors and other market participants tend to react more to bad news than they do to good news. The kurtosis coefficient is very high. Kurtosis is a measure of the thickness of the tails of the return distribution, and the result shows evidence of leptokurtosis with a value greater than three. That is, it surpasses the kurtosis of the Normal distribution, which is three, and that suggests a fat-tailed distribution for describing the return series. Lastly, the outcome of the Shapiro-Wilk test rejects the assumption of normality with a p-value = 0. This implies that the assumption of a Normal distribution for modelling the volatility persistence of the returns is not realistic.

3.1.3 Tests for Stationarity

This study further used three stationarity tests involving the Augmented-Dickey Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) tests to test for stationarity in the price and return series. We begin with the test for stationarity of the price index in Panel A of Table 2. The table displays the outcomes of the ADF and PP unit root tests, and they indicate that the null hypothesis of non-stationarity, which implies the presence of a unit root in the price series, cannot be rejected because the p-values are greater than 0.05. Hence, the price index is non-stationary.

[11] Annualised volatility for 252 trading days (i.e., one year) $= \sqrt{252} \times \sqrt{\text{variance}} = 20.32\%$ (see [3]).

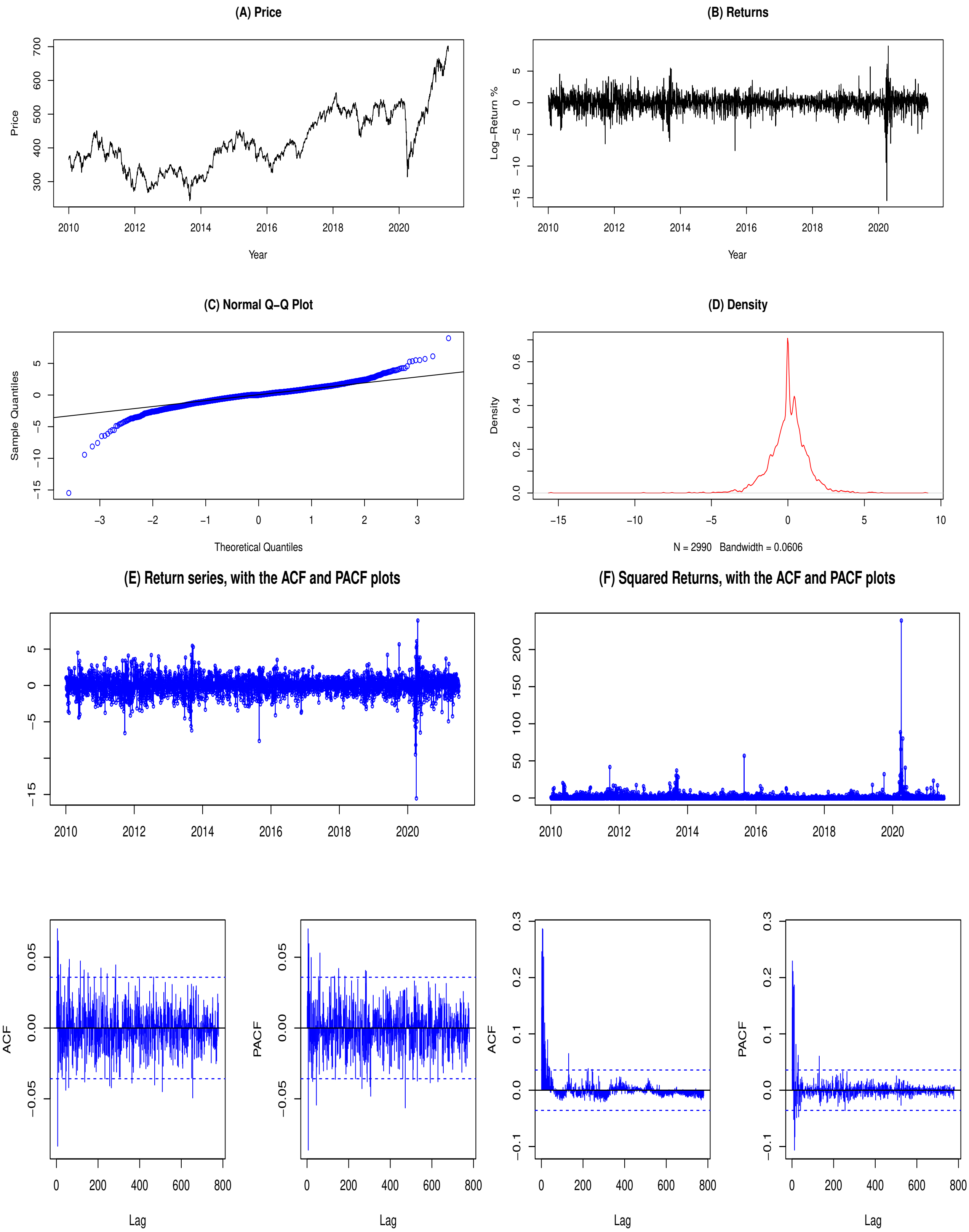

Fig. 1: Panels (A) and (B), respectively, show the plots of the price and return series, while Panels (C) and (D) display the quantile-quantile (Q-Q) and density plots, respectively, for the S&P Indian Index. Panels (E) and (F) display the ACF and PACF plots of the returns and squared returns, respectively.

Table 1: Descriptive Statistics and Engle's ARCH Test Outcomes.

Panel A Returns Descriptive Statistics			**Panel B** Engle's ARCH Test Outcomes				
				PQ test		LM test	
			Lag order	PQ	P-value	LM	P-value
Mean	0.0215		4	452	0	3446	0
Variance	1.6383		8	1043	0	1130	0
Skewness	-0.9525		12	1361	0	641	0
Kurtosis	11.6371		16	1434	0	443	0
Shapiro-Wilk	0.9222		20	1487	0	343	0
p-value	0.0000		24	1502	0	283	0

Note: LM is the Lagrange-Multiplier statistic, while PQ is the Portmanteau-Q statistic in Panel B.

Furthermore, the KPSS test that is based on the null hypothesis of stationarity is rejected because the p-value < 0.05 (see Panel A of Table 2). This further indicates that the price index is non-stationary.

Next, for the return series in Panel B of Table 2, the results of the ADF and PP unit root tests indicate that the null hypothesis of non-stationarity, which implies the presence of a unit root in the returns is rejected because the p-values are less than 0.05. Hence, the return series is stationary. Also, we cannot reject the KPSS test with a null hypothesis of stationarity because the p-value > 0.05, as presented in Panel B of Table 2. This further implies that the return series is stationary.

Table 2: Tests for Stationarity of the Price Index and Return Series.

Panel A Price Index				**Panel B** Returns		
Estimate	Lag order	P-value		Estimate	Lag order	P-value
ADF Test				ADF test		
-2.4231	9	0.3992		-15.6340	9	0.0100
PP Test				PP Test		
-10.9580	9	0.4986		-3072.1	9	0.0100
KPSS Test				KPSS Test		
19.6090	9	0.0100		0.1052	9	0.1000

Note: Panel A comprises of the outcomes of the ADF and PP unit root tests, and the outcomes of the KPSS stationarity test for the price index. Panel B contains the outcomes of the three tests for the returns.

3.1.4 Tests for Autocorrelation and ARCH Effects

Next, autocorrelation and ARCH effects (or heteroscedasticity) are removed by fitting ARMA-fGARCH models with each of the ten innovations to the stationary returns. ARMA(1,1) model as shown in Equation (54) is used as the best, among the candidates' ARMA(m,n) models, to capture autocorrelation in the residuals. Table 3 shows the results of the Weighted Ljung-Box test [94] on both the standardised residuals (denoted as WLB SRs) and the standardised squared residuals (WLB SSRs) from the fit of the ARMA(1,1) model. The table shows that for both instances, the p-values of the test at lag order 5 are greater than 5% under the ten innovation distributions. Based on this, the null hypothesis of "no serial correlation" in the returns cannot be rejected. Hence, serial correlation is removed in the residuals.

$$r_t = \varsigma_0 + \varsigma_1 r_{t-1} + \psi_1 \varepsilon_{t-1} + \varepsilon_t \tag{54}$$

After removing serial correlation in the returns, we carry out Engle's ARCH test [28] using the Portmanteau-Q (PQ) and Lagrange Multiplier (LM) tests to examine the existence of ARCH effects or heteroscedasticity in the residuals. Both tests are carried out based on the null hypothesis of homoscedasticity in the residuals of an ARIMA model. The results from the two tests reveal highly significant p-values of 0 from lag order 4 to 24 as shown in Panel B of Table 1. Therefore, we reject the null hypothesis of "no ARCH effects" in the residuals, which implies the presence of volatility clustering in the returns. Based on this, we fit candidates' fGARCH(p,q) models, with each of the ten distributed errors, to capture the ARCH effects in the returns. Hence, following Engle and Patton [3], we use the Bayesian Information Criterion (BIC) [95] and found that the best candidate model, to capture heteroscedasticity in the returns, among the fGARCH(p,q) class for $p \in [1,5]$ and $q \in [1,2]$ is the parsimonious fGARCH(1,1) (see Equation 55).

$$\sigma_t^{\gamma} = \omega + \alpha_1 \sigma_{t-1}^{\gamma} (|z_{t-1} - \zeta_{21}| - \zeta_{11}\{z_{t-1} - \zeta_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma}. \tag{55}$$

3.1.5 Residual Diagnostic Test

Following the fit of the fGARCH(1,1) model to the returns, we carry out some diagnostic checks using the weighted ARCH LM test to ascertain if heteroscedasticity has been removed. Based on the diagnostic results, it can be seen from Table 3 that the p-value of the "ARCH LM (5)" statistic at lag order 5 exceeds 5% under each of the ten assumed errors. Hence, this implies that the ARCH effect is captured since we cannot reject the null hypothesis of "no ARCH effects" in the residuals.

3.1.6 Selection of an Optimal Error Distribution

Next, the selection of the most adequate innovation distribution (among the ten innovation assumptions) that can be used to describe the returns for estimating the volatility persistence is carried out in Table 3. Models' comparison and selection are made using four information criteria that consist of the Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Shibata Information Criterion (SIC), and Hannan-Quinn Information Criterion (HQIC) (see [34, 95]). The assumed innovation with the minimum (or lowest) value of information criteria will be the most adequate innovation distribution required to estimate the persistence. To begin with, Table 3 shows that all, but one, of the fGARCH volatility parameter estimates ($\hat{\omega}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, $\hat{\zeta}_{11}$, and $\hat{\zeta}_{21}$) under the ten assumed errors are highly significant at 1% level. To be precise, the only exclusions to the 1% significance are the $\hat{\omega}$ which is significant at 5% under the Student's t, and the $\hat{\zeta}_{11}$ which is nearly insignificant across the board. These highly significant outcomes in the parameters reveal the presence of volatility clustering in the conditional variance. Moreover, the strongly significant $\hat{\zeta}_{21}$ indicates that the market is driven more by short volatility shocks. In other words, the effects of short volatility shocks are more pronounced in the market than those of large shocks.

For model selection, Table 3 shows that the four information criteria have their lowest values under the NIG innovation distribution. Therefore, the NIG innovation (fitted with the fGARCH(1,1) model) is the most suitable to describe the market returns when the underlying error distribution is unknown for volatility modelling. This result is evident in the claim that the NIG distribution is analytically tractable, and it can adequately model the skewness of financial market variables, such as equity prices, exchange rates, and interest rates [96]. Moreover, it has been applied many times for financial applications both as the unconditional return distribution and as the conditional distribution of a GARCH model (see [96]). Hence, this makes the distribution attractive for financial modelling applications.

3.1.7 Persistence in Volatility and Mean Reversion

The estimated volatility persistence under this optimal NIG error distribution is 0.9749. This indicates that the volatility of the S&P Indian equity market's returns displays considerably high persistence. To determine if a model has adequately captured all the persistence present in the variance of returns, Engle and Patton [3] suggested that the standardised squared residuals (SSRs, henceforth) should be serially uncorrelated. Hence, from our outcomes in Table 3, the p-value of the Weighted Ljung-Box test of the SSRs is greater than 5% under the NIG innovation (and generally under all the remaining nine assumed innovation distributions)[12], implying that the SSRs are serially uncorrelated. This indicates that the fGARCH(1,1) model fitted with the NIG assumed innovation has adequately captured all of the persistence present in the variance of returns.

Broadly speaking, the outcomes of our study show that the mean and variance equations are well specified, with no evidence of correlation in the standardised residuals (SRs) and SSRs of the model. The outcomes of the calculated ARCH LM test also show that the ARCH effect is filtered out in the residuals since we cannot reject the null hypothesis of "no ARCH effects" based on the ARCH LM results. Thus, the estimated fGARCH(1,1) model in this study is adequately specified and correctly fitted, and the persistence is also adequately captured by the model. Figure 2 shows the conditional volatility of the fGARCH(1,1) model fitted with the NIG assumed distribution, and it displays alternating phases of lower and higher volatility in the returns.

[12] These outcomes are observed up to the ninth lags for both SRs and SSRs under the ten error assumptions.

Table 3: The Empirical Outcomes of fitting ARMA(1,1)-fGARCH(1,1) models to the return data.

	(A) Normal	(B) skew-Normal	(C) Student's t	(D) skew-Student's t	(E) GED
$\hat{\varsigma}_0$	0.0045	-0.0073	0.0462**	0.0110	0.0389**
$\hat{\varsigma}_1$	0.5358*	0.5260**	0.3417	0.3545	0.2327*
$\hat{\psi}_1$	-0.4668**	-0.4612***	-0.2724	-0.2872	-0.1767*
$\hat{\omega}$	0.0482*	0.0453*	0.0441**	0.0447*	0.0460*
$\hat{\alpha}$	0.0769*	0.0844*	0.0871*	0.0903*	0.0812*
$\hat{\beta}$	0.8046*	0.8168*	0.8054*	0.8075*	0.8093*
$\hat{\zeta}_{11}$	-0.1721	-0.1842***	-0.1682	-0.1893	-0.1725
$\hat{\zeta}_{21}$	1.3867*	1.3282*	1.3281*	1.3301*	1.3288*
$\hat{\gamma}=\hat{\delta}$	2.1253*	1.9995*	1.9723*	2.0131*	2.0521*
Persistence ($\hat{P}$)	0.9726	0.9735	0.9728	0.9760	0.9711
WLB SRs (5)	0.7405	0.6924	0.7014	0.8040	1.5361
p-value (5)	(1.0000)	(1.0000)	(1.0000)	(1.0000)	(0.9973)
WLB SSRs (5)	2.927	3.039	2.761	3.094	3.034
p-value (5)	(0.4205)	(0.4000)	(0.4524)	(0.3902)	(0.4009)
ARCH LM (5)	3.9170	3.8687	3.4943	3.5942	3.7311
p-value (5)	(0.1816)	(0.1862)	(0.2257)	(0.2144)	(0.1999)
AIC	3.0541	3.0426	3.0085	3.0006	3.0052
BIC	3.0722	3.0627	3.0286	3.0227	3.0253
SIC	3.0541	3.0426	3.0085	3.0006	3.0052
HQIC	3.0606	3.0499	3.0157	3.0086	3.0124
	(F) skew-GED	**(G) GHYP**	**(H) NIG**	**(I) GHST**	**(J) JSU**
$\hat{\varsigma}_0$	0.0064	0.0108	**0.0106**	-0.0027	0.0097
$\hat{\varsigma}_1$	0.2684**	0.3145	**0.3208**	0.3712	0.3312
$\hat{\psi}_1$	-0.2017***	-0.2501	**−0.2556**	-0.3052	-0.2651
$\hat{\omega}$	0.0458*	0.0447*	**0.0443***	0.0434*	0.0442*
$\hat{\alpha}$	0.0831*	0.0878*	**0.0900***	0.0963*	0.0916*
$\hat{\beta}$	0.8116*	0.8104*	**0.8113***	0.8187*	0.8112*
$\hat{\zeta}_{11}$	-0.1796	-0.1910	**−0.1919**	-0.1998***	-0.1918
$\hat{\zeta}_{21}$	1.3298*	1.3263*	**1.3191***	1.2903*	1.3151*
$\hat{\gamma}=\hat{\delta}$	2.0496*	2.0335*	**1.9993***	1.8975*	1.9821*
Persistence ($\hat{P}$)	0.9735	0.9748	**0.9749**	0.9756	0.9754
WLB SRs (5)	1.1550	1.0025	**0.9421**	0.8057	0.8857
p-value (5)	(0.9999)	(1.0000)	(**1.0000**)	(1.0000)	(1.0000)
WLB SSRs (5)	3.146	3.171	**3.127**	3.029	3.100
p-value (5)	(0.381)	(0.3767)	(**0.3843**)	(0.4017)	(0.3891)
ARCH LM (5)	3.8738	3.6558	**3.6329**	3.7242	3.6114
p-value (5)	(0.1857)	(0.2078)	(**0.2102**)	(0.2006)	(0.2125)
AIC	3.0008	2.9995	**2.9990**	3.0030	2.9992
BIC	3.0229	3.0236	**3.0210**	3.0251	3.0213
SIC	3.0008	2.9994	**2.9989**	3.0030	2.9992
HQIC	3.0087	3.0081	**3.0069**	3.0110	3.0072

Note: WLB SRs (SSRs) denote the Weighted Ljung-Box test for standardised residuals (standardised squared residuals), where "(5)" is lag order 5. The p-value at 5% level of significance is presented in the round bracket for each error. The "$*$", "$**$" and "$***$" are 1%, 5%, and 10% significance levels, respectively. The information criteria are computed as AIC = $-2L/N+2p/N$, BIC = $-2L/N+p\log_e(N)/N$, HQIC = $-2L/N+2p\log_e(\log_e(N))/N$, and SIC = $-2L/N+\log_e((N+2p)/N)$, where N denotes the sample size, L is the log-likelihood of the maximum likelihood of unknown parameter vector $L(\Theta)$, and p denotes the number of estimated parameters [34, 95].

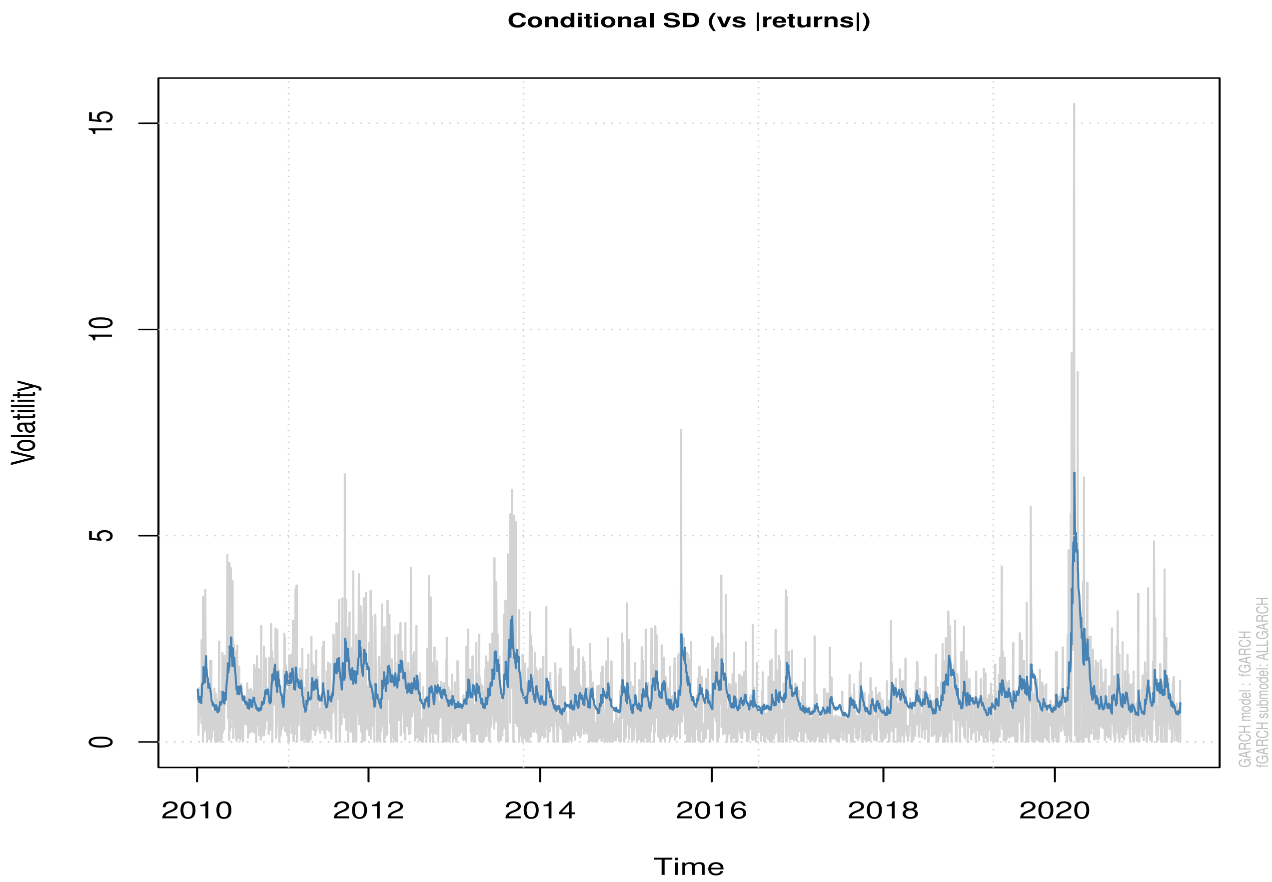

Fig. 2: Conditional volatility of the fGARCH(1,1) model fitted with the NIG assumed error.

The estimated outcome of 0.9749 is reasonably close to one, which is a sign of considerable persistence in the volatility. The existence of long memory in financial assets' volatility relates with the persistence of shocks in returns [23]. However, a crucial test to determine the presence of a long memory feature is to examine it in the return and volatility of the series [32]. Hence, this study used the semiparametric and parametric approaches to estimate the fractional differencing parameter d to determine the existence of long memory in the returns and the return volatility, i.e., squared returns and absolute values of returns [51]. Following the steps of Bobeica and Bojesteanu [23], we used a parametric approach through the FIGARCH model, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak [56] (GPH, henceforth) and Robinson [72], the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator. The long memory outputs are obtained by calculating the fractional differencing parameter d in the returns and volatility using these estimators. We decided to use the LW and ELW methods as improvements to the GPH technique due to the acclaimed bias in the estimation of the GPH estimator (see [31, 68, 74]).

The parametric FIGARCH model is applied with fractional white noise ARFIMA$(0,d,0)$ specification [20]. For the semiparametric approach, we follow Geweke and Porter-Hudak [56] to use the sample selection function $m = g(T) = T^{\alpha}$, $0 < \alpha < 1$ (see [22]). Hence, we used the optimal rate $m = T^{4/5}$ for the number of frequencies, with a power equal to 0.8 (see [68, 77, 97, 98]). This is also the default bandwidth applied in Vera-Valdés [24], where m is the bandwidth that specifies the number of Fourier frequencies and T is the sample size (see [98]).

The outcomes of the estimation of d in the ARFIMA process by the GPH, LW, ELW, and FIGARCH methods for the returns, absolute returns, and squared returns are reported in Table 4. The three semiparametric estimators yield closely related results for d within 0 and 0.5, indicating the presence of long memory in the series. The outcomes of d through the parametric ARFIMA-FIGARCH method are also within 0 and 1, indicating reasonable evidence of long memory in the conditional variance of the series. These results suggest that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns compared to the squared and absolute returns.

Table 4: Long memory estimation through the fractional integration parameter d.

	Returns	Absolute returns	Squared returns
Panel A: Semiparametric estimation methods			
GPH estimator	0.0389	0.2828	0.2749
LW estimator	0.0347	0.2611	0.2461
ELW estimator	0.0504	0.2779	0.2628
Panel B: Parametric estimation method			
ARFIMA-FIGARCH process	–	0.1178	0.1278

Note: Estimated outcomes of the fractional integration parameter d for the returns, absolute returns, and squared returns (conditional volatility) using the semiparametric (in Panel A) and parametric (in Panel B) approaches.

The obtained outcome of 0.9749 indicates that the market's returns volatility displays considerable persistence, with a volatility half-life of about 27 days. This suggests the presence of considerable long memory in the return's volatility, but it is still mean reverting since the outcome is significantly less than one. This implies that even if it takes a while, the volatility process does go back to its mean (see [3]). With this decay rate of 0.9749, we plot the decay process using the $(0.9749)^{\text{day}}$ following the method of Chen and Shen [99], Chiang et al. [100], and Chou [6] for the first 100 days as shown in Figure 3. Panel A of the plots displays the decay of shock impacts through the fGARCH(1,1)-NIG model, and it shows that the impact decays slowly (see [3]), dropping to half the intensity in about 27 days. The limit of the decay sequence is zero [3] as shown by the decline towards zero. This confirms that the volatility process is mean reverting. The unconditional variance of the fGARCH(1,1)-NIG model for the S&P Indian market over the sample period is 1.7659, implying a mean annualised volatility of 21.10%, which is quite close to the sample estimate of the unconditional volatility presented in Section 3.1.2.

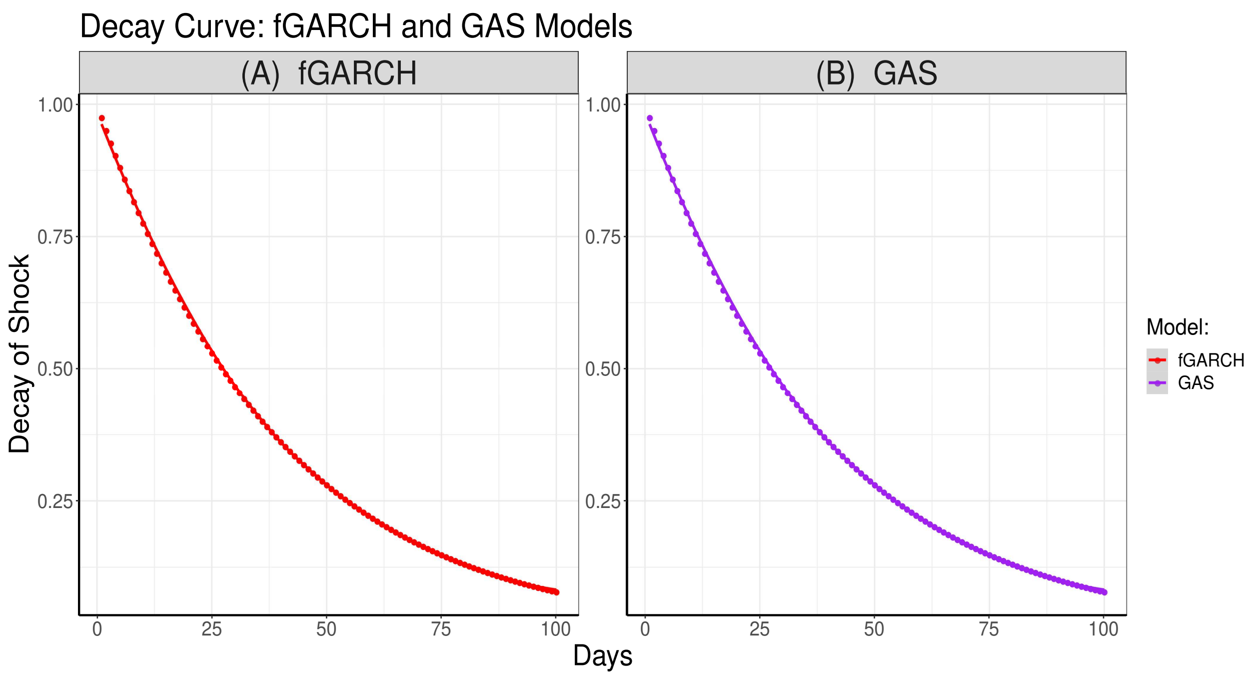

Fig. 3: Panels (A) and (B) display the decay of volatility persistence through the fGARCH(1,1)-NIG and GAS-AST1 models, respectively. The two decay curves follow the same trajectory because the persistence estimates from the two models are the same (i.e., 0.9749).

From the outcomes of the ARMA(1,1)-fGARCH(1,1) models in Table 3, the mean and variance equations of the fGARCH(1,1) model when fitted with the assumed NIG innovation can be stated as:

$$
\begin{aligned}
r_t &= \mu_t + \varepsilon_t \\
&= \varsigma_0 + \varsigma_1 r_{t-1} + \psi_1 \varepsilon_{t-1} + \varepsilon_t \\
&= 0.0106 + 0.3208 r_{t-1} - 0.2556 \varepsilon_{t-1} + \varepsilon_t \\
\sigma_t^{\gamma} &= \omega + \alpha_1 \sigma_{t-1}^{\gamma}(|z_{t-1} - \zeta_{21}| - \zeta_{11}\{z_{t-1} - \zeta_{21}\})^{\delta} + \beta_1 \sigma_{t-1}^{\gamma} \\
\sigma_t^{1.9993} &= 0.0443 + 0.0900 \sigma_{t-1}^{1.9993}(|z_{t-1} - 1.3191| + 0.1919\{z_{t-1} - 1.3191\})^{1.9993} + 0.8113 \sigma_{t-1}^{1.9993}
\end{aligned}
$$

3.2 Application of the GAS Model

Next, we present the outcomes of the estimations involving the observation-driven GAS model. Hence, among seven selected assumed innovations, we use this study to determine an optimal or the most adequate assumed innovation distribution that captures volatility persistence estimation through the GAS model. The seven assumed innovations are the Normal, skew-Normal, Student's t, skew-Student's t, asymmetric Student's t with two tail decay parameters (AST), asymmetric Student's t with one tail decay parameter (AST1), asymmetric Laplace distribution (ALD). Details on these error distributions can be found in [5, 34, 43, 49].

Model selection and comparisons are carried out using two information criteria AIC and BIC as shown in Table 5. It is observed from the table that the GAS parameter estimates $\hat{\kappa}_\sigma$, $\hat{a}_\sigma$ and $\hat{b}_\sigma$ are statistically significant at 1% level under the seven innovation distributions, except for the $\hat{\kappa}_\sigma$ that is insignificant under the Student's t distribution and is 5% significant under the skew-Student's t distribution. The table's results further show that both the AIC and BIC have their lowest values under the AST1 innovation. Hence, the volatility persistence in this market's returns can be most adequately described through the GAS model fitted with the AST1 assumed innovation.

Table 5: Empirical outcomes of the GAS modelling on the real return data.

	Normal	skew-Normal	Student's t	skew-Student's t	AST	AST1	ALD
$\hat{\kappa}_\sigma$	0.0080*	0.0039*	-0.0030	0.0036**	0.0227*	**0.0226***	0.0048**
$\hat{a}_\sigma$	0.1239*	0.0305*	0.2048*	0.0499*	0.0509*	**0.0508***	0.0666*
$\hat{b}_\sigma$	0.9689*	0.9696*	0.9743*	0.9746*	0.9747*	**0.9749***	0.9724*
AIC	9275.222	9259.497	9109.057	9103.281	9094.407	**9092.654**	9110.983
BIC	9299.234	9289.512	9139.072	9139.299	9136.429	**9128.672**	9140.998

Note: The empirical outcomes of the fit of the GAS model. The "*" and "**" denote 1% and 5% levels of significance, respectively.

The estimated volatility persistence $\hat{b}_\sigma$ under this optimal AST1 distributed error is 0.9749. The outcome indicates that the GAS parameter $\hat{b}_\sigma$ in **B** is estimated close to unity with a volatility half-life of about 27 days, which implies a considerable persistent dynamic process for $\boldsymbol{\vartheta}_t$. This persistence outcome of 0.9749 coincides with the persistence outcome obtained from the best model fGARCH(1,1)-NIG (i.e., the fGARCH(1,1) model fitted with the NIG assumed distribution) in Section 3.1.7. The high persistence outcome is consistent with the findings of Pandey and Kumar [10], who used the GARCH(1,1) model on the Indian S&P CNX NIFTY 50 for the sample period 1997 to 2012 and found high persistence in the volatility process. However, this study used a more robust approach that involved a comparative use of the omnibus fGARCH and GAS models to estimate the persistence of the return volatility.

Panel B of Figure 3 shows that the impact of shocks through the GAS-AST1 model decays toward zero mean reversion, where the persistence impact dropped by half in about 27 days. That is, it follows the same trajectory as in fGARCH(1,1)-NIG model in Panel A, thus both models have the same persistence outcomes. The outcome further suggests the existence of long memory in the volatility of the returns. In addition, the estimated value of the unconditional scale (volatility or variance) under the GAS-AST1 model for the S&P Indian market over the sample period is 2.4605, implying a mean annualised volatility of 24.90%. To summarise, this study shows that when the underlying true innovation distribution is unknown, the fGARCH model fitted with the NIG assumed innovation and the GAS model fitted with the AST1 are the most suitable to describe the returns for volatility persistence estimation in the S&P Indian market.

From the outputs of the GAS-AST1 model in Table 5, the equation of the GAS model for the time-varying scale (or volatility) parameter can be stated as:

$$\boldsymbol{\vartheta}_{t+1} = 0.0226 + 0.0508\boldsymbol{s}_t + 0.9749\boldsymbol{\vartheta}_t. \tag{56}$$

3.3 Application of the Beta-Skew-t-EGARCH Model

These estimates of the persistence from both the fGARCH and GAS models indicate that the returns volatility exhibits considerable long memory since the persistence tends towards (or is close to) one. Moreover, their mean reversions (from the half-life computations) are somehow slow, which implies that the volatility of returns approaches the average or long-run volatility slowly. Based on this, we carry out further investigations on the long-memory behaviour of volatility decomposition into long-term and short-term components, and possible asymmetry (or leverage effect), fat-tails and skewness in the process using the one- and two-component Beta-Skew-t-EGARCH models, and then compare their outcomes. To be precise, we use both the one- and two-component models to estimate leverage effects, fat-tails, and skewness in the returns. We further use the two-component model to investigate the long-memory decomposition of volatility. That is, the two-component model is used to determine if the persistence of volatility can be decomposed into long-term and short-term processes. The results of fitting these one- and two-component Beta-Skew-t-EGARCH models are then compared to determine a superior fit. The conditional error of the Beta-Skew-t-EGARCH model is distributed as a skew Student's t (see [1, 2]). Skewness η in the two models is applied through the method proposed by Fernández [85] (see [1, 2, 34]). Hence, $\eta < 1 (\eta > 1)$ implies left (right) skewness.

We begin with the one-component modelling. Panel A of Figure 4 displays the plot of the fitted conditional standard deviations for the one-component model. The plot shows that the return series is characterised by time-varying volatility with a strong volatility spike caused by the global COVID-19 pandemic crisis in 2020. The outcomes of estimation through the one-component Beta-Skew-t-EGARCH model are presented in Panel A of Table 6. From the table (in Panel A), the estimated degrees of freedom $\hat{\nu}$ in the skew Student's t innovation is 6.2102, which is a reasonably fat-tailed conditional Student's t density. The skewness estimate $\hat{\eta}$ is about 0.8709, which relates to pronounced negative skewness ($\eta < 1$) in the residuals z_t.

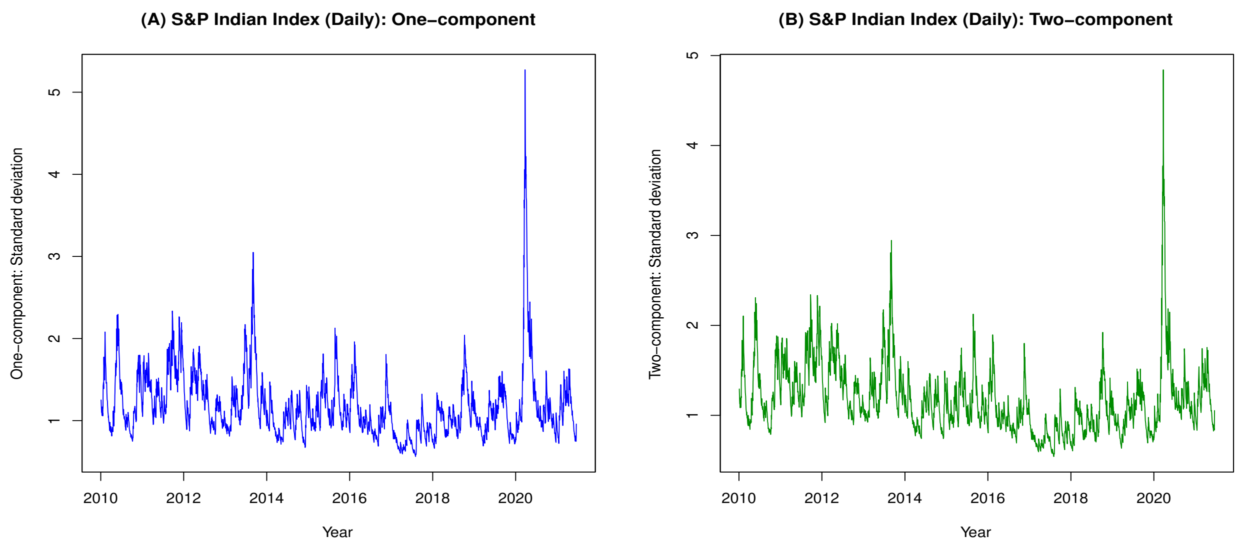

Fig. 4: Panels (A) and (B) present the fitted conditional standard deviations of the one- and two-component models, respectively.

The leverage effect estimate $\hat{\kappa}^*$ is positive, which indicates that large negative returns are being followed by higher volatility. This implies that negative shocks or bad news will impact future volatility more than positive shocks or good news of the same size. In other words, the outcome shows an asymmetric effect with a stronger impact from negative

Table 6: Empirical and Simulation Outcomes of the Beta-Skew-t-EGARCH Model.

	Panel A		**Panel B**	
	Empirical Outcomes Beta-Skew-t-EGARCH Model		**Simulation Outcomes Beta-Skew-t-EGARCH Model**	
	One-Component	**Two-Component**	**One-Component**	**Two-Component**
$\hat{\omega}$ [se]	0.0142 [0.0482]	0.0529 [0.1293]	0.0212 [0.0430]	0.0461 [0.1106]
$\hat{\phi}_1$ [se]	0.9721 [0.0056]	0.9988 [0.0015]	0.9771 [0.0034]	0.9979 [0.0012]
$h2l_{\phi_1}$ (Days)		577		
$\hat{\phi}_2$ [se]	–	0.9550 [0.0092]	–	0.9529 [0.0078]
$h2l_{\phi_2}$ (Days)		15		
$\hat{\kappa}_1$ [se]	0.0402 [0.0054]	0.0076 [0.0030]	0.0413 [0.0041]	0.0151 [0.0038]
$\hat{\kappa}_2$ [se]	–	0.0301 [0.0061]	–	0.0238 [0.0062]
$\hat{\kappa}^*$ [se]	0.0337 [0.0039]	0.0379 [0.0044]	0.0324 [0.0030]	0.0414 [0.0035]
$\hat{\eta}$ [se]	0.8709 [0.0207]	0.8742 [0.0208]	0.8611 [0.0163]	0.8613 [0.0162]
$\hat{\nu}$ [se]	6.2102 [0.6855]	6.3534 [0.7171]	6.0746 [0.4639]	6.3577 [0.5032]
BIC	3.0255	3.0266		

Note: $\hat{\omega}$ estimates the long-term log-volatility, $\hat{\phi}_1$ and $\hat{\phi}_2$ are estimators for the long-term and short-term persistence parameters, respectively, and $\hat{\kappa}_1$ ($\hat{\kappa}_2$) estimates the long-run (short-run) response to shocks. $h2l_{\phi_1}$ and $h2l_{\phi_2}$ are measures of the half-life for $\hat{\phi}_1$ and $\hat{\phi}_2$, respectively. $\hat{\kappa}^*$ estimates the leverage parameter, while $\hat{\eta}$ and $\hat{\nu}$ are estimates of the skewness and degrees of freedom, respectively. [se] in square bracket is the standard error of the estimated parameter.

shocks. In the event of asymmetric effects, Mishra [101] reported that investors' attention becomes more short-term focused. Investors tend to constantly review their investment portfolios for liquidity and performance, even if such investments are purchased with a long-term view. This could have an adverse influence on economic growth and business investment spending because investors tend to move their funds to more liquid and less risky assets as a result of such effects.

Next is the two-component modelling. Panel B of Figure 4 displays the plot of the fitted conditional standard deviations for the two-component model. Panel A of Table 6 presents the outcomes of the estimations involving the two-component model. From the table, the degrees of freedom estimate in the skew Student's t innovation is 6.3534, suggesting fat-tails in the conditional Student's t density. The estimated skewness is about 0.8742, which indicates a pronounced negative skewness. This implies that the risk of a large negative demeaned stock return is greater than that of a large positive demeaned stock return. The persistence of shocks in the long-run component $\hat{\phi}_1$ is very high at 0.9988, with a mean-reversion half-life ($h2l_{\phi_1}$) of about 577 days (i.e., about a year and seven months). This long-run half-life outcome suggests that the long-run effect of the volatility, which was partly caused by the 2020/2021 global COVID-19 pandemic crisis, would persist for as long as about a year and seven months before returning halfway back to the normal state. This indicates that even if the volatility of returns appears to have quite a long memory, it will still mean revert since the persistence estimate is less than one [3, 100]. This implies that even though it takes a long time to revert, the volatility process does go back to its mean [3, 100]. Due to the persistence caused by the pandemic outbreak, the Indian economy contracted by 6.6 percent during the fiscal year 2021 but staged a mild recovery in the fiscal year 2022 when it grew 8.7 percent [102].

The persistence of shocks in the short-run component $\hat{\phi}_2$ is 0.9550, with a half-life ($h2l_{\phi_2}$) outcome of about 15 days. Hence, the short-run component decays much faster than the long-run component. In other words, it can be seen from the persistence outcomes that $\hat{\phi}_2 < \hat{\phi}_1 < 1$ and from the half-life results that $h2l_{\phi_2} < h2l_{\phi_1}$, hence the short-run component decays more quickly than the long-run component that dominates the volatility persistence process. Panel A (Panel B) of Figure 5 shows that the decay of persistence impact in the long-term (short-term) reached half in about 577 days (15 days) and it continues towards zero mean reversion.

The parameter estimate $\hat{\kappa}_1$ that indicates the long-run response of volatility to shocks is 0.0076, while the estimate $\hat{\kappa}_2$ for the short-run response to shocks is 0.0301. These estimates show a significant discrepancy, where the outcomes reveal that the unexpected arrival of news influences the short-run component considerably more than the long-run component. In other words, the response to the effect of shocks in the short-run is higher than in the long-run volatility. More precisely, the long-run component displays smaller proportional effects of about 20 percent response to volatility shocks as compared to 80 percent from the short-run component[13]. However, even though the short-run component has a stronger shock effect, it is short-lived (see [86, 99, 100] for related outcomes).

[13] The total displayed response to shocks via $\hat{\kappa}_1$ and $\hat{\kappa}_2$ from the model as shown in Table 6 is 0.0377 (i.e., 0.0076 + 0.0301). Hence, there are 20.16% $\approx$ 20% and 79.84% $\approx$ 80% responses to volatility shocks by the long-term and short-term components, respectively.

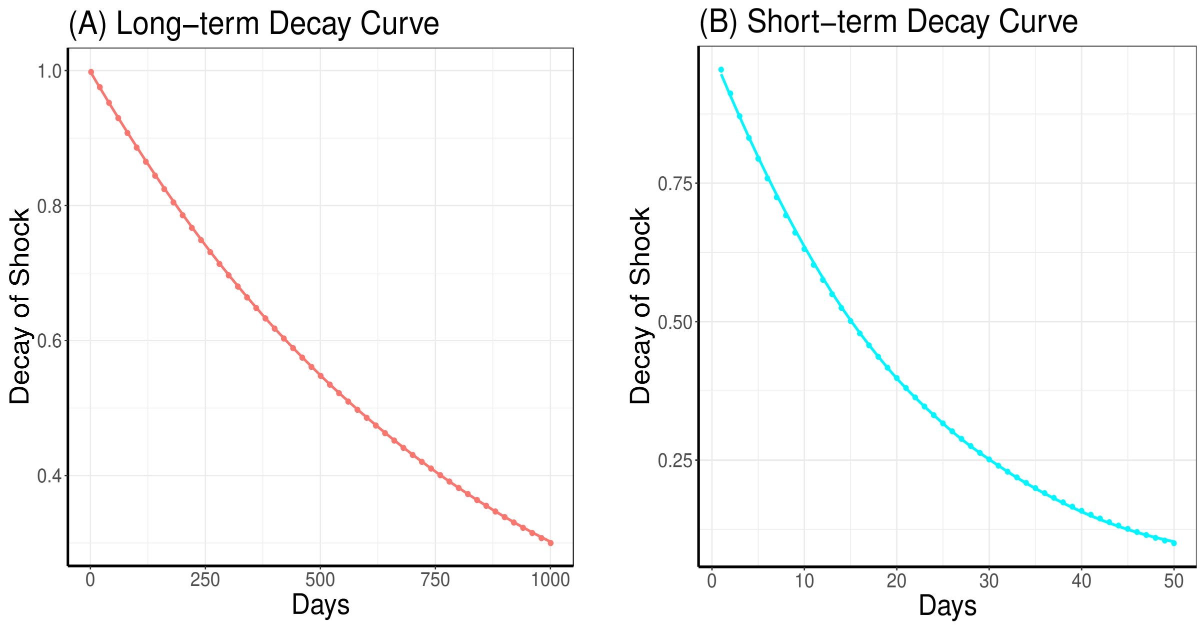

Fig. 5: Panels (A) and (B) present the long-term and short-term decay curves, respectively.

To summarise, this study finds the existence of both short-term and long-term volatility in the persistence process, where the response to the effect of shocks in the short-run is much higher than in the long-run volatility. This infers that higher volatility in the process is mostly due to the short-run volatility increase. However, the short-run volatility fluctuation is brief, while the long-run mean-reversion of volatility persistence will dominate thereafter. Precisely, the short-run effect is big but short-lived. Although the long-run component displays a pronounced longer persistence into the future, its response to volatility shocks is much lower than that of the transient short-run component. This means that investment and other market risks in the long term seem to be considerably under control in the market.

3.4 Simulation Study

Following the empirical outcomes, we run a set of Monte Carlo experiments using the true parameter outcomes from the Beta-Skew-*t*-EGARCH specification to further ascertain the validity of both the one- and two-component Beta-Skew-*t*-EGARCH model's results.

For the one-component simulation, we use sample size $N = 5{,}000$ simulated returns, generated from the true parameters $\omega = 0.0142$, $\phi_1 = 0.9721$, $\kappa_1 = 0.0402$, $\kappa^* = 0.0337$, $\eta = 0.8709$, and $\nu = 6.2102$. These true parameter values are empirical outcomes (i.e., MLE estimates) from fitting the first order one-component Beta-Skew-*t*-EGARCH model to the real Indian returns data. We use seed value 12345 for the simulation. Next, we fit the one-component Beta-Skew-*t*-EGARCH model to the simulated dataset and obtained the outcomes as presented in Panel B of Table 6 under the "One-Component" model. From the table, the estimated degrees of freedom $\hat{\nu}$ in the skew Student's t is 6.0746, the leverage parameter estimate $\hat{\kappa}^*$ is 0.0324, while the skewness estimate $\hat{\eta}$ is about 0.8611. These outcomes are quite close to the empirical results in Panel A of the table.

For the two-component simulation, we follow the same steps by using $N = 5{,}000$ simulated returns, generated from the true parameters $\omega = 0.0529$, $\phi_1 = 0.9988$, $\phi_2 = 0.9550$, $\kappa_1 = 0.0076$, $\kappa_2 = 0.0301$, $\kappa^* = 0.0379$, $\eta = 0.8742$, and $\nu = 6.3534$. We use seed value 12345 for the simulation. The true parameter values are obtained by fitting the first order two-component Beta-Skew-*t*-EGARCH model to the real Indian returns data. Next, we fit the two-component Beta-Skew-*t*-EGARCH model to the simulated dataset and obtained the outcomes as presented in Panel B of Table 6 under the "Two-Component" model. The estimates of the long-run $\hat{\phi}_1$ and short-run $\hat{\phi}_2$ are 0.9979 and 0.9529, respectively, and they are quite close and consistent with the empirical outcomes in Panel A of the table.

Moreover, given a 95% nominal recovery level, the True Parameter Recovery (TPR) outcome for the leverage parameter estimate $\hat{\kappa}^*$ in the one-component model is 91.34%, while the TPR outcome for the long-run (short-run) estimate is 94.91% (94.79%) in the two-component model. These outcomes indicate a good performance of the

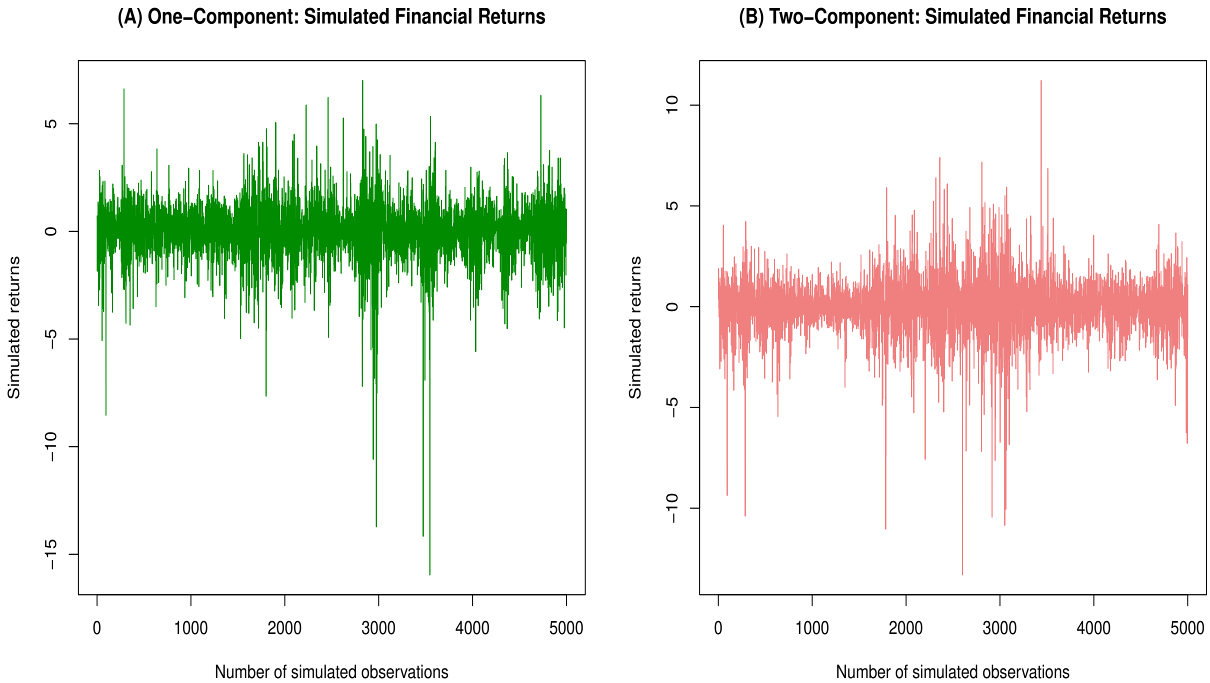

Fig. 6: Panels (A) and (B) display the plots for one-component simulated returns and two-component simulated returns, respectively.

simulation experiments. Hence, the simulation experiments performed considerably well with suitably valid outcomes. Thus, the leverage effect is evident, with long-memory decomposition into the long-run and short-run components. The response to volatility shocks is more pronounced in the short-lived short-run component than it is in the long-run component. Panels A and B of Figure 6 display the plots for one- and two-component simulated returns, respectively. Both plots show that the simulation return series is characterised by time-varying volatility.

3.5 Model Comparison

Next, we compare the results from the one- and two-component Beta-Skew-t-EGARCH model through the BIC[14] values, and the outcomes suggest that the one-component model outperforms the two-component model. This outcome is consistent with that of Harvey and Sucarrat [1], where the one-component model also performed better than the two-component model; but the outcome is in contrast with the findings of Sucarrat [2], where the two-component model outperformed the one-component model. Moreover, as reported by Harvey and Sucarrat [1], the use of the two-component model does not always give a better fit.

3.6 Discussion

In this study, we used three autoregressive models comprising the fGARCH, GAS, and Beta-Skew-t-EGARCH models to estimate six features of return volatility that are relevant for robust risk management in the S&P Indian market index. These features are stylised facts that characterise the market, and they include volatility persistence, mean-reversion, asymmetry (or leverage effect), skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components. The ability of these models to capture these stylised facts provided a high degree of robustness

[14] We used BIC for model selection because it shows consistency as the sample size increases, such that the criterion will select a true model of finite dimension if it is included among the candidate models [103]. In other words, the BIC chooses the true model with probability one, with the assumption that the true model is among the set of selected candidate models (see [104, 105]). Moreover, using consistency property, the BIC discourages overparameterisation where it imposes more heavy penalties than the AIC on model complexity [95]. Hence, the BIC potentially favours highly parsimonious models which are neither overly simple nor overly complex but usually lie between the two [106].

for volatility modelling in the market's returns. Both simulations and empirical evidence were used to show the accuracy of the estimations.

To begin with, the study comparatively used the robust fGARCH and GAS models to estimate the magnitude and dynamics of the persistence in conditional volatility using the returns from the market index. The outcomes of the estimations found the NIG and AST1 assumed errors as the most adequate (or optimal) error distributions to use with the fGARCH and GAS models, respectively, for volatility modelling when the underlying error distribution is unknown in the returns. Hence, the NIG (AST1) distribution may be widely used with the fGARCH (GAS) model to improve the accuracy of volatility modelling for risk measures in finance and other areas. Appropriate risk management with proper economic policy implementation could create a channel for profit maximisation by financial institutions and individual investors [107].

By fitting each of the optimal assumed innovations to their respective model, the study found considerably high volatility persistence in the market returns. The high persistence suggests the presence of volatility clustering in the returns. Knowledge about the clustering of volatility allows market agents to adopt dynamic and flexible trading strategies that are suitable either for high-volatility or low-volatility regimes [10]. The study further showed that the fGARCH and GAS models performed equally well in the volatility persistence (and mean-reversion) estimations when fitted with their respective optimal assumed errors. Next, we comparatively used the one- and two-component Beta-Skew-t-EGARCH models to estimate other features of the return volatility that include leverage effect or asymmetry, skewness, fat-tails, and the long-memory behaviour of volatility decomposition into long-term and short-term components. Specifically, we used both the one- and two-component models to estimate leverage effects, fat-tails, and skewness in the returns. Through the one- and two-component models, our findings show that negative skewness and leverage effects are pronounced, with considerable fat-tails in the conditional density. The leverage estimate is positive, which indicates that large negative returns are being followed by higher volatility. The pronounced negative skewness estimate implies that the risk of a large negative stock return is greater than that of a large positive stock return.

Also, we used a parametric approach through the ARFIMA-FIGARCH models, and three semi-parametric approaches via the log periodogram estimator of Geweke and Porter-Hudak (GPH), the local Whittle (LW) estimator, and the exact local Whittle (ELW) estimator to estimate and determine the presence of long memory in the returns, absolute returns, and squared returns. The results of the estimations indicate that the daily returns, squared returns, and absolute returns exhibit long memory, hence, shocks decay at a slower rate. However, the persistence is lower in the returns when compared to the squared and absolute returns.

Furthermore, we used the two-component Beta-Skew-t-EGARCH model to investigate the long-memory decomposition of volatility. Through this two-component model, the study found the existence of both long-run and short-run components of volatility in the persistence process, but the response to the effect of shocks in the short-run is higher than in the long-run volatility. This response to shock effects is also part of the findings of the fGARCH modelling. This implies that higher volatility in the process is mostly due to the short-run volatility increase. Further findings through the two-component Beta-Skew-t-EGARCH model using the half-life estimation showed that the short-run volatility fluctuation reverts much faster to the mean or normal volatility state than the long-run volatility persistence. Consequently, these results show that with the arrival of news in the stock market, the long-run component displays a much lower response to the effect of shocks (or change to volatility) than the transient short-run component. In summary, the short-run effect is big but short-lived, while the long-run effect is much lower but persists into the future. This means that investment and other market risks in the long term seem to be considerably under control in the market.

Lastly, a comparison of the two versions of the Beta-Skew-t-EGARCH model showed that the one-component model outperformed the two-component model. These discussed outcomes summarily answered the four research questions, and the study shows that the market returns are characterised by the six stated volatility features, namely, pronounced persistence, mean reversion, leverage effect or volatility asymmetry, conditional skewness, conditional fat-tailedness, and the long memory behaviour of volatility decomposition into long-term and short-term components.

4 Conclusion

In conclusion, this study largely contributes to the literature by comparing the fit of two robust models involving the fGARCH and GAS. The fGARCH model uses the dynamics of the residuals to drive the conditional volatility while the GAS model uses the dynamics of the conditional score to drive the time-varying parameters of the time series process. To the best of the authors' knowledge, the comparison of these two robust models has not been carried out in previous studies for modelling the dynamics of volatility. Risk management systems are highly dependent on the underlying assumed distribution, and the identification of a distribution that adequately captures every aspect of the given financial data. This may be of great benefit to investors and risk managers [108]. Hence, our findings showed that to determine a reliable volatility (risk) modelling approach in a financial time series with unknown underlying error distribution, the NIG (AST1) assumed error could be recommended for use with the fGARCH (GAS) model. The study further revealed considerably high volatility persistence, negative skewness, leverage effect and fat-tails in the S&P Indian financial market returns.

Lastly, our findings from the long-memory behaviour of volatility decomposition revealed that although the response to shocks is greater in the short-term component, it is however short-lived. On the contrary, despite a high degree of persistence in the long-term component, market information or unexpected news arrival only has a low long-run impact on the market. Based on this, the long-run investment risks within the Indian stock market seem to be under control. Hence, our findings suggest that rational investors should try to stay calm with the arrival of unexpected news or unforeseen events in the Indian stock market because the long-run effect of such news will not be severe, and the market will eventually return to its normal state. With the presence of short-term and long-term components and their impacts on the market, this study also suggests that market managers and government should make efforts to understand the implications of changes in their system of trading and policies implemented. Such moves (or actions) will facilitate improvement in the market activities and further enable them to better control risks in the market.

For future studies, the authors intend to explore the functionalities of other robust time-varying models for volatility modelling and forecasting in multiple market indices. Specifically, the authors intend to use the apARCH model [42] as an extension of the GARCH model, and the score-driven extensions involving the two-component Beta-t-QVAR-M-lev model proposed by Haddad et al. [18], and the Beta-t-EGARCH model with random shifts (RS-Beta-t-EGARCH model) developed by Alanya-Beltran [45].

Acknowledgement

The authors are grateful to Thomson Reuters Datastream for providing the data, and to the University of the Witwatersrand and the University of Venda for their resources.

References

[1] A. Harvey and G. Sucarrat, EGARCH models with fat tails, skewness and leverage. Computational Statistics and Data Analysis, **26**, 320-338 (2014).
[2] G. Sucarrat, betategarch: Simulation, estimation and forecasting of first-order beta-skew-t-EGARCH models. The R Journal, **5**, 137-147 (2013).
[3] R.F. Engle and A.J. Patton, What good is a volatility model?. Quantitative Finance, **1**, 237-245 (2001).
[4] J. Porterba and L. Summers, The persistence of volatility and stock market fluctuations. American Economic Review, **76**, 1143-1151 (1986).
[5] D. Ardia, K. Boudt, and L. Catania, Generalized autoregressive score models in R: The GAS package. Journal of Statistical Software, **88**, 1-28 (2019).
[6] R.Y. Chou, Volatility persistence and stock valuations : Some empirical evidence using GARCH. Journal of Applied Econometrics, **3**, 279-294 (1988).
[7] T. Takaishi, Volatility estimation using a rational GARCH model. Quantitative Finance and Economics, **2**, 127-136 (2018).
[8] L. Feng and Y. Shi, A simulation study on the distributions of disturbances in the GARCH model. Cogent Economics and Finance, **5**, 1355503 (2017).
[9] L. Hentschel, All in the family nesting symmetric and asymmetric GARCH models. Journal of Financial Economics, **39**, 71-104 (1995).
[10] R. Pandey and A. Kumar, Modelling persistence in conditional volatility of asset returns. Afro-Asian Journal Finance and Accounting, **7**, 16-34 (2017).
[11] R.F. Engle and T. Bollerslev, Modelling the persistence of conditional variances. Econometric Reviews, **5**, 1-50 (1986).
[12] D. Oh and A. Patton, The time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads. Journal of Business and Economic Statistics, **36**, 181-195 (2016).
[13] D. Creal, B. Schwaab, S. Koopman, and A. Lucas, Observation-driven mixed-measurement dynamic factor models with an application to credit risk. The Review of Economics and Statistics, **96**, 898-915 (2014).
[14] F. Blasques, S. Koopman, and A. Lucas, Maximum likelihood estimation for generalized autoregressive score models. TI 2014-029/III, Tinbergen Institute Discussion Paper, 1-65 (2014).
[15] L. Catania and A. Billé, Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Applied Econometrics, **32**, 1178-1196 (2017).
[16] A. Opschoor, P. Janus, A. Lucas, and D. Van Dijk, New HEAVY models for fat-tailed realized covariances and returns. Journal of Business and Economic Statistics, **36**, 643-657 (2018).
[17] P. Gorgi, P. Hansen, P. Janus, and S. Koopman, Realized Wishart-GARCH: a score-driven multi-asset volatility model. Journal of Financial Econometrics, **17**, 1-32 (2019).
[18] M.F.C. Haddad, S. Blazsek, P. Arestis, F. Fuerst, and H.H. Sheng, The two-component Beta-t-QVAR-M-lev: a new forecasting model. Financial Markets and Portfolio Management, **37**, 379-401 (2023).

[19] F. Sowell, Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics, **53**, 165-188 (1992).
[20] Y. Cheung, Test for fractional integration: A Monte Carlo investigation. Journal of Time Series Analysis, **14**, 331-345 (2008).
[21] T. Bollerslev and J. Wright, Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. Journal of Econometrics, **98**, 81-106 (2000).
[22] G. Shea, Uncertainty and implied variance bounds in long-memory models of the interest rate term structure. Empirical Economics, **16**, 287-312 (1991).
[23] G. Bobeica and E. Bojesteanu, Long memory in volatility. An investigation on the central and eastern European exchange rates. European Research Studies, **11**, 7-18 (2008).
[24] J. Vera-Valdés, LongMemory.jl: Generating, estimating, and forecasting long memory models in Julia. arXiv:2401.14077v1, (2024).
[25] Thomson reuters datastream. [Online]. Available at: Subscription Service (Accessed: June 2021), 2021.
[26] S. Hertog, P. Gerland, and J. Wilmoth, India overtakes China as the world's most populous country. UN DESA Population Division, World Population Prospects: Summary of Results, 1-5 (2022).
[27] T. Bollerslev, Generalized autoregressive conditional heteroskedastic. Journal of Econometrics, **31**, 307-327 (1986).
[28] R.F. Engle, Autoregressive conditional heteroscedacity with estimates of variance of United Kingdom inflation. Econometrica, **50**, 987-1008 (1982).
[29] A.C. Harvey, Dynamic models for volatility and heavy tails: With applications to financial and economic time series. Cambridge University Press: New York, USA. **37**, 1-261 (2013).
[30] R. Baillie, T. Bollerslev, and H. Mikkelsen, Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, **74**, 3-30 (1996).
[31] Baillie, R., Long memory processes and fractional integration in econometrics, Journal of Econometrics, **73**, 5-59 (1996).
[32] N. Rafik, N. Esmaeil, G. Nadiya, and A. Ashkan, Long memory analysis: An empirical investigation. International Journal of Economics and Financial Issues, **4**, 16-26 (2014).
[33] A. Ghalanos, rugarch: Univariate GARCH models. R package version 1.4-7. (2022). https://cran.r-project.org/web/packages/rugarch/rugarch.pdf (accessed on 30 August 2023).
[34] A. Ghalanos, Introduction to the rugarch package. (Version 1.3-8), (2018). https://mirrors.nic.cz/R/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf (accessed on 30 August 2023)
[35] J.M. Zakoian, Threshold heteroscedastic models. Journal of Economic Dynamics and Control, **18**, 931-955 (1994).
[36] R.F. Engle and V.K. Ng, Measuring and testing the impact of news on volatility. The journal of Finance, **48**, 17749-1778 (1993).
[37] S.J. Taylor, Modelling financial time series. World Scientific Publishing Company Pte. Ltd., 1-297 (1986).
[38] G.W. Schwert, Stock volatility and the crash of '87. The Review of Financial Studies, **3**, 77-102 (1990).
[39] M.L. Higgins, and A.K. Bera, A class of nonlinear Arch models. International Economic Review, **33**, 137-158 (1992).
[40] B.D. Nelson, Conditional heteroskedasticity in asset returns : A new approach. Econometrica, **59**, 347-370 (1991).
[41] L.R. Glosten, R. Jagannathan, and D.E. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance, **48**, 1779-1801 (1993).
[42] Z. Ding, C.W. Granger, and R.F. Engle, A long memory property of stock market returns and a new model. Journal of Empirical Finance, 83-106 (1993).
[43] D. Creal, S.J. Koopman, and A. Lucas, Generalized autoregressive score models with applications. Journal of Applied Econometrics, **28**, 777-795 (2013).
[44] L. Catania, D. Ardia, and K. Boudt, Generalized autoregressive score models. Package 'GAS' Version 0.3.3 (2020). https://doi.org/10.18637/jss.v088.i06 (accessed on 6 January 2025)
[45] W. Alanya-Beltran, Modelling stock returns volatility with dynamic conditional score models and random shifts. Finance Research Letters, **45**, 102121 (2022).
[46] D. Creal, S.J. Koopman, and A. Lucas, Generalized autoregressive score models. Aenorm, **20**, 37-41 (2012).
[47] S.J. Koopman, A. Lucas, and M. Zamojski, Dynamic term structure models with score driven time varying parameters: estimation and forecasting. NBP Working Paper, No. 258, 3-80 (2017).
[48] F. Blasques, D.D. Creal, P. Janus, S.J. Koopman, A. Lucas, M. Scharth, and B. Schwaab, Generalized autoregressive score models for time-varying parameters: new models and applications, (2014). https://www.gasmodel.com/2014GAS_ISFrotterdam.pdf (accessed on 26 June 2024).
[49] D. Creal, S.J. Koopman, and A. Lucas, A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. Journal of Business and Economic Statistics, **29**, 552-563 (2011).
[50] J. Barkoulas, W. Labys, and J. Onochie, Long memory in futures prices. Financial Review, **34**, 91-100 (1999).
[51] K. Brania and H. Gurgul, The impact of estimation methods and data frequency on the results of long memory assessment. Managerial Economics, **16**, 7-37 (2015).
[52] H. Hurst, Long term storage capacity of reservoirs. Proceedings of the American Society of Civil Engineers, **11**, 770-99 (1951).
[53] B. Mandelbrot, A fast fractional Gaussian noise generator, Water Resources Research, **7**, 543-553 (1971).
[54] B. Mandelbrot and J. Van Ness, Fractional Brownian motion, fractional noises and applications, SIAM Review, **10**, 422-437 (1968).
[55] J. Hosking, Fractional differencing. Biometrika, **68**, 165-176 (1981).
[56] J. Geweke and S. Porter-Hudak, The estimation and application of long memory time series models. Journal of Time Series Analysis, **4**, 221-238 (1983).

[57] K. Saha, V. Madhavan, and G. Chandrashekhar, Pitfalls in long memory research. Cogent Economics and Finance, **8**, 1733280 (2020).
[58] J. Beran, Statistics for long-memory processes, Chapman and Hall, (1994). https://api.semanticscholar.org/CorpusID:62495143 (accessed on 25 August 2024).
[59] J. Vera-Valdés, Temperature anomalies, long memory, and aggregation. Econometrics, **9**, 1-22 (2021).
[60] S. Holan, T. McElroy, and S. Chakraborty, A Bayesian approach to estimating the long memory parameter. Bayesian Analysis, **4**, 159-190 (2009).
[61] C. Granger and R. Joyeux, An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis, **1**, 15-29 (1980).
[62] W. Wang, P. Van Gelder, J. Vrijling, and X. Chen, Detecting long-memory: Monte Carlo simulations and application to daily streamflow processes. Hydrology and Earth System Sciences, **11**, 851-862 (2007).
[63] C. Baum, S. Hurn, and K. Lindsay, Local Whittle estimation of the long-memory parameter. The Stata Journal, **20**, 565-583 (2020).
[64] C. Baum and V. Wiggins, Estimate long memory in a time series via modified log-periodogram regression, (2024). http://fmwww.bc.edu/RePEc/bocode/m/modlpr.html (accessed on 5 August 2024).
[65] R. Dahlhaus, Efficient parameter estimation for self-similar processes. Annals of Statistics, **17**, 1749-1766 (1989).
[66] R. Fox and M. Taqqu, Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. Annals of Statistics, **14**, 517-532 (1986).
[67] S. Lopes, B. Olbermann, and V. Reisen, Non-stationary Gaussian ARFIMA processes: Estimation and application, Brazilian Review of Econometrics Rio de Janeiro, **22**, 103-126 (2002).
[68] H. Boubaker, G. Canarella, R. Gupta, and S. Miller, Long-memory modeling and forecasting: evidence from the U.S. historical series of inflation. Studies in Nonlinear Dynamics and Econometrics, **25**, 289-310 (2021).
[69] J. Fouquau and P. Spieser, Stock returns memories: a "stardust" memory?. Dans Finance, **35**, 57-85 (2014).
[70] C. Conrad and B. Haag, Inequality constraints in the fractionally integrated GARCH model. Journal of Financial Econometrics, **4**, 413-449 (2006).
[71] Y. Xunyu, G. Ping, and L. Handong, Improving estimation of the fractionally differencing parameter in the SARFIMA model using tapered periodogram. Economic Modelling, **46**, 167-179 (2015).
[72] P. Robinson, Gaussian semiparametric estimation of long range dependence. Annals of Statistics, **23**, 1630-1661 (1995a).
[73] P. Robinson, Time series with strong dependence. Advances in econometrics, 6th world congress (Cambridge University Press, Cambridge) (1990).
[74] C. Agiakloglou, P. Newbold and M. Wohar, Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis, **14**, 235-246 (1992).
[75] C. Velasco, Non-stationary log-periodogram regression. Journal of Econometrics, **91**, 325-371 (1999).
[76] C. Kim and P. Philips, Log periodogram regression in the nonstationary case. Mimeo, Yale University (1999).
[77] C. Hurvich, R. Deo and J. Brodsky, The mean squared error of geweke and porter-hudak's estimator of the long memory parameter of a long-memory time series. Journal of Time Series Analysis, **19**, 1095-1112 (1998).
[78] P. Whittle, Hypothesis testing in time series analysis. Uppsala: Almqvist and Wiksells, (1951). https://search.worldcat.org/title/Hypothesis-testing-in-time-series-analysis/oclc/22153644 (accessed on 27 August 2024).
[79] Y. Cheung and F. Diebold, On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. Journal of Econometrics, **62**, 301-316 (1994).
[80] H. K*ü*nsch, Statistical aspects of self-similar processes. in Prokhorov, Y. and Sazanov, V.V. (eds), Proceedings of the First World Congress of the Bernoulli Society. Utrecht, VNU Science Press, 67-74 (1987).
[81] P. Robinson, Log-periodogram regression of time series with long range dependence. Annals of Statistics, **23**, 1048-1072 (1995b).
[82] K. Shimotsu and P. Phillips, Exact local Whittle estimation of fractional integration. Annals of Statistics, **33**, 1890-1933 (2005).
[83] A. Harvey and T. Chakravarty, Beta-t-EGARCH. Cambridge working papers in economics 0840. Faculty of Economics, University of Cambridge, p137 (2008).
[84] D. Creal, S. Koopman, and A. Lucas, A general framework for observation driven time-varying parameter models. Tinbergen Institute Discussion Paper, Amsterdam, TI, 2008-108/4 (2008).
[85] C. Fern*á*ndez and M. Steel, On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association, **93**, 359-371 (1998).
[86] R. Engle and G. Lee, A permanent and transitory component model of stock return volatility, in R.F. Engle and H. White (eds.), cointegration, causality, and forecasting: A festschrift in honor of Clive W.J. Granger. Oxford University Press, New York, 475-497 (1999).
[87] S. Alizadeh, M. Brandt and F. Diebold, Range-based estimation of stochastic volatility models. Journal of Finance, **57**, 1047-1092 (2002).
[88] P. Chalmers, Introduction to Monte Carlo simulations with applications in R using the SimDesign package. Journal of Finance, 1-46 (2019).
[89] T. Samuel, C. Chimedza, and C. Sigauke, Simulation framework to determine suitable innovations for volatility persistence estimation: The GARCH approach. Journal of Risk and Financial Management, **16**, 392 (2023).
[90] T. Hilary, Descriptive statistics for research. (2002). https://www.stats.ox.ac.uk/pub/bdr/IAUL/Course1Notes2.pdf (accessed on 6 January 2025).

[91] T. Bollerslev, Conditionally heteroskedasticity time series model for speculative prices and rates of returns. The Review of Economic and Statistics, **69**, 542-547 (1987).
[92] M. Eling, Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models?. Insurance: Mathematics and Economics, **59**, 45-56 (2014).
[93] A. Azzalini, A class of distributions which includes the normal ones, Scandinavian Journal of Statistics, **12**, 171-178 (1985).
[94] T.J. Fisher and C.M. Gallagher, New weighted portmanteau statistics for time series goodness of fit testing. Journal of the American Statistical Association, **107**, 777-787 (2012).
[95] F. Li, A.S. Cohen, S. Kim, and S. Cho, Model selection methods for mixture dichotomous IRT models, Applied Psychological Measurement, **33**, 353-373 (2009).
[96] K. Aas and I.H. Haff, The generalized hyperbolic skew student's t-distribution, Journal of Financial Econometrics, **4**, 275-309 (2006).
[97] D. Andrews and P. Guggenberger, A bias-reduced log periodogram regression estimator for the long-memory parameter. Econometrica, **71**, 675-712 (2003).
[98] C. Leschinski, M. Voges, and K. Wenger, Long memory time series. Package 'LongMemoryTS' Version 0.1.0. https://cran.r-project.org/web/packages/GAS/GAS.pdf (accessed on 22 July 2024).
[99] S. Chen and C. Shen, GARCH, jumps, and permanent and transitory components of volatility: The case of the Taiwan exchange rate. Mathematics and Computers in Simulation, **67**, 201-216 (2000).
[100] S.M. Chiang, C.P. Yeh, and C.L. Chiu, Permanent and transitory components in the Chinese stock market: The ARJI-Trend model, Taylor and Francis, **45**, 35-55 (2006).
[101] P.K. Mishra, A GARCH model approach to capital market volatility: The case of India, Indian Journal of Economics & Business, **9**, 631-641 (2010).
[102] CRISIL, Indian economy: States of aftermath, CRISIL: An S&P Global Economy, (2022). https://www.crisil.com/en/home/our-analysis/reports/2022/07/crisil-insights-indian-economy-states-of-aftermath.html (Accessed on 6 January 2025).
[103] W. Christensen, Model selection using information criteria (Made Easy in SAS®), University of California, Los Angeles, wchristensen@ucla.edu, (2018). https://support.sas.com/resources/papers/proceedings18/2587-2018.pdf (accessed on 6 January 2025).
[104] R. Shibata, An optimal selection of regression variables. Biometrika, **68**, 45-54 (1981).
[105] R. Nishii, Asymptotic properties of criteria for selection of variables in multiple regression. Annals of Statistics, **12**, 758-765 (1984).
[106] P. Breheny, Model selection I, BST 760: Advanced regression, 1-32 (2013).
https://myweb.uiowa.edu/pbreheny/uk/teaching/760-s11/notes/2-17.pdf (accessed on 23 August 2024).
[107] B. Dralle, Modelling volatility in financial time series, Master of science in statistics, University of KwaZulu-Natal, (2011).
[108] K. Chinhamu, C.K. Huang, and D. Chikobvu, Evaluating risk in gold prices with generalized hyperbolic and stable distributions, Proceedings of the 57th Annual Conference of SASA, 17-24 (2017).