University of Helsinki
Faculty of Science
Masters programme in Mathematics and Statistics

Master's thesis

Univariate and Multivariate GARCH models

Chamini Wijeratne

Supervisor: Leena Kalliovirta

May 08, 2025

Title: Univariate and Multivariate GARCH models
Author: Chamini Wijeratne
Month and year: May 2025
Page count: 49 pp.

Abstract:

Understanding and forecasting financial market volatility is essential for effective risk management, asset allocation and pricing decisions. Traditional models often rely on assumptions of constant variance and symmetric responses to shocks, which fail to capture common features in financial return data, such as volatility clustering, asymmetric effects and dynamic correlations. This thesis applies a combination of univariate and multivariate GARCH models to study the volatility and correlation behavior of two major U.S. stocks, General Electric (GE) and IBM using monthly return data from 1926 to 2008. The univariate analysis begins with the GARCH(1,1) model to establish a benchmark for volatility persistence. To capture asymmetry in shock responses, the EGARCH(1,1) model is used. However, residual diagnostics indicated autocorrelation in the standardized residuals from the EGARCH(1,1) model, leading to the adoption of an extended AR(1)-EGARCH(1,1) model. This modification improves model fit and resolves the residual autocorrelation issue, offering a more robust specification for capturing both volatility asymmetry and return autocorrelation.

In the multivariate setting, the Dynamic Conditional Correlation, DCC-GARCH model is employed to model time-varying correlations while the BEKK-GARCH model provides insight into volatility spillovers and risk transmission across the two stocks. The results confirm persistent volatility clustering and significant asymmetry in responses to negative shocks. Furthermore, the correlation between GE and IBM varies over time, increasing during financial stress and decreasing in calmer periods. The DCC-GARCH model effectively captures these evolving correlations, whereas the BEKK-GARCH model highlights directional volatility transmission. These findings have practical relevance for investors and policymakers. Dynamic models of volatility and correlation support better diversification strategies and help identify periods of elevated market risk. By combining univariate and multivariate GARCH frameworks, including the extended AR(1)-EGARCH(1,1) model, this thesis offers an example of volatility dynamics and cross-asset interactions in financial markets.

Keywords: Volatility Modeling, GARCH, EGARCH, AR(1)-EGARCH(1,1), DCC-GARCH, BEKK-GARCH, Financial Risk, Time-Varying Correlations.

Contents

1. Introduction

Financial markets exhibit time-varying volatility and complex interdependencies among asset returns, making the modeling of volatility a crucial aspect of financial analysis. The presence of volatility clustering, where large fluctuations tend to be followed by further large fluctuations challenges the assumption of constant variance in classical financial models. Additionally, financial return series often display asymmetric volatility effects, where negative shocks lead to larger increases in volatility than positive shocks of the same magnitude. To address these stylized facts, time series models such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family have been widely adopted for modeling and forecasting financial market volatility (987; Engle, 1982; 308; Bollerslev, 1986).

This thesis focuses on modeling the volatility and correlation dynamics of General Electric (GE) and IBM monthly stock returns using both univariate and multivariate GARCH-type models. GE, a multinational industrial conglomerate, and IBM, a global technology company, represent two major sectors of the U.S. stock market. Examining their return behavior enables the analysis of volatility persistence, asymmetric effects, and time-varying comovement between sectors. The univariate modeling begins with the benchmark GARCH(1,1) specification, which captures volatility clustering and persistence. To address the asymmetric impact of shocks, the EGARCH(1,1) model is employed, allowing negative shocks to have a different influence on volatility than positive ones (317; Nelson, 1991). However, diagnostic checks revealed significant autocorrelation in the standardized residuals of the EGARCH(1,1) model for GE returns. To improve model adequacy, an autoregressive term was introduced in the conditional mean equation, resulting in the AR(1)-EGARCH(1,1) model. This modification successfully eliminated residual autocorrelation, improved model fit and maintained the ability to capture asymmetric volatility behavior.

For the multivariate analysis, the Dynamic Conditional Correlation, DCC-GARCH model is used to explore the evolution of time-varying correlations between GE and IBM, while the BEKK-GARCH model helps examine volatility spillovers and shock transmission across assets (349; Engle, 2002; 215; Engle and Kroner, 1995). The empirical findings confirm that volatility clustering and asymmetry are consistent features of both stocks. The AR(1)-EGARCH(1,1) model provides a better fit in the univariate case by improving residual behavior, while the DCC-GARCH model effectively captures the dynamic correlation structure, especially during turbulent market periods. The BEKK-GARCH model offers further insights into how volatility in one stock may influence the other, illustrating the interconnectedness of financial assets.

These results have important implications for risk management, asset allocation and regulatory policy. A better understanding of time-varying volatility and correlations allows for more informed portfolio diversification and helps identify systemic risk. This thesis applies standard univariate and multivariate GARCH models, including AR(1)-EGARCH(1,1), to provide a structured analysis of volatility dynamics and time-varying correlations in financial markets. The remainder of the thesis is organized as follows. Section 2 presents the theoretical background, introducing key concepts and detailing the univariate and multivariate GARCH-type models used in the analysis. Section 3 describes the data and reports the empirical results of the fitted models for GE and IBM stock returns. Section 4 concludes the study by summarizing the main findings and discussing their implications for financial modeling and risk management.

2. Univariate and Multivariate GARCH Models

Modeling financial market volatility is essential for risk management, asset pricing and portfolio allocation. Traditional econometric models assume constant variance over time, an assumption that does not hold in empirical financial data. One well-documented characteristic of financial returns is volatility clustering, where periods of high volatility tend to be followed by high volatility and periods of low volatility tend to be followed by low volatility. This phenomenon, first noted by (215; Mandelbrot, 1963) and later reinforced by (276; Fama, 1965), suggests that large price changes are more likely to be followed by further large changes, whether positive or negative, rather than by small changes. The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by (308; Bollerslev, 1986), extends the earlier Autoregressive Conditional Heteroskedasticity (ARCH) model of (987; Engle, 1982) by incorporating lagged conditional variances to better capture the persistence of volatility. A univariate GARCH model describes the conditional variance of a single financial time series, allowing for time-varying volatility that depends on past squared innovations and past variances. The persistence of volatility is determined by the sum of model parameters, which approaches or even surpasses unity for highly persistent volatility processes (310; Bollerslev, 1986).

While univariate GARCH models effectively describe the volatility of a single asset, financial markets often exhibit interdependencies between multiple assets. Multivariate GARCH (MGARCH) models extend the univariate framework by accounting for time-varying covariances among multiple financial assets. One widely used specification is the Dynamic Conditional Correlation GARCH (DCC-GARCH) model, proposed by Engle in 2002, which models conditional correlations dynamically while maintaining computational feasibility. Another prominent approach is the BEKK-GARCH model, introduced by Baba, Engle, Kraft and Kroner in 1991, which provides a flexible structure for capturing volatility spillovers and cross-asset risk transmission. Multivariate GARCH models consider the joint behavior of financial returns and their conditional covariance structure. The choice between different MGARCH models depends on the trade-off between flexibility and computational efficiency, as more complex models require greater computational resources (78; Engle and Kroner, 1995). Simpler models, while computationally efficient, may fail to capture the dynamic interdependencies between assets, leading to biased risk estimates. Empirical studies have shown that GARCH models effectively capture key stylized facts of financial returns, such as volatility clustering, leptokurtosis and leverage effects (see e.g. 317; Glosten et al., 1993; 322; Andersen et al., 2003). However, model selection and estimation remain challenging due to the large number of parameters in multivariate settings. Diagnostic checks for univariate models often include the Ljung-Box test for residual autocorrelation, while in multivariate settings, model adequacy is typically assessed using likelihood-based information criteria and residual analysis tailored to vector processes (see e.g. ; Tsay, 2010).

2.1 Basic Concepts

2.1.1 Random variable

A random variable provides a formal framework for representing uncertain outcomes in probability theory and statistical modeling. In this thesis, we focus on continuous vector-valued random variables, which are particularly relevant for modeling multivariate financial time series. Formally, a vector-valued random variable is a measurable function $X : \Omega \to \mathbb{R}^m$, where Ω is the sample space and $\mathbb{R}^m$ is the m-dimensional real space (see e.g. Feller, 1971). The distribution of such a random variable is characterized by its cumulative distribution function (CDF), defined as

$$F_X(x) = \mathbb{P}(X_1 \leq x_1, \ldots, X_m \leq x_m), \tag{2.1}$$

where $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$. If the CDF is absolutely continuous, then there exists a joint probability density function (PDF) $f_X(x)$ such that

$$F_X(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} f_X(u_1, \ldots, u_m)\, du_m \cdots du_1. \tag{2.2}$$

$$f_X(x) = \frac{\partial^m}{\partial x_1 \cdots \partial x_m} F_X(x). \tag{2.3}$$

(see e.g. Casella and Berger, 2002). The expectation $\mathbb{E}[X] \in \mathbb{R}^m$ and the covariance matrix $\mathrm{Cov}(X) \in \mathbb{R}^{m \times m}$ are fundamental descriptors of the distribution's central tendency and dispersion. The existence of these unconditional moments is a prerequisite for defining their conditional counterparts. In time series models, the conditional expectation and conditional covariance matrix describe how the distribution of a random variable evolves over time, given past information. These conditional quantities play a central role in financial econometrics, where models such as ARCH and GARCH are used to capture time-varying features of financial data, such as volatility clustering and dynamic correlations (see e.g. Engle, 1982; Bollerslev, 1986).

2.1.2 Stochastic Process

A stochastic process is a collection of random variables indexed by time, commonly used to model systems that evolve under uncertainty. Formally, a stochastic process $\{X_t\}_{t\in T}$ is a family of random variables indexed by time t (see e.g. Hamilton, 1994). In this thesis, we focus on discrete-time stochastic processes, where the index set $T \subset \mathbb{Z}$ represents discrete time points at which observations are recorded. Discrete-time models are particularly convenient for financial applications because data are often collected at regular intervals such as daily, weekly or monthly. However, this regularity is an idealization, in practice, trading calendars are influenced by weekends, holidays and other market closures, which introduce irregularities in the observation times. Despite these limitations, discrete-time models remain widely used due to their tractability and compatibility with standard econometric techniques, especially for volatility modeling using GARCH-type frameworks (see e.g. Tsay, 2010).

2.1.3 Stationarity

Stationarity is a fundamental concept in time series analysis, ensuring that the statistical properties of a process remain stable over time. A stochastic process $\{X_t\}_{t\in\mathbb{Z}}$ is said to be strictly stationary if the joint distribution of the process remains invariant under time shifts. In the univariate case, for any positive integer h and any selection of time points $t_1, t_2, \dots, t_k \in \mathbb{Z}$, the condition $(X_{t_1}, X_{t_2}, \dots, X_{t_k}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$ holds, where $\stackrel{d}{=}$ denotes equality in distribution. In the multivariate case, where each observation $X_t \in \mathbb{R}^m$, the joint distribution of the vector-valued process $(X_{t_1}, X_{t_2}, \dots, X_{t_k}) \in \mathbb{R}^{km}$ must also remain unchanged under shifts in time. It is important to note that strict stationarity is a distributional property and does not imply the existence of any moments such as the mean or variance. In empirical modeling, it is common practice to search for transformations that render the data strictly stationary, as the statistical model is defined under this assumption. After estimating the model, one must verify that the parameter values indicate no violation of stationarity conditions.

Since strict stationarity is difficult to verify and does not concern moment conditions, a weaker version known as weak stationarity (or covariance stationarity) is often assumed. The concept of weak stationarity plays a central role in time series analysis. A stochastic process $\{X_t\}$ is said to be weakly stationary if its mean is constant over time, its variance is finite and time-invariant and its autocovariance function $\gamma(h) = \mathrm{Cov}(X_t, X_{t+h})$ depends only on the lag $h \in \mathbb{Z}$ and not on the specific time point $t \in \mathbb{Z}$. In the univariate setting, this means that $\mathbb{E}[X_t] = \mu$ and $\mathrm{Var}(X_t) = \mathbb{E}[(X_t - \mu)^2] = \sigma^2$ for all $t \in \mathbb{Z}$, while $\gamma(h) = \mathbb{E}[(X_t - \mu)(X_{t+h} - \mu)]$ holds for all $h \in \mathbb{Z}$. In the multivariate setting, if $X_t \in \mathbb{R}^m$, weak stationarity requires that the mean vector $\mathbb{E}[X_t]$ be constant, the covariance matrix $\mathrm{Cov}(X_t)$ remain finite and time-invariant and the cross-covariance matrices $\Gamma(h) = \mathrm{Cov}(X_t, X_{t+h}) \in \mathbb{R}^{m\times m}$ depend only on the lag h. Weak stationarity provides the foundation for defining and estimating conditional variance models such as GARCH, as it guarantees that the second-order properties of the process remain stable over time.

The assumption of stationarity is indispensable in time series modeling because without it, the statistical model cannot be properly defined. This applies to commonly used models such as the autoregressive moving average (ARMA) model and the generalized autoregressive conditional heteroskedasticity (GARCH) model. If the observed series is non-stationary, estimation techniques may yield misleading or spurious results and valid inference becomes impossible. Therefore, the modeller must search for an appropriate transformation that renders the series stationary, for example by differencing the observed data when non-stationarity arises in the mean. Tests such as the Augmented Dickey-Fuller (ADF) test can be used to detect the presence of a unit root, which is a specific type of non-stationarity. However, this form of non-stationarity is not applicable in the context of GARCH models, where non-stationarity arises in the conditional variance and not due to unit roots. Understanding both strict and weak forms of stationarity, in both univariate and multivariate contexts, is essential for constructing reliable time series models and conducting valid statistical inference.

2.1.4 White Noise

A white noise process is a fundamental building block in time series modeling and plays a crucial role in defining the error structures of models such as ARMA and GARCH. In the univariate case, a stochastic process $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is said to be (weak) white noise if it satisfies $\mathbb{E}[\varepsilon_t] = 0$, $\mathrm{Var}(\varepsilon_t) = \sigma^2 < \infty$, and $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t \neq s$ (see e.g., Nelson, 1991). This definition implies that the process consists of uncorrelated random variables with constant mean and variance over time. In contrast, strong (or strict) white noise requires that the random variables ε_t are independent and identically distributed (iid), meaning that in addition to being uncorrelated, all joint distributions factorize and higher-order dependencies are absent. In the multivariate setting, let $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ be a sequence of m-dimensional random vectors, $\varepsilon_t \in \mathbb{R}^m$. The process is said to be multivariate white noise if $\mathbb{E}[\varepsilon_t] = 0$, $\mathrm{Cov}(\varepsilon_t) = \Sigma_\varepsilon < \infty$, and $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t \neq s$, where Σ_ε denotes the constant, finite covariance matrix of the process. If, in addition, $\{\varepsilon_t\}$ consists of iid random vectors, the process is called multivariate strong white noise. This distinction is essential in multivariate model specifications, particularly in models such as vector autoregressions (VAR) and multivariate GARCH, where assumptions about the innovation process influence the validity of parameter estimates and inference procedures.

2.1.5 Conditional mean and covariance

In time series analysis, the conditional mean and conditional covariance are fundamental concepts for characterizing the dependence structure of stochastic processes. Let $\{\boldsymbol{X}_t\}_{t\in\mathbb{Z}}$ be an $\mathbb{R}^m$ valued stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_{t-1}$ denote the sigma-algebra generated by the history of the process up to time $t-1$, i.e., $\mathcal{F}_{t-1} = \sigma(\boldsymbol{X}_{t-1}, \boldsymbol{X}_{t-2}, \ldots)$. The conditional mean of $\boldsymbol{X}_t$, denoted by $\mathbb{E}[\boldsymbol{X}_t \mid \mathcal{F}_{t-1}]$, is the best mean-square predictor of $\boldsymbol{X}_t$ given past information. In the univariate case, for a real-valued process $\{X_t\}$, the conditional mean $\mathbb{E}[X_t \mid \mathcal{F}_{t-1}]$ exists if $X_t \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, i.e., if $\mathbb{E}[|X_t|] < \infty$. This conditional expectation captures the dynamic behavior of the process and is used extensively in modeling frameworks such as autoregressive and moving average processes (see e.g., Hamilton, 1994; Tsay, 2005). The conditional covariance of two real-valued random variables X_t and Y_t, given $\mathcal{F}_{t-1}$, is defined as

$$\mathrm{Cov}(X_t, Y_t \mid \mathcal{F}_{t-1}) = \mathbb{E}\left[(X_t - \mathbb{E}[X_t \mid \mathcal{F}_{t-1}])(Y_t - \mathbb{E}[Y_t \mid \mathcal{F}_{t-1}]) \mid \mathcal{F}_{t-1}\right]. \quad (2.4)$$

In the multivariate setting, the conditional covariance matrix of $\boldsymbol{X}_t$ given $\mathcal{F}_{t-1}$ is defined as $\boldsymbol{H}_t = \mathrm{Cov}(\boldsymbol{X}_t \mid \mathcal{F}_{t-1})$, which is assumed to be a symmetric positive definite $m \times m$ matrix. The (i, j)-th element of $\boldsymbol{H}_t$ is given by $\mathrm{Cov}(X_{it}, X_{jt} \mid \mathcal{F}_{t-1})$, where X_{it} and X_{jt} denote the i-th and j-th components of $\boldsymbol{X}_t$. Positive definiteness ensures the existence of $\boldsymbol{H}_t^{-1}$, which is essential in likelihood based inference and in defining derived quantities. The conditional correlation matrix, denoted by $\boldsymbol{R}_t$, captures the pairwise conditional correlations among components of $\boldsymbol{X}_t$ and is obtained by standardizing $\boldsymbol{H}_t$ as

$$\boldsymbol{R}_t = [\mathrm{diag}(\boldsymbol{H}_t)]^{-1/2} \boldsymbol{H}_t [\mathrm{diag}(\boldsymbol{H}_t)]^{-1/2}, \quad (2.5)$$

where $\mathrm{diag}(\boldsymbol{H}_t)$ denotes the diagonal matrix consisting of the variances of each component of $\boldsymbol{X}_t$. These conditional second order properties play a central role in univariate and multivariate GARCH type models, where time varying volatility and correlation structures are modeled explicitly.

2.1.6 Likelihood Function

The likelihood function is a key concept in statistical inference, both in the frequentist and Bayesian frameworks. In the context of time series models, it plays a central role in parameter estimation, model evaluation and hypothesis testing. The likelihood function is derived from the joint distribution of the observed data, reflecting how likely the observed outcomes are given a set of parameters. Let $y_1, y_2, \ldots, y_T$ denote a univariate or multivariate time series, where each y_t is an m-dimensional vector representing the observation at time t, i.e., $y_t \in \mathbb{R}^m$. Here, T represents the total number of observations.

Suppose the data are generated from a parametric model with an unknown parameter vector $\theta \in \Theta \subset \mathbb{R}^d$, where Θ denotes the parameter space and d is the dimension of the parameter vector. The joint probability density function of $y_1, y_2, \ldots, y_T$ conditional on the initial values F_0, is denoted by

$$f(y_1, y_2, \ldots, y_T \mid \theta, F_0).$$

This joint density reflects the probability distribution of the entire sequence of observations $y_1, \ldots, y_T$ given the parameter vector θ and the initial information F_0. The likelihood function, which is the probability of observing the data given the model parameters is obtained by taking the joint density

$$L(\theta) = f(y_1, y_2, \ldots, y_T \mid \theta, F_0). \tag{2.6}$$

The initial values F_0 correspond to the presample observations required to initialize the recursion in models such as GARCH. These initial values can be treated as fixed constants, known values or drawn from a stationary distribution, depending on the estimation strategy employed. The specific strategy used for initializing the model and estimating the parameters will be discussed in detail in the empirical section of this thesis. If the data-generating process satisfies a Markov property or if we can factor the joint density recursively, the likelihood function can be expressed as a product of conditional densities. The Markov property here implies that the conditional distribution of y_t, given the entire past $\{y_{t-1}, y_{t-2}, \ldots\}$, depends only on a finite subset, typically the most recent observations. This assumption allows us to write the joint density function of $(y_1, y_2, \ldots, y_T)$ as

$$f(y_1, y_2, \ldots, y_T \mid \theta) = \prod_{t=1}^{T} f(y_t \mid \mathcal{F}_{t-1}, \theta), \tag{2.7}$$

where $\mathcal{F}_{t-1} = \sigma(y_{t-1}, y_{t-2}, \ldots)$ is the sigma-algebra generated by past values up to time $t-1$ and $f(\cdot \mid \mathcal{F}_{t-1}, \theta)$ denotes the conditional density of y_t given past information and parameters. This factorization follows from the chain rule of probability and holds under the assumption of conditional independence given $\mathcal{F}_{t-1}$ (see e.g., Meitz and Saikkonen, 2016). The maximum likelihood estimator (MLE) is obtained by maximizing the log-likelihood function

$$\ell_T(\theta) = \sum_{t=1}^{T} \log f(y_t \mid \mathcal{F}_{t-1}, \theta). \tag{2.8}$$

Asymptotic Properties of the MLE

Under suitable regularity conditions (see e.g., Teräsvirta, Tjøstheim, and Granger, 2010; Comte and Lieberman, 2003), the MLE $\hat{\theta} \in \mathbb{R}^d$, where $\theta \in \Theta \subset \mathbb{R}^d$ is the parameter vector, is consistent and asymptotically normal. Specifically, we have

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}_d(0, I^{-1}(\theta_0)), \tag{2.9}$$

where $\theta_0 \in \mathbb{R}^d$ is the true parameter vector and $I(\theta_0)$ is the $d \times d$ Fisher information matrix, defined as

$$I(\theta_0) = \mathbb{E}\left[-\frac{\partial^2 \ell_T(\theta)}{\partial\theta\partial\theta'}\bigg|_{\theta=\theta_0} \right]. \tag{2.10}$$

In the context of time series models, this setup must often be extended to account for conditional likelihoods, where the conditioning on past observations is essential for model specification. The information matrix in such cases is evaluated conditionally, and inference is based on the conditional distribution of the data (see e.g., Hamilton, 1994). This asymptotic distribution justifies the use of the inverse of the observed or expected Fisher information matrix to estimate the covariance matrix of $\hat{\theta}$. The standard errors of the parameter estimates are calculated as the square roots of the diagonal elements of this estimated covariance matrix. These standard errors are essential for constructing confidence intervals and conducting hypothesis testing. Thus, a clear understanding of the likelihood function and its asymptotic properties is crucial for statistical inference in time series models. The estimation process reflects the underlying data structure and ensures that the model is both statistically sound and practically informative (See e.g. ; Tsay, 2010).

2.1.7 Model Selection Criteria: AIC, BIC, HQC and FPE

Selecting an appropriate model is crucial to ensuring accurate predictions and meaningful interpretations. Several statistical criteria help determine the best model by balancing goodness-of-fit and model complexity. Among the most widely used are the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Hannan-Quinn Criterion (HQC) and the Final Prediction Error (FPE) criterion introduced by Shibata. These criteria assess models by penalizing excessive parameterization to avoid overfitting while ensuring a good fit to the data (See e.g. ; Tsay, 2010).

Akaike Information Criterion (AIC) The Akaike Information Criterion (AIC), introduced by Akaike (1973), is a widely used method for comparing statistical models by balancing goodness-of-fit with model complexity. It is defined as

$$\text{AIC} = -2 \log L(\hat{\theta}) + 2k, \tag{2.11}$$

where $\log L(\hat{\theta})$ is the log-likelihood evaluated at the maximum likelihood estimator $\hat{\theta}$ and k is the total number of estimated parameters in the model. This formulation penalizes models with more parameters to avoid overfitting. A lower AIC value indicates a better trade-off between model fit and simplicity. However, AIC does not include a penalty that increases with the sample size. As a result, it is inconsistent, meaning that it does not guarantee selection of the true model even as the sample size $T \to \infty$. In contrast, other criteria such as the Bayesian Information Criterion (BIC) and the Hannan–Quinn Criterion (HQC) introduce sample size dependent penalties and are consistent under general conditions. These criteria are often preferred in theoretical model selection settings where consistency is desirable (see e.g., Tsay, 2010).

Bayesian Information Criterion (BIC) The Bayesian Information Criterion (BIC), also known as the Schwarz criterion (Schwarz, 1978), extends model selection principles by incorporating a penalty term that increases with sample size. It is defined as

$$\text{BIC} = -2 \log L(\hat{\theta}) + k \log T, \tag{2.12}$$

where $\log L(\hat{\theta})$ is the log-likelihood evaluated at the maximum likelihood estimator $\hat{\theta}$, k is the number of estimated parameters and T is the sample size. The penalty term $k \log T$ increases with T, which discourages overfitting in larger samples more strongly than AIC. BIC is consistent, meaning that under regularity conditions, it selects the true model with probability approaching one as $T \to \infty$. This makes it more reliable than AIC in large sample scenarios where the goal is to identify the correct model structure (see e.g., Tsay, 2010).

Hannan-Quinn Criterion (HQC) The Hannan–Quinn Criterion (HQC), introduced by Hannan and Quinn (1979), offers a middle ground between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) in terms of penalizing model complexity. It is defined as

$$\text{HQC} = -2 \log L(\hat{\theta}) + 2k \log(\log T), \tag{2.13}$$

where $\log L(\hat{\theta})$ is the log likelihood evaluated at the maximum likelihood estimator $\hat{\theta}$, k is the number of estimated parameters and T is the sample size. The penalty term $2k \log(\log T)$ grows with the sample size but at a slower rate than the BIC's $k \log T$. Unlike AIC, which applies a constant penalty and BIC, which increases the penalty more strongly, HQC's slower growing penalty allows it to balance model fit and parsimony effectively. HQC is also consistent, meaning it tends to select the correct model as $T \to \infty$ and it has been shown to perform well in finite samples (see e.g., Tsay, 2010).

Final Prediction Error Criterion (FPE) The criterion introduced by Shibata (1976), originally known as the Final Prediction Error (FPE) criterion, applies AIC to the order selection of autoregressive models and is particularly relevant for forecasting accuracy. Unlike BIC, which aims for consistent model identification, the FPE criterion prioritizes predictive performance. It is asymptotically equivalent to AIC and tends to select models that minimize the mean squared prediction error. While this approach may lead to overfitting in terms of parameter count, it often improves short term forecast performance (see e.g., Tsay, 2010). Each model selection criterion has its advantages depending on the modeling objective. AIC and the FPE criterion are well suited for short term forecasting due to their focus on predictive accuracy, even at the risk of overfitting. BIC, with its stronger penalty, is more appropriate when the goal is to identify the true data generating process. HQC offers a balanced alternative, penalizing complexity less aggressively than BIC but still ensuring model consistency. In empirical applications, it is common practice to consider multiple criteria to make a robust model selection decision (see e.g., Tsay, 2010).

2.1.8 Residuals

Residuals play a crucial role in statistical modeling, as they help evaluate whether the assumptions underlying a statistical model are valid, i.e., whether the model is correctly specified. A residual is the difference between the observed value and the corresponding predicted value from the model. Formally, if y_t represents the observed time series and $\hat{y}_t$ is the fitted value, the residual at time t is given by

$$e_t = y_t - \hat{y}_t. \tag{2.14}$$

A well specified model should produce residuals that resemble strong white noise, a process that is uncorrelated, has constant variance, a mean close to zero and follows a specified distribution (see e.g., Tsay, 2010). This is distinct from weak white noise, which only requires zero mean, constant variance and lack of autocorrelation, without any distributional assumption. The strong white noise assumption is particularly critical in likelihood based inference, as estimation and hypothesis testing procedures rely on the specified distribution (often Gaussian). If the assumed distribution is incorrect, the inference may be invalid and the model misleading. Diagnostic checks on residuals are essential for model validation. In univariate time series models, the autocorrelation function (ACF) and partial autocorrelation function (PACF) of residuals are used to assess remaining temporal dependencies. The Ljung–Box test is often applied to examine whether autocorrelations in residuals are jointly significant over several lags (see e.g., Tsay, 2010). However, this test is limited to univariate series. To test for conditional heteroskedasticity, visual inspection of squared residuals and formal procedures such as the ARCH-LM test are used. If volatility clustering is detected, a GARCH type model may be more suitable. In GARCH modeling, assessing the adequacy of the conditional variance specification is important. This involves analyzing standardized residuals, defined as $\hat{\varepsilon}_t = e_t/\hat{\sigma}_t$, where $\hat{\sigma}_t$ is the estimated conditional standard deviation. These standardized residuals should again resemble strong white noise.

A common diagnostic for conditional variance misspecification is to analyze the ACF and PACF of squared standardized residuals and to apply the Ljung–Box test to those series. The appropriate reference for this test is McLeod and Li (1983), who proposed diagnostic checking of ARMA models using autocorrelations of squared residuals. In multivariate settings, such as VAR or multivariate GARCH models, residual diagnostics are extended to vector valued processes. For autocorrelation, multivariate portmanteau tests such as the multivariate Ljung–Box test or the test of Hosking (1980) can be employed. These tests generalize univariate autocorrelation diagnostics to the multivariate case. To detect heteroskedasticity in multivariate residuals, multivariate ARCH-LM type tests can also be employed. Note that Hosking's (1980) test may also be interpreted as a test for heteroskedasticity in multivariate systems. Assessing the distributional assumption in multivariate models involves testing whether the standardized residual vectors follow a multivariate normal distribution. Formal tests for this purpose include the Doornik–Hansen test (see e.g., Doornik and Hansen, 2008) and the Henze–Zirkler test (see e.g., Henze and Zirkler, 1990). Additionally, graphical tools such as Q–Q plots of the marginal distributions of residuals and Mahalanobis distance plots of the residual vectors help evaluate deviations from multivariate normality (see e.g., Tsay, 2010). Accurate modeling in both univariate and multivariate contexts depends critically on thorough residual diagnostics. These checks ensure the reliability of estimation and inference and offer guidance for model refinement when mis-specification is detected.

2.1.9 Volatality clustering

Volatility clustering is a well documented phenomenon in financial time series, where large changes in asset returns, whether positive or negative are often followed by large changes, while small changes tend to be followed by small changes. This suggests that financial markets experience alternating periods of turbulence and stability rather than exhibiting constant variability (see e.g., Tsay, 2010). The concept was first observed by Mandelbrot (1963), who noted that financial returns exhibit periods of persistent high or low volatility rather than being independent over time (201; Mandelbrot, 1963).

Mathematically, volatility clustering can be described with other models, including time-varying volatility models such as GARCH or stochastic volatility models. Volatility clustering is commonly attributed to the behavior of investors and market participants. According to Cont (2007), this phenomenon arises due to the gradual diffusion of information in financial markets, where news impacts trading behavior over time rather than instantaneously (310; Cont, 2007). Additionally, factors such as investor herding and feedback trading contribute to clustering effects by causing prolonged market fluctuations. This serves as a preliminary explanation, then leads to one mathematical definition. Recall, there are other options to describe and model volatility clustering, than the autoregressive conditional heteroskedasticity (ARCH) models and its extensions like GARCH. Empirical studies show that volatility clustering is present across various asset classes and time periods. It is a crucial property in risk management and financial modeling, as it affects option pricing, value-at-risk calculations and portfolio optimization. Since volatility is predictable to some extent, accounting for its persistence allows for more accurate forecasting and risk assessment in financial markets (see e.g.; Andersen et al., 2003).

2.1.10 Student's *t*-distribution

In the context of GARCH type models, it is common to assume that the error term follows a standardized Student's t -distribution instead of the normal distribution. This choice allows for fat tails, which are frequently observed in financial return series. The standardized t-distribution with ν degrees of freedom has mean zero and variance equal to one when $\nu > 2$. The probability density function (pdf) of a standardized t-distributed random variable ε_t is given by

$$f(\varepsilon_t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}} \left(1 + \frac{\varepsilon_t^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad \nu > 2, \tag{2.15}$$

where $\Gamma(\cdot)$ denotes the gamma function (see e.g., Meitz & Saikkonen, 2016). If the model for the conditional mean is $y_t = \mu_t + \sigma_t \varepsilon_t$ and ε_t follows the standardized t-distribution above, then the conditional density of y_t given the information set $\mathcal{F}_{t-1}$ is

$$f(y_t \mid \mathcal{F}_{t-1}) = \frac{1}{\sigma_t} \cdot \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu-2)\pi}} \left[1 + \frac{1}{\nu-2}\left(\frac{y_t-\mu_t}{\sigma_t}\right)^2\right]^{-\frac{\nu+1}{2}}. \tag{2.16}$$

The corresponding log-likelihood function for a sample of T observations is

$$\ell_T(\theta) = \sum_{t=1}^{T} \Bigg[\log \Gamma\left(\frac{\nu+1}{2}\right) - \log \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log\left((\nu-2)\pi\right) \\ - \log \sigma_t - \frac{\nu+1}{2}\log\left(1 + \frac{(y_t-\mu_t)^2}{(\nu-2)\sigma_t^2}\right)\Bigg], \qquad (2.17)$$

where θ denotes the vector of parameters in the model, including the additional shape parameter ν (see e.g., Meitz & Saikkonen, 2016). The shape parameter ν is estimated in both univariate and multivariate GARCH type models. The parameter ν captures the heaviness of the tails and allows the models to accommodate excess kurtosis observed in financial return data. Its inclusion enables a more accurate specification of the error distribution in the presence of conditional heteroskedasticity.

2.2 General Aspects of Models of Conditional Variance

Suppose y_t is a weakly and strictly stationary process for which $\mathbb{E}_{t-1}(y_t) = 0$ and therefore $\mathbb{E}(y_t) = 0$. A general model for the conditional variance of y_t can be written as

$$y_t = h_t^{1/2}\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0,1), \qquad (2.18)$$

where h_t denotes the conditional variance of y_t given past observations and ε_t is assumed to be independently and identically distributed with zero mean and unit variance. It is further assumed that ε_t is independent of the past values of y_t, meaning that it is orthogonal to the information set available at time $t-1$. This assumption implies that h_t and ε_t are independent of each other. Given this model, the conditional expectation of y_t is

$$\mathbb{E}_{t-1}(y_t) = \mathbb{E}_{t-1}(h_t^{1/2}\varepsilon_t) = h_t^{1/2}\mathbb{E}_{t-1}(\varepsilon_t) = h_t^{1/2}\mathbb{E}(\varepsilon_t) = 0.$$

The first equality follows from the model specification. The second uses the fact that h_t is measurable with respect to the information set $\mathcal{F}_{t-1}$ and can be treated as constant under the conditional expectation. The third equality follows from the assumption that ε_t is independent of the past and hence $\mathbb{E}_{t-1}(\varepsilon_t) = \mathbb{E}(\varepsilon_t)$. Since ε_t is assumed to have zero mean, the final result is zero. To compute the conditional variance of y_t, we use the definition

$$\text{Var}_{t-1}(y_t) = \mathbb{E}_{t-1}(y_t^2) - [\mathbb{E}_{t-1}(y_t)]^2 = \mathbb{E}_{t-1}(y_t^2).$$

since $\mathbb{E}_{t-1}(y_t) = 0$. Substituting from the model,

$$\mathbb{E}_{t-1}(y_t^2) = \mathbb{E}_{t-1}(h_t\varepsilon_t^2) = h_t\mathbb{E}_{t-1}(\varepsilon_t^2) = h_t\mathbb{E}(\varepsilon_t^2) = h_t.$$

Thus, the conditional variance of y_t is h_t and the term $h_t^{1/2}$ represents the conditional standard deviation of y_t, which is commonly referred to as volatility. Because y_t is assumed to be stationary, its unconditional second moment $\mathbb{E}(y_t^2) = \text{Var}(y_t)$ is constant over time. From the model and the independence between h_t and ε_t, it follows that

$$\text{Var}(y_t) = \mathbb{E}(y_t^2) = \mathbb{E}(h_t \varepsilon_t^2) = \mathbb{E}(h_t)\mathbb{E}(\varepsilon_t^2) = \mathbb{E}(h_t),$$

as $\mathbb{E}(\varepsilon_t^2) = 1$. Therefore, for the process y_t to be stationary, the conditional variance process h_t must also be stationary (see e.g., Meitz & Saikkonen, 2016). Finally, under the assumptions made, the process y_t is uncorrelated over time. That is, for all $k \neq 0$,

$$\text{Cov}(y_t, y_{t-k}) = 0,$$

because ε_t is independent across time and h_t is a function of past values, which ensures no linear dependence between y_t and its lagged values.

2.3 Autoregressive Conditional Heteroskedasticity (ARCH) Model

The Autoregressive Conditional Heteroskedasticity (ARCH) model, introduced by (307; Engle, 1982) captures time varying volatility in financial time series by modeling the conditional variance of the error term as a function of past squared observations. This modeling approach was motivated by the empirical observation that financial asset returns exhibit volatility clustering. The ARCH(s) model is defined by the equations

$$y_t = \mu + h_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } (0, 1), \tag{2.19}$$

$$h_t = \omega + \sum_{i=1}^{s} \alpha_i (y_{t-i} - \mu)^2, \tag{2.20}$$

where μ denotes the conditional mean of the process, h_t is the conditional variance and the parameters satisfy $\omega > 0$ and $\alpha_i \geq 0$ for all i, ensuring a non-negative variance process (see e.g., Meitz and Saikkonen, 2016). A sufficient condition for weak stationarity of the ARCH process is

$$\sum_{i=1}^{s} \alpha_i < 1,$$

which guarantees that the unconditional variance of y_t is finite. While weak stationarity ensures that the first and second moments of the process are time invariant, strict stationarity requires that the full distribution of y_t remains constant over time. For ARCH processes, strict stationarity can also be established under appropriate conditions on the coefficients and the distribution of the innovation process (see e.g., Francq and Zakoïan, 2010).

2.4 Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by (308; Bollerslev, 1986), extends the ARCH framework by including lagged conditional variances in addition to past squared observations. This structure provides a more parsimonious and flexible approach to modeling volatility clustering in financial time series. The GARCH(p, s) model is defined as

$$y_t = \mu + h_t^{1/2}\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0,1), \tag{2.21}$$

$$h_t = \omega + \sum_{i=1}^{s} \alpha_i (y_{t-i} - \mu)^2 + \sum_{j=1}^{p} \beta_j h_{t-j}, \tag{2.22}$$

where μ denotes the conditional mean of the return process and h_t is the conditional variance. The parameters satisfy $\omega > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$ to ensure that the variance is positive and well-defined (see e.g. Meitz and Saikkonen, 2016). The inclusion of lagged variances h_{t-j} distinguishes the GARCH model from the ARCH model, allowing it to capture long range volatility persistence. A sufficient condition for weak stationarity of the GARCH(p, s) process is

$$\sum_{i=1}^{s} \alpha_i + \sum_{j=1}^{p} \beta_j < 1,$$

which ensures that the unconditional variance of y_t exists and is finite (see e.g. Meitz and Saikkonen, 2016). In addition to weak stationarity, strict stationarity of GARCH models can be established under more technical conditions. Among various specifications, the GARCH(1,1) model is the most widely used in empirical finance due to its simplicity and effectiveness in capturing volatility dynamics (see e.g. Tsay, 2010).

2.5 Exponential GARCH (EGARCH) Model

The Exponential GARCH (EGARCH) model, introduced by (315; Nelson, 1991) addresses certain limitations of standard GARCH models by allowing for asymmetric effects in volatility dynamics. Unlike symmetric GARCH models, which assume that positive and negative shocks of equal magnitude have the same effect on volatility, EGARCH models incorporate leverage effects i.e., the impact of a shock on volatility depends on its sign and magnitude. The direction of this asymmetry is governed by the sign of the parameter γ. The EGARCH(1,1) model consists of the following conditional mean and variance equations

$$y_t = \mu + h_t^{1/2}\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0,1) \tag{2.23}$$

$$\log h_t = \omega + \beta \log h_{t-1} + \alpha \left| \frac{y_{t-1} - \mu}{\sqrt{h_{t-1}}} \right| + \gamma \frac{y_{t-1} - \mu}{\sqrt{h_{t-1}}}, \tag{2.24}$$

Here, h_t denotes the conditional variance and μ is the conditional mean. The parameter β captures the persistence of volatility, α measures the magnitude effect of standardized shocks and γ represents the asymmetry or leverage effect. A key advantage of the EGARCH model is that the logarithmic specification of h_t ensures its positivity without imposing non negativity constraints on the parameters.

2.6 Maximum Likelihood Estimation for ARCH, GARCH and EGARCH Models

Maximum Likelihood Estimation (MLE) is commonly used to estimate the parameters of volatility models such as ARCH, GARCH and EGARCH. The general approach assumes that the error terms ε_t follow a specific distribution, typically the normal or t-distribution. MLE is used to estimate the parameters that maximize the likelihood of observing the given data based on the conditional distribution of the returns. For all three models, the log-likelihood function is generally expressed as

$$\ell_T(\theta) = \sum_{t=1}^{T} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(h_t) - \frac{(y_t - \mu)^2}{2h_t} \right], \tag{2.25}$$

where, y_t is the observed value at time t, μ is the conditional mean of the process, h_t is the conditional variance at time t and θ represents the parameters to be estimated, including μ, ω, α_i, β_j, γ and other model specific parameters. The parameters of the model, including μ, ω, α_i, β_j and γ (for the EGARCH model) are estimated by maximizing this log-likelihood function. In practice, the true distribution of the errors ε_t may deviate from normality and in such cases, the MLE framework can still be applied by using Quasi-Maximum Likelihood Estimation (QMLE). QMLE is an alternative estimation method that is particularly useful when the distribution of the error terms is unknown or not normal. Unlike basic MLE, QMLE involves using a pseudo-likelihood function based on the assumption that the distribution of the error term does not need to be fully specified but can still be treated as if it were. Under regularity conditions, QMLE provides consistent and asymptotically normal estimates of the model parameters, even if the distribution of ε_t deviates from normality (see e.g., Bollerslev, 1986; Engle, 1982). Another important consideration in MLE is the choice of initial values for the parameters. Since the likelihood function is often non-linear and may have multiple local maxima, selecting appropriate starting values is crucial for the optimization process. A common practice is to use estimates from simpler models, such as the sample mean for μ and the sample variance for ω, as initial values.

2.7 Autoregressive Exponential Generalized Autoregressive Conditional Heteroskedasticity (AR(1)-EGARCH(1,1)) Model

The AR(1)-EGARCH(1,1) model integrates an autoregressive process for the conditional mean with an EGARCH(1,1) specification for the conditional variance. This combined structure is particularly useful in financial time series modeling, as asset returns often display both autocorrelation and volatility clustering (See e.g. ; Bollerslev, 1986). We assume the error process ε_t follows the conditional distribution $\varepsilon_t \sim$ i.i.d.$(0, h_t)$, where h_t is the conditional variance. The return process then satisfies the AR(1) specification

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t. \tag{2.26}$$

Here, μ is a constant and ϕ captures the autoregressive dependence. Under the stationarity condition $|\phi| < 1$, the unconditional mean of y_t is given by $\mathbb{E}(y_t) = \mu/(1-\phi)$ (See e.g. ; Meitz and Saikkonen, 2016), which is obtained as a function of the parameters μ and ϕ. The conditional variance is modeled using the EGARCH(1,1) specification

$$\log h_t = \omega + \beta \log h_{t-1} + \alpha \left| \frac{y_{t-1} - \mu}{\sqrt{h_{t-1}}} \right| + \gamma \frac{y_{t-1} - \mu}{\sqrt{h_{t-1}}}. \tag{2.27}$$

This formulation captures both the magnitude effect of past standardized shocks (via α) and their asymmetry (via γ). The sign of γ determines whether positive or negative shocks have a greater effect on volatility and the logarithmic form ensures positivity of h_t without non-negativity constraints on the parameters. Assuming conditional normality, the log-likelihood function for the AR(1)-EGARCH(1,1) model over T observations is given by

$$\ell_T(\theta) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\log h_t + \frac{(y_t - \mu - \phi y_{t-1})^2}{h_t}\right), \tag{2.28}$$

where $\theta = (\mu, \phi, \omega, \alpha, \beta, \gamma)$ denotes the full parameter vector.

2.8 Multivariate Volatility Models

Early research in financial econometrics primarily focused on modeling the volatility of individual assets. However, financial markets are interconnected and changes in volatility often spill over across multiple assets. This interdependence is particularly evident during periods of financial distress when market fluctuations in one asset class or geographic region can significantly impact others (210; Engle, 2002). Understanding these relationships is crucial for financial applications such as asset pricing, risk management and portfolio allocation. Multivariate volatility models extend univariate approaches to account for these dynamic relationships. The Multivariate GARCH (MGARCH) family of models allows for the estimation of time-varying covariances and correlations, which are essential for capturing systemic risk in financial markets (See e.g ; Tsay, 2010). One of the earliest models in this category is the Constant Conditional Correlation (CCC) model introduced by (90; Bollerslev, 1990), which assumes that correlations between assets remain constant over time. However, this assumption is often unrealistic in practice, as financial market correlations tend to vary dynamically in response to economic and market conditions (142; Engle and Sheppard, 2001).

To address this limitation, the Dynamic Conditional Correlation (DCC) model was developed, allowing correlations to evolve over time (143; Engle, 2002). The DCC model is widely used in portfolio optimization since correlations between assets determine the extent to which diversification can reduce portfolio risk. In financial theory, a portfolio consists of a collection of financial assets and diversification aims to reduce total risk by combining assets with less than perfect correlation (189; Markowitz, 1952). By capturing time-varying correlations, the DCC model enhances the ability of investors to construct well-diversified portfolios and optimize their risk-return trade-offs. Another widely used model in multivariate volatility analysis is the BEKK-GARCH model, which provides a more flexible structure for modeling volatility spillovers across assets (220; Engle and Kroner, 1995). Unlike DCC models, BEKK does not impose a predetermined correlation structure, allowing for a more general representation of interdependencies. However, this flexibility comes at the cost of increased computational complexity, making it less practical for large-scale financial applications (See e.g. ; Tsay, 2010).

The choice of a multivariate volatility model depends on the specific research question and computational feasibility. While BEKK-GARCH is useful for examining detailed volatility transmission mechanisms, the DCC and ADCC models provide a more parsimonious approach suitable for large datasets. Given the importance of capturing both volatility clustering and correlation dynamics, this thesis focuses on time-varying correlation models such as DCC and BEKK. These models offer practical advantages in empirical finance by enabling the estimation of evolving market relationships and improving risk management strategies. Multivariate volatility modeling remains a vital area of research in financial econometrics. By extending univariate models to account for interdependencies, MGARCH models provide a robust framework for analyzing systemic risk, portfolio diversification and financial contagion effects. These models contribute to our understanding of how volatility propagates through financial markets, enhancing decision making for both investors and policymakers.

2.8.1 Constant Conditional Correlation (CCC) Model

The Constant Conditional Correlation (CCC) model, introduced by (See e.g. ; Bollerslev, 1990) extends the univariate GARCH framework to a multivariate setting by modeling time-varying conditional variances while assuming that correlations between assets remain constant. This approach simplifies the estimation of the covariance structure in financial markets, making it computationally efficient (See e.g. ; Tsay, 2010). As already explained, in the univariate case, the conditional variance of a single time series $\{y_t\}$ is given by

$$y_t = h_t^{1/2}\varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0,1), \tag{2.29}$$

where h_t represents the conditional variance of y_t given past observations. This formulation captures the heteroskedastic nature of financial time series. In the multivariate extension, the vector of asset returns, denoted as $\mathbf{y}_t$ (a $d \times 1$ vector), follows the process

$$\mathbf{y}_t = H_t^{1/2}\boldsymbol{\varepsilon}_t, \tag{2.30}$$

as shown in Teräsvirta, Tjøstheim and Granger (2010, p. 202, Eq. 8.65), where $\mathbf{H}_t$ is a $d \times d$ conditional variance-covariance matrix and $\boldsymbol{\varepsilon}_t$ is an i.i.d. process of zero-mean $d \times 1$ vectors with covariance matrix I_d, the $d \times d$ identity matrix. As explained in Teräsvirta, Tjøstheim and Granger (2010), in the CCC model, the structure of $\mathbf{H}_t$ is specified as

$$\mathbf{H}_t = D_t R D_t, \tag{2.31}$$

where R is a $d \times d$ constant correlation matrix containing the pairwise correlations between assets and D_t is a diagonal matrix with the square roots of individual conditional variances on its diagonal. The (i, j)-element of the matrix $\mathbf{H}_t$, denoted $H_{t,ij}$, is given by

$$H_{t,ij} = \sqrt{h_{t,i}}\sqrt{h_{t,j}}\rho_{ij}, \tag{2.32}$$

where ρ_{ij} is the (i, j)-element of the constant correlation matrix R and $h_{t,i}$ and $h_{t,j}$ are the conditional variances of assets i and j, respectively. Each individual conditional variance $h_{t,i}$ is modeled using a univariate GARCH(p, s) process of the form

$$h_{t,i} = c_i + \sum_{j=1}^{s} \alpha_{i,j} y_{t-j,i}^2 + \sum_{k=1}^{p} \beta_{i,k} h_{t-k,i}, \tag{2.33}$$

where $c_i > 0$, $\alpha_{i,j} \geq 0$ and $\beta_{i,k} \geq 0$ for all i, j and k. The indices of $\alpha_{i,j}$ and $\beta_{i,k}$ emphasize that each asset may have its own GARCH parameters, accounting for asset-specific volatility dynamics. Estimation of the CCC model typically follows a two-step procedure. In the first step, univariate GARCH models are estimated separately for each asset using the maximum likelihood estimation (MLE) theory explained earlier. This provides the time-varying conditional variances $h_{t,i}$ for each asset. In the second step, the constant correlation matrix R is estimated based on the standardized residuals

$$z_{t,i} = \frac{y_{t,i}}{\sqrt{h_{t,i}}}, \tag{2.34}$$

and the correlation coefficient ρ_{ij} is computed as

$$\rho_{ij} = \frac{1}{T}\sum_{t=1}^{T} z_{t,i}z_{t,j}, \tag{2.35}$$

where T is the total number of observations. This estimator for R uses the outer product of standardized residuals, ensuring symmetry and unit diagonals. The full log-likelihood function for the CCC model is provided in Teräsvirta, Tjøstheim and Granger (2010, p. 206, Eq. 8.72) and can be written as

$$\ell_T(\theta) = -\frac{dT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} \log|\mathbf{H}_t| + \mathbf{y}_t'\mathbf{H}_t^{-1}\mathbf{y}_t, \tag{2.36}$$

where $\mathbf{y}_t$ is the vector of returns and $\mathbf{H}_t$ is constructed according to equation (2.31). Although the CCC model offers computational advantages, its main limitation is the assumption of constant correlations. Empirical evidence shows that asset correlations often vary over time, particularly during financial crises (Cappiello et al., 2006).

2.8.2 Dynamic Conditional Correlation (DCC) Model

The Dynamic Conditional Correlation (DCC) model, introduced by (200; Engle and Sheppard, 2001), extends the Constant Conditional Correlation (CCC) model by allowing correlations between assets to vary over time. Unlike the CCC model, which assumes constant correlations, the DCC model captures the dynamic nature of financial markets, making it more suitable for modeling evolving relationships between asset returns. As in the CCC model, the conditional covariance matrix $\mathbf{H}_t$ in the DCC model is expressed as

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{P}_t\mathbf{D}_t, \tag{2.37}$$

where $\mathbf{D}_t$ is a diagonal matrix of time-varying standard deviations obtained from univariate GARCH models and $\mathbf{P}_t$ is the time-varying conditional correlation matrix. The marginal conditional variances $h_{t,i}$ are modeled using univariate GARCH processes for $i = 1, \ldots, d$, based on the maximum likelihood estimation (MLE) theory discussed earlier. Here, d denotes the number of assets under consideration. The standardized residuals used to estimate the dynamic correlations are defined as

$$z_{t,i} = \frac{y_{t,i}}{\sqrt{h_{t,i}}}, \tag{2.38}$$

where $y_{t,i}$ is the return of asset i at time t and $h_{t,i}$ is its conditional variance. These standardized residuals are used to estimate the correlation matrix $\mathbf{P}_t$ (See e.g. ; Tsay, 2010). The estimation of $\mathbf{P}_t$ involves modeling the dynamics of conditional correlations using a GARCH-type process applied to the outer products of the standardized residuals. According to Teräsvirta, Tjøstheim and Granger (2010), different estimation strategies can be considered for estimating the parameters of the DCC-GARCH model. The parameters of the individual univariate GARCH equations that define $\mathbf{D}_t$ can be estimated separately by assuming an identity matrix for the correlation component, i.e., $\mathbf{P}_t = \mathbf{I}_d$. Once the univariate models are estimated, the standardized residuals $\mathbf{z}_t$ are used to estimate the parameters governing the time-varying correlations. Specifically, the dynamic process for the correlation matrix is given by

$$\mathbf{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + a\mathbf{z}_{t-1}\mathbf{z}'_{t-1} + b\mathbf{Q}_{t-1}, \tag{2.39}$$

$$\mathbf{P}_t = \operatorname{diag}(\mathbf{Q}_t)^{-1/2}\mathbf{Q}_t \operatorname{diag}(\mathbf{Q}_t)^{-1/2}, \tag{2.40}$$

where a and b are scalar parameters that determine the dynamics of the conditional correlation process and $\bar{\mathbf{Q}}$ is the unconditional covariance matrix of the standardized residuals $\mathbf{z}_t$. The matrix $\mathbf{Q}_t$ is a positive definite matrix that evolves over time and captures the dynamic dependence structure, while $\mathbf{P}_t$ is obtained by normalizing $\mathbf{Q}_t$ to have ones on the diagonal, thereby forming a proper correlation matrix. The parameters a and b are constrained such that $a \geq 0$, $b \geq 0$ and $a + b < 1$ to ensure the positive definiteness and stationarity of the process. These are additional parameters introduced in the DCC model to govern the short-run persistence and decay rate of past shocks to the conditional correlation. A high value of a implies that recent shocks have a significant impact on correlation dynamics, while a high value of b indicates a strong persistence of past conditional correlations. Under suitable regularity conditions, the two-step estimation procedure yields consistent, though generally inefficient, parameter estimates. Engle (2002) also proposed simplified approaches for estimating the correlation parameters more efficiently. Compared to fully parameterized multivariate GARCH models such as BEKK, the DCC model significantly reduces computational complexity while maintaining flexibility, making it particularly useful for modeling high-dimensional financial datasets (215; Cappiello, Engle, and Sheppard, 2006).

2.8.3 BEKK-GARCH Model

The BEKK-GARCH model, introduced by (15; Baba et al., 1991), is a multivariate extension of the GARCH framework that flexibly models the conditional covariance matrix of asset returns. Unlike the Constant Conditional Correlation (CCC) and Dynamic Conditional Correlation (DCC) models, which impose specific correlation structures, the BEKK formulation allows for more complex volatility dynamics and interdependencies, making it particularly useful for analyzing volatility spillovers and cross-market effects (88; Engle and Kroner, 1995). The general multivariate GARCH model poses challenges for ensuring the positive definiteness of the conditional covariance matrix. The BEKK model addresses this by structuring the conditional covariance matrix $\mathbf{H}_t$ in a way that guarantees positive definiteness by construction. The BEKK-GARCH model with a single lag ($K = 1$) is specified as

$$\mathbf{H}_t = \mathbf{C}\mathbf{C}' + \mathbf{A}'\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\mathbf{A} + \mathbf{B}'\mathbf{H}_{t-1}\mathbf{B}, \tag{2.41}$$

where $\mathbf{H}_t$ is the conditional covariance matrix of the vector of asset returns $\mathbf{y}_t$, $\mathbf{C}$ is a constant lower triangular matrix, and $\mathbf{A}$ and $\mathbf{B}$ are constant square parameter matrices that capture the influence of past returns and past covariances, respectively. This form can be viewed as a restricted version of the general VEC model and provides a parsimonious yet flexible representation of time-varying covariances. The log-likelihood function for the BEKK-GARCH model is analogous to that of the CCC-GARCH model and is given by

$$\ell_T(\theta) = -\frac{mT}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log\det\mathbf{H}_t(\theta) - \frac{1}{2}\sum_{t=1}^{T}\mathbf{y}_t'\mathbf{H}_t^{-1}(\theta)\mathbf{y}_t, \tag{2.42}$$

where θ denotes the vector of model parameters, including the unique elements of $\mathbf{C}$, $\mathbf{A}$ and $\mathbf{B}$ and T is the sample size. Estimation of θ is typically performed using the Quasi-Maximum Likelihood (QML) method. While the structure of the BEKK model ensures positive definiteness of $\mathbf{H}_t$, it introduces computational complexity. As noted by (88; Engle and Kroner, 1995), computing analytical derivatives of the log-likelihood is challenging due to the need for recursive differentiation. They recommended using numerical derivatives and optimization via the BHHH (Berndt–Hall–Hall–Hausman) algorithm. However, (102; Hafner and Herwartz, 2008) derived analytical expressions for the first and second derivatives of $\mathbf{H}_t$ in the BEKK model with $K = 1$, making exact gradient-based estimation feasible and improving efficiency and accuracy.

Additionally, (103; Comte and Lieberman, 2003) established the asymptotic properties of the QML estimators for the BEKK model under standard regularity conditions, including consistency and asymptotic normality. Despite its flexibility, the BEKK model can be computationally demanding due to the number of parameters, especially in higher dimensions. Nevertheless, it remains widely used in empirical finance for analyzing volatility transmission and interdependence across markets.

2.8.4 Factor GARCH Models

Factor GARCH models offer an alternative approach to modeling multivariate volatility by reducing dimensionality through latent factors. These models assume that the volatilities and covariances of multiple assets are driven by a smaller number of unobservable risk factors, simplifying estimation while still capturing the essential dynamics of market risk (120; Engle, 2002). This makes Factor GARCH models particularly useful for large financial datasets, where fully parameterized models like BEKK become computationally infeasible. Instead of estimating all individual asset covariances directly, Factor GARCH models express the conditional covariance matrix as a function of a few common risk factors. These factors represent shared sources of volatility that affect multiple assets, improving computational efficiency while maintaining interpretability (see e.g. ; Tsay, 2010). The estimation process is typically carried out using numerical methods such as maximum likelihood estimation (MLE) or the generalized method of moments (GMM) (130; Cappiello et al., 2006). Model selection is usually guided by information criteria such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC).

One potential advantage of Factor GARCH models is their capacity to provide economic interpretations, as the estimated factors can often be interpreted as corresponding to fundamental drivers of market risk (135; Engle and Sheppard, 2001). This interpretability makes them particularly valuable in portfolio risk management, where identifying common sources of volatility aids in constructing diversified portfolios. However, a major challenge in using these models is correctly specifying the number of latent factors, as an incorrect choice can lead to biased estimates and model misspecification (See e.g. ; Tsay, 2010). Additionally, Factor GARCH models may struggle to capture asset-specific risks that are not driven by common factors, necessitating extensions such as nonlinear or multivariate factor GARCH variants to improve flexibility and accuracy in capturing volatility dynamics (145; Engle, 2002). While both BEKK and Factor GARCH models provide valuable insights into multivariate volatility modeling, their application depends on the research objective. BEKK is better suited for detailed analysis of volatility transmission and spillovers, while Factor GARCH models are more practical for high-dimensional problems where dimensionality reduction is necessary for efficient computation.

2.9 Motivation for Model Selection

This thesis investigates the dynamics of financial market volatility using both univariate and multivariate GARCH type models. The primary objective is to assess how well different models capture volatility clustering, asymmetry and time-varying correlations in the returns of major U.S. stocks. The models discussed so far are not presented merely for completeness, they form the methodological backbone for the empirical analysis carried out in the later sections of this work. Their inclusion is motivated by the need to evaluate the performance of both traditional and more flexible multivariate specifications such as DCC and BEKK in capturing interconnected risks. By comparing model performance, the thesis aims to identify which frameworks provide the most accurate and interpretable insights for risk modeling and portfolio management.

3. Data

3.1 Data Description

This study utilizes a dataset of monthly stock returns for General Electric (GE) and IBM, spanning from 1926 to 2008. The data is obtained from Tsay (2020), which originally sources financial market data from the Center for Research in Security Prices (CRSP) at the University of Chicago Booth School of Business (300; CRSP, 2021). The dataset is used to analyze volatility dynamics and correlation structures between these two assets. Data processing and statistical computations were conducted using R, a widely used statistical computing environment (R Core Team, 2021). Specifically, the `rugarch` package (Ghalanos, 2022) was used for univariate GARCH modeling, while the `rmgarch` package (Ghalanos, 2022) was employed for multivariate GARCH modeling, enabling the estimation of dynamic conditional correlations and other key properties of volatility.

To provide an initial understanding of the dataset, Figure 3.1 presents the time series plot of GE and IBM monthly returns. The figure illustrates fluctuations in returns over time, highlighting periods of increased volatility, which often coincide with major economic downturns. One prominent example is the Great Depression of the 1930s, triggered by the 1929 stock market crash, which led to prolonged economic contraction, high unemployment, and extreme financial instability (301; Romer, 2003). During this period, stock returns exhibited severe volatility, with rapid declines and temporary recoveries reflecting the uncertainty in financial markets. Another notable period of heightened volatility is the 2008 Global Financial Crisis, caused by the collapse of the housing market and the failure of major financial institutions, leading to a sharp decline in stock prices worldwide (302; Reinhart & Rogoff, 2009). Similar patterns of increased market turbulence can also be observed during other recessions, such as the 1973–75 oil crisis recession (303; Hamilton, 1983) and the dot-com bubble burst of the early 2000s (304; Shiller, 2005). These financial crises contributed to drastic fluctuations in asset returns, which are evident in the time series plot.

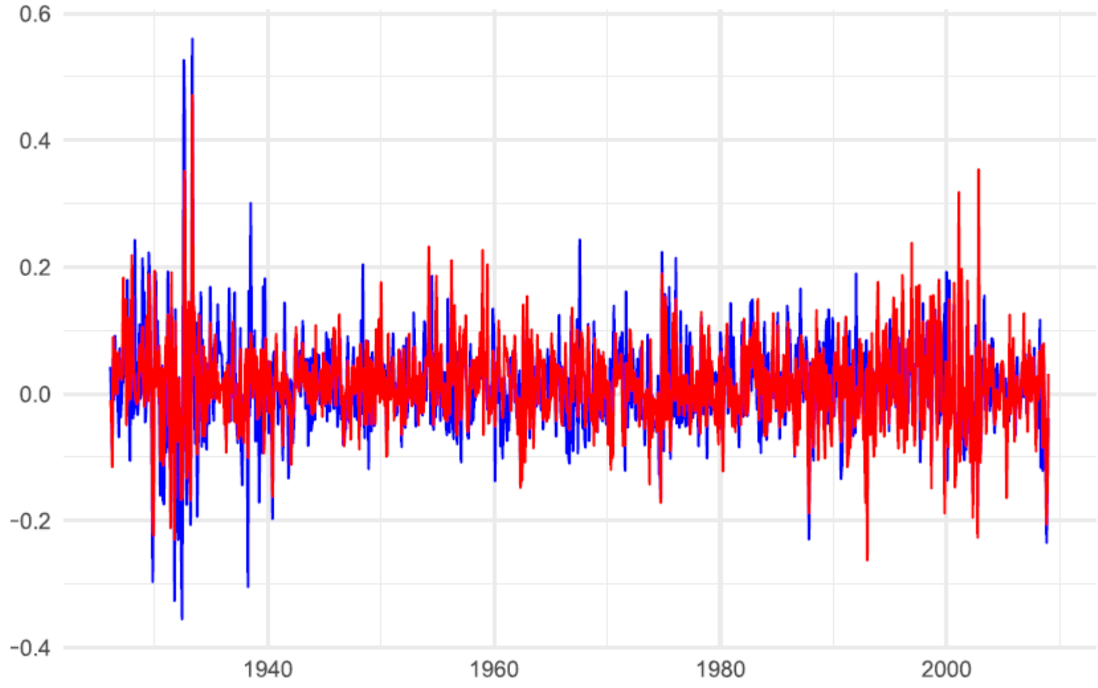


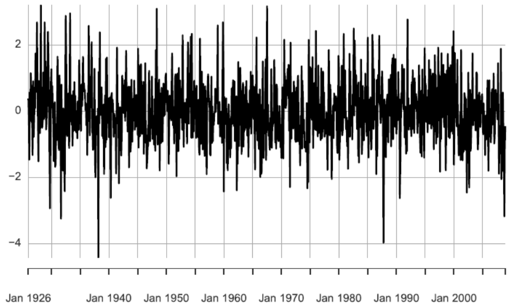
Figure 3.1: The figure presents the time series of monthly stock returns for GE in blue and IBM in red from 1926 to 2008.

Table 3.1: Estimated Parameters for GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on GE Monthly Stock Returns. Estimation was carried out using Maximum Likelihood Estimation (MLE).

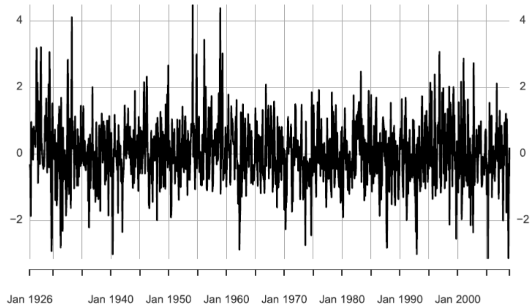
Parameter	GARCH(1,1)	EGARCH(1,1)	AR(1)-EGARCH(1,1)
μ	0.013 (0.002)	0.011 (0.002)	0.011 (0.002)
ω	0.0002 (0.00006)	-0.222 (0.072)	-0.218 (0.068)
α_1	0.112 (0.019)	-0.051 (0.024)	-0.053 (0.024)
β_1	0.863 (0.020)	0.958 (0.013)	0.958 (0.013)
γ_1	—	0.228 (0.038)	0.225 (0.037)
ϕ	—	—	0.182 (0.057)
ν	—	9.331 (2.564)	9.382 (2.593)

Table 3.2: Log-likelihood, AIC, BIC for GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on GE Monthly Stock Returns.

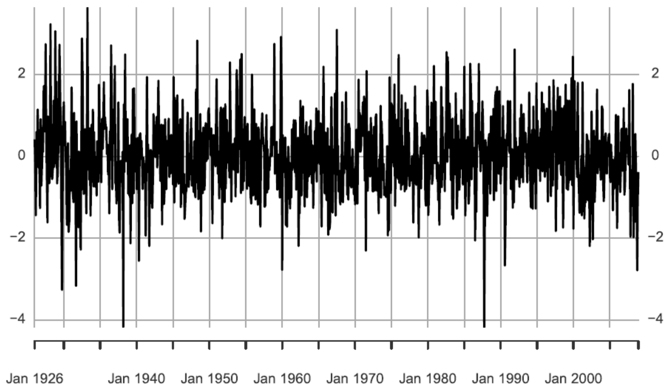
Criterion	GARCH(1,1)	EGARCH(1,1)	AR(1)-EGARCH(1,1)
Log-Likelihood	1252	1262	1264
AIC	-2.51	-2.52	-2.52
BIC	-2.49	-2.49	-2.49
Hannan-Quinn (HQIC)	-2.50	-2.51	-2.51
Shibata	-2.51	-2.52	-2.52



GARCH(1,1)



EGARCH(1,1)



AR(1)-EGARCH(1,1)

Figure 3.2: Time Series Plots of Standardized Residuals from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on GE Monthly Stock Returns.

3.2 Univariate GE models

GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models were estimated to analyze the volatility dynamics of GE stock returns. The primary distinctions among these models lie in how they account for volatility persistence, asymmetry, and the conditional mean structure. GARCH(1,1) assumes symmetric responses of volatility to shocks (308; Bollerslev, 1986), whereas EGARCH(1,1) incorporates asymmetry, allowing negative shocks to affect volatility differently from positive ones (325; Nelson, 1991). AR(1)-EGARCH(1,1) extends this by including an autoregressive term in the conditional mean equation to capture potential serial correlation in the return series. When comparing the estimated parameters (see Table 3.1), GARCH(1,1) reveals high volatility persistence with $\hat{\alpha}_1 + \hat{\beta}_1 \approx 0.97$, which still implies weak stationarity. The estimated EGARCH(1,1) model indicates some asymmetry, though the value of $\hat{\gamma}_1$ is moderate (approximately 0.23). The AR(1)-EGARCH(1,1) model also displays evidence of asymmetry and strong persistence ($\hat{\beta}_1 = 0.95$), along with a statistically significant AR(1) term ($\hat{\phi} = 0.19$). In the GE return series, this result suggests that past returns influence current returns both in the conditional mean and in the conditional variance.

Regarding model selection, the AR(1)-EGARCH(1,1) model has the highest log-likelihood (1266) and the lowest AIC (-2.53), BIC (-2.50), and HQIC (-2.52) among the three, indicating a marginally superior fit (see Table 3.2). To evaluate model adequacy, residual diagnostics are essential. The time series plots of standardized residuals (see Figure 3.2) show that all three models produce residuals that appear stationary and free from strong patterns. Notably, the residuals from the AR(1)-EGARCH(1,1) model exhibit slightly reduced volatility clustering, indicating a better representation of the dynamics in the return series.

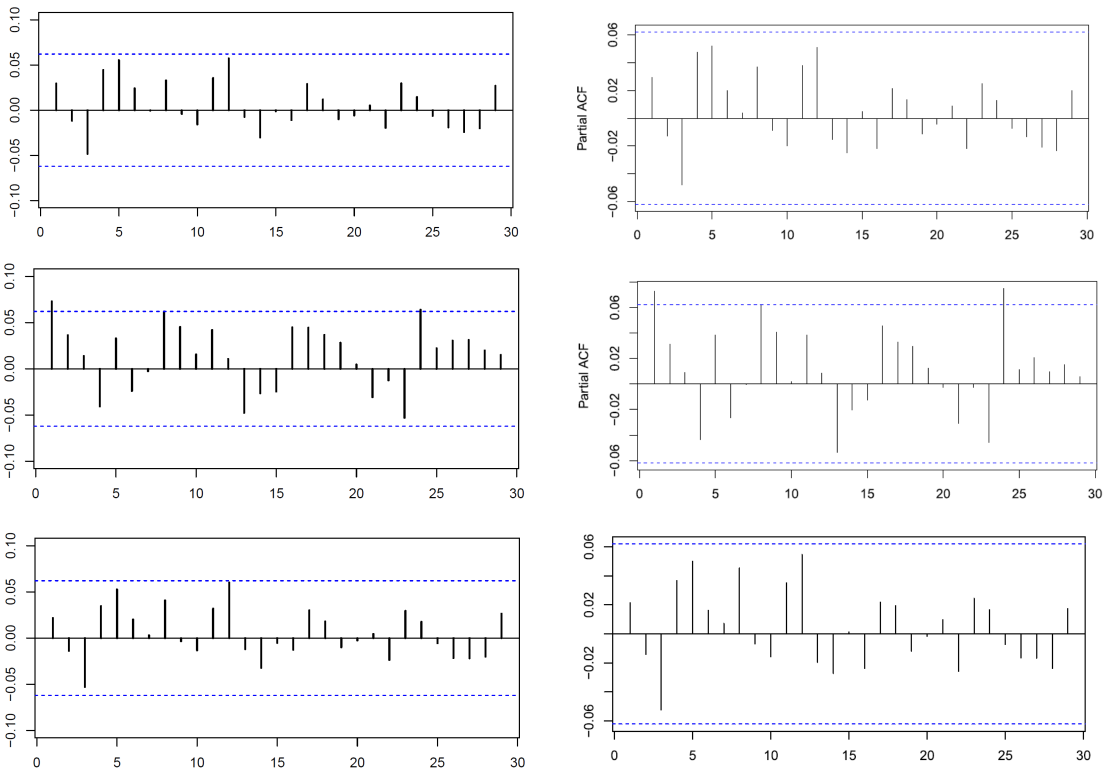


Figure 3.3: Autocorrelation (ACF) and Partial Autocorrelation (PACF) Plots of Standardized Residuals from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on GE Monthly Stock Returns. The first row shows GARCH(1,1), the second row EGARCH(1,1) and the third row AR(1)-EGARCH(1,1), with ACF on the left and PACF on the right within each row. The dashed blue lines represent approximate 95% pointwise confidence intervals, calculated as $\pm 1.96/\sqrt{T}$, where T is the sample size.

The ACF and PACF plots of standardized residuals in Figure 3.3 provide important diagnostic tools for assessing model adequacy. In the case of the GARCH(1,1) model, both ACF and PACF plots show no significant autocorrelation, indicating that the model sufficiently captures the conditional heteroskedasticity present in the GE return series. However, for the EGARCH(1,1) model, the PACF plot reveals a significant spike at lag 1 that exceeds the 95% pointwise confidence interval. While this might suggest some residual autocorrelation, it is important to note that these bounds are only valid for each lag individually. Since each plot include 30 lags, making joint inference based solely on pointwise intervals can be misleading particularly because the probability of observing at least one significant spike under the null increases with the number of lags. This limitation highlights the importance of complementing visual inspection with formal diagnostic tests.

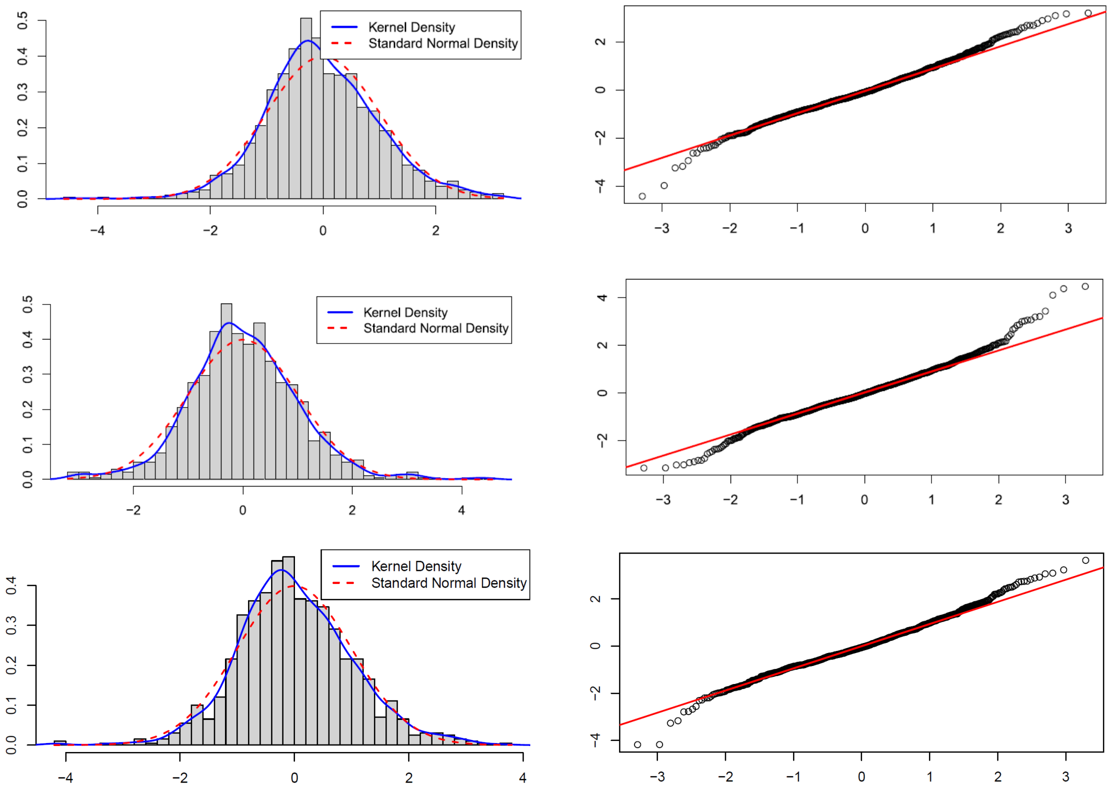


Figure 3.4: Histogram and QQ Plots of Standardized Residuals from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on GE Monthly Stock Returns. Each row corresponds to a specific model. GARCH(1,1) (top), EGARCH(1,1) (middle) and AR(1)-EGARCH(1,1) (bottom).

To address this issue, an AR(1) term was incorporated into the conditional mean equation, resulting in the AR(1)-EGARCH(1,1) model. This modification successfully eliminates the spike at lag 1 in the PACF plot of the standardized residuals. The absence of significant autocorrelations in both ACF and PACF confirms that the AR(1)-EGARCH(1,1) model better captures the dynamics of the return process and ensures that the residuals behave like white noise. Therefore, incorporating an AR term improves the model's specification by addressing remaining linear dependencies in the return series. Although the histogram and Q-Q plot of standardized residuals from the AR(1)-EGARCH(1,1) model (Figure 3.4) show improved fit around the center of the distribution compared to the GARCH and EGARCH models, heavy tails and asymmetry are still evident particularly in the upper tail. Q-Q plots assess the normality of residuals, while histograms visualize their distribution. This indicates that the normality assumption does not fit the data, a common limitation in financial return modeling (317; Nelson, 1991).

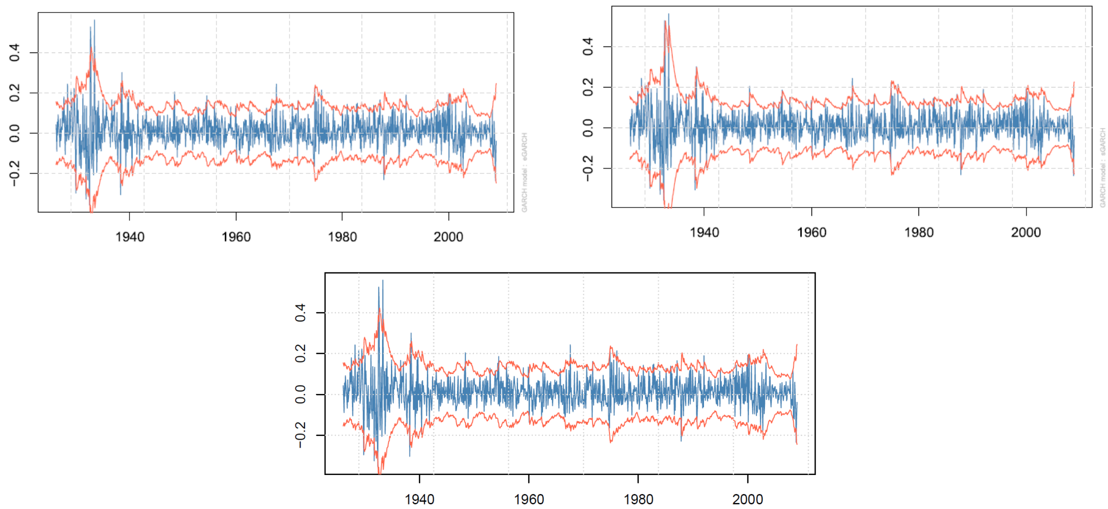


Figure 3.5: Conditional standard deviation plots estimated from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models based on GE monthly stock returns. Each plot shows the evolution of conditional volatility over time. The blue line represents the estimated conditional standard deviation and the red lines correspond to the ± 2 standard error bands, calculated using the conditional variance output from each model under the assumption of conditional normality (317; Nelson, 1991).

The estimation of all models in Table 3.1 was based on the normal likelihood function, which assumes conditional normality of the residuals. The evidence of non-normality observed in diagnostic plots suggests that a more flexible distribution, such as the Student's - t, could provide a better fit by accommodating the heavy tails commonly observed in financial time series. Volatility clustering is effectively captured by all three models GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) as illustrated in the conditional standard deviation plots in Figure 3.5. These plots exhibit alternating periods of high and low volatility, where large shocks are followed by further large shocks (of either sign) and small shocks tend to follow small shocks. The GARCH(1,1) model provides a baseline representation of time-varying volatility, while the EGARCH(1,1) model captures asymmetric effects by allowing for the possibility that negative and positive shocks impact volatility differently. The AR(1)-EGARCH(1,1) model further improves the specification by accounting for autocorrelation in the return series, resulting in smoother volatility dynamics and a better fit to the observed data.

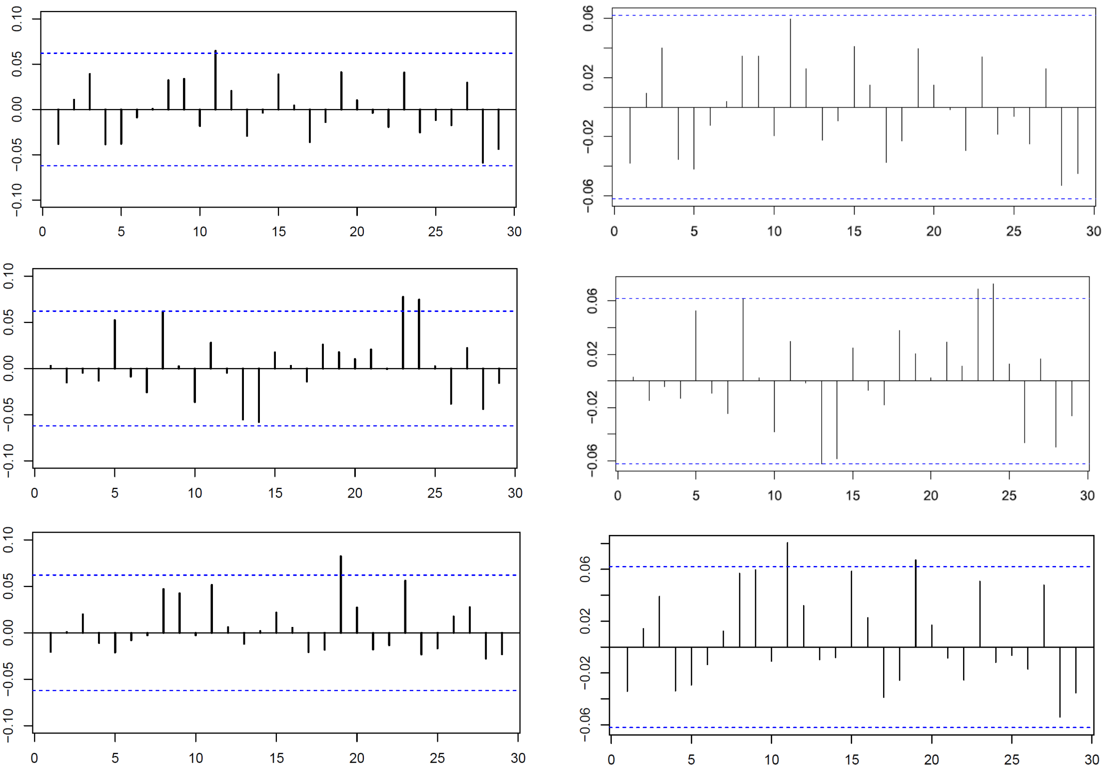


Figure 3.6: Autocorrelation (ACF) and Partial Autocorrelation (PACF) plots of squared standardized residuals from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models based on GE monthly stock returns. The first row presents the ACF (left) and PACF (right) for the GARCH(1.1) model, the second for the EGARCH(1,1) model and the third for the AR(1)-EGARCH(1,1) model. The blue dashed lines represent the 95% pointwise confidence bounds.

The adequacy of each volatility model was further assessed by examining the autocorrelation (ACF) and partial autocorrelation (PACF) plots of the squared standardized residuals as shown in Figure 3.6. Ideally, for a well-specified GARCH-type model, there should be no significant autocorrelation left in the squared standardized residuals, that is, all autocorrelation and partial autocorrelation spikes should lie within the 95% confidence bounds (317; Tsay, 2010). In the case of the GARCH(1,1) model, both the ACF and PACF plots show no significant spikes beyond lag 1, indicating that the model has successfully captured the second-order dependencies in the data. The absence of autocorrelation in the squared standardized residuals suggests a reasonably good fit (317; Tsay, 2010).

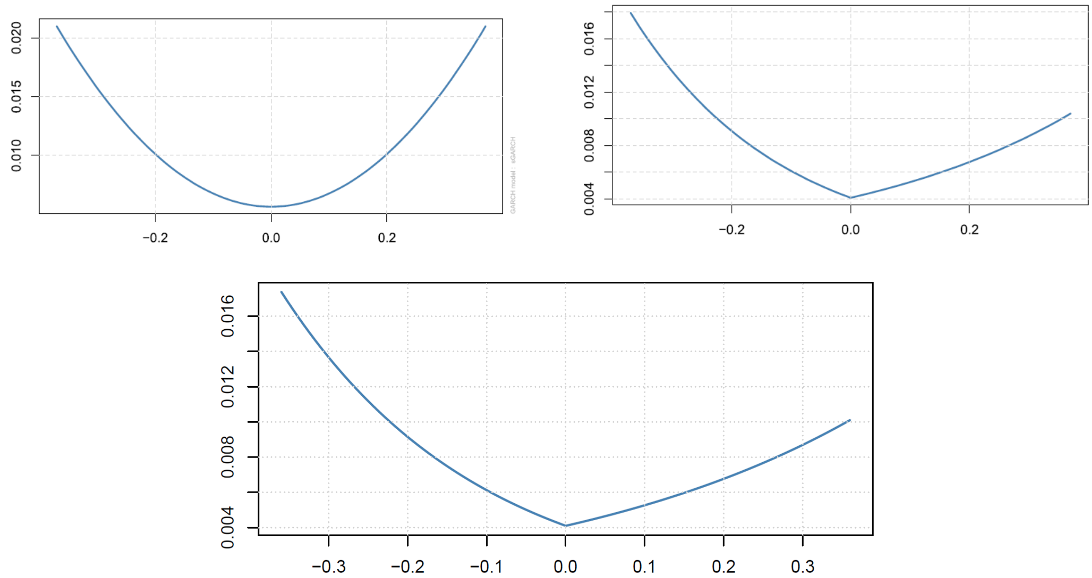


Figure 3.7: News impact curves from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models based on GE monthly stock returns. The top row displays the GARCH and EGARCH models, while the bottom row presents the AR(1)-EGARCH(1,1) model.

For the EGARCH(1,1) model, while most autocorrelations lie within the confidence bands, notable spikes appear in the PACF at lags 23 and 24, slightly exceeding the 95% pointwise confidence intervals. These deviations may raise concerns about higher-order effects in the volatility dynamics. However, considering that approximately 30 lags are examined, observing one or two spikes outside the confidence bounds is not unexpected even under the null hypothesis of no autocorrelation. The 95% pointwise confidence intervals imply that each individual lag has a 5% chance of exceeding the bounds by random variation alone. Therefore, occasional exceedances do not necessarily imply model misspecification. This understanding supports the adequacy of the EGARCH(1,1) model, although it may slightly underfit certain long-memory or periodic features in the volatility process (317; Nelson, 1991). The PACF plot of the AR(1)-EGARCH(1,1) model shows a mild spike at lag 12, suggesting the possible presence of calendar or seasonal effects not explicitly modeled (317; Glosten, Jagannathan, and Runkle, 1993). Still, overall residual diagnostics support the model's suitability. Both ACF and PACF plots show minimal autocorrelation, and the AR(1)-EGARCH(1,1) model achieves the highest log-likelihood and the most favorable information criteria among the three candidate models. Thus, despite the isolated spike at lag 12, the AR(1)-EGARCH(1,1) model can be considered well specified.

Figure 3.7 illustrates an important property of volatility models the response of conditional volatility to new information, also known as the news impact curve. The GARCH(1,1) model displays a symmetric response, meaning that positive and negative shocks of equal size have the same effect on volatility. In contrast, the EGARCH(1,1) model captures asymmetry: negative shocks lead to larger increases in volatility than positive ones. This is consistent with empirical findings in financial markets, where adverse news tends to elicit stronger volatility reactions (317; Nelson, 1991). The AR(1)-EGARCH(1,1) model also accommodates asymmetric effects by incorporating an indicator function that differentiates the impact of positive and negative shocks. As shown in Figure 3.7, the news impact curve lies above the origin for negative innovations, confirming the presence of a leverage effect. While all three models GARCH(1,1), EGARCH(1,1), and AR(1)-EGARCH(1,1) successfully capture volatility clustering, the asymmetric models offer important improvements. EGARCH(1,1) incorporates asymmetry through a logarithmic specification of conditional variance (317; Nelson, 1991), while the AR(1)-EGARCH(1,1) model achieves it using an indicator function that allows different responses to positive and negative shocks (317; Glosten, Jagannathan, and Runkle, 1993). Among the models considered, the AR(1)-EGARCH(1,1) provides the best fit. It demonstrates the strongest performance based on log-likelihood and information criteria, the lowest residual autocorrelation, and clear evidence of leverage effects confirmed by both residual diagnostics and the news impact curve.

3.3 Univariate IBM Models

GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models were estimated for IBM stock returns. Since the general methodology and most of the graphical patterns are similar to those observed for GE, this section highlights only the key differences across the three models when applied to IBM. Particular attention is given to volatility asymmetry, persistence, model fit and residual diagnostics to evaluate how well each model captures the characteristics of the IBM return series.

Table 3.3: Estimated Parameters for GARCH(1,1), EGARCH(1,1), and AR(1)-EGARCH(1,1) Models Based on IBM Monthly Stock Returns. Estimation was carried out using Maximum Likelihood Estimation (MLE). The standard errors (in parentheses) are calculated using the asymptotic covariance matrix of the MLE, obtained directly from the Hessian matrix evaluated at the parameter estimates.

Parameter	**GARCH(1,1)**	**EGARCH(1,1)**	**AR(1)-EGARCH(1,1)**
μ	0.014 (0.0020)	0.012 (0.0019)	0.012 (0.0020)
ω	0.000 (0.0001)	−0.288 (0.1069)	−0.275 (0.1007)
α_1	0.111 (0.0261)	−0.060 (0.0268)	−0.061 (0.0270)
β_1	0.825 (0.0403)	0.947 (0.0198)	0.949 (0.0186)
γ_1	—	0.210 (0.0421)	0.207 (0.0413)
ν	—	6.753 (1.3599)	6.903 (1.4093)
ϕ	—	—	0.050 (0.0342)

Table 3.4: Log-likelihood and Information Criteria for GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) Models Based on IBM Monthly Stock Returns.

Criterion	**GARCH(1,1)**	**EGARCH(1,1)**	**AR(1)-EGARCH(1,1)**
Log-Likelihood	1277	1301	1303
AIC	-2.56	-2.60	-2.60
BIC	-2.54	-2.57	-2.57
Hannan-Quinn	—	—	-2.59
Shibata	—	—	-2.60

Table 3.3 presents the estimated parameters for the GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models, with standard errors reported in parentheses. Table 3.4 summarizes the model selection criteria, including log-likelihood, AIC and BIC values. Based on these statistics, both asymmetric models outperform the symmetric GARCH(1,1) model. Among them, AR(1)-EGARCH(1,1) shows the highest log-likelihood and the lowest AIC, indicating the best overall fit. Volatility persistence remains high across all models. In GARCH(1,1), the sum of the ARCH and GARCH coefficients is approximately $\hat{\alpha}_1+\hat{\beta}_1 \approx 0.94$, indicating long memory in volatility. Both asymmetric models capture the leverage effect through a significant $\hat{\gamma}_1$ parameter. For EGARCH(1,1), $\hat{\gamma}_1 = 0.209$ and for AR(1)-EGARCH(1,1), $\hat{\gamma}_1 = 0.207$, confirming the presence of asymmetric volatility behavior in IBM stock returns (317; Nelson, 1991). These findings are consistent with those observed in the GE analysis, further supporting the use of models that account for asymmetry in financial return volatility.

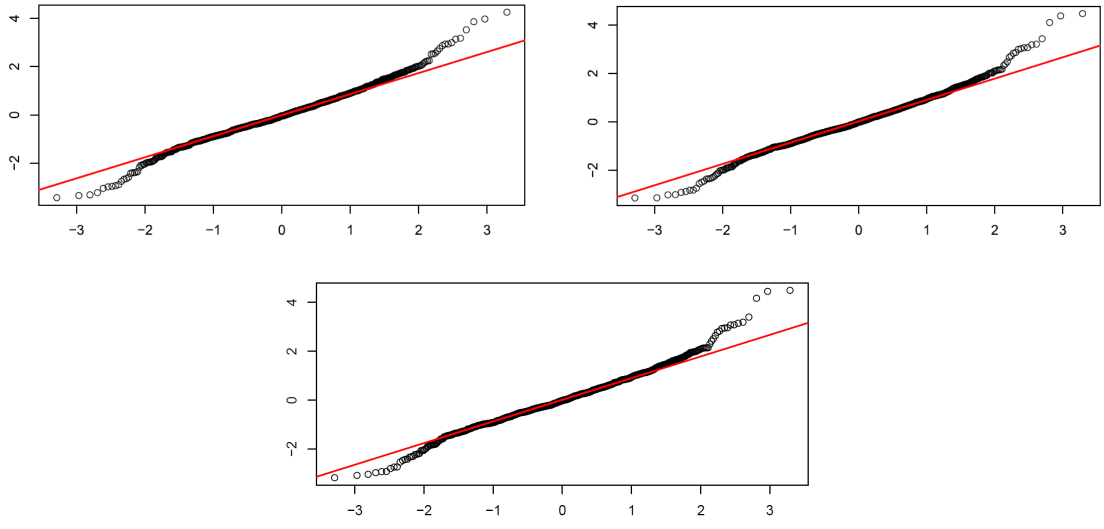


Figure 3.8: QQ plots of standardized residuals from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models based on IBM monthly stock returns. The top row shows the QQ plots for GARCH(1,1) and EGARCH(1,1) models and the bottom row shows the QQ plot for AR(1)-EGARCH(1,1). The red line represents the theoretical quantiles under the standard normal distribution, providing a benchmark for assessing deviations in the residuals.

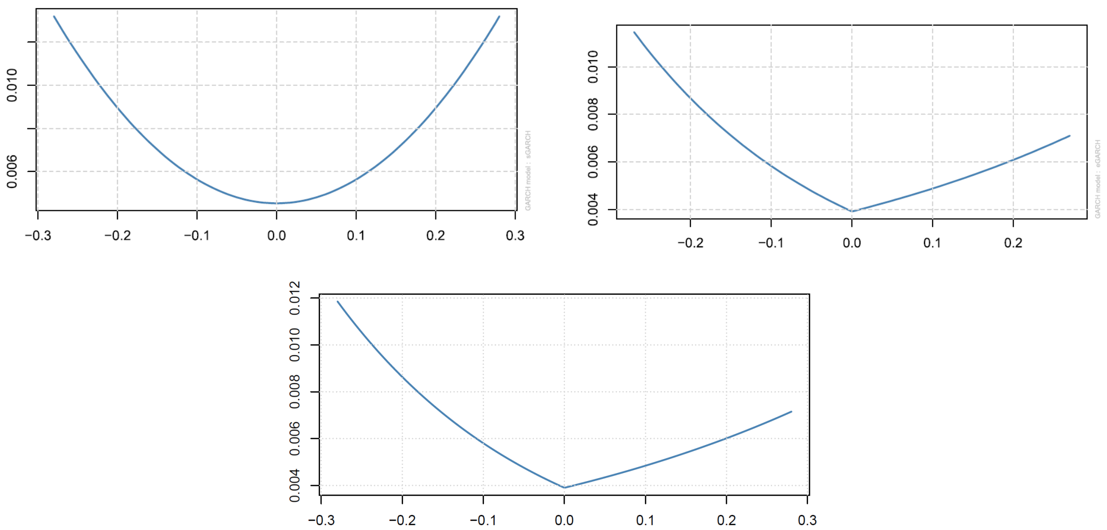


Figure 3.9: News impact curves from GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) models based on IBM monthly stock returns. The top row shows the GARCH(1,1) and EGARCH(1,1) curves and the bottom row presents the AR(1)-EGARCH(1,1) curve.

The QQ plots of standardized residuals from all three models GARCH(1,1), EGARCH(1,1) and AR(1)-EGARCH(1,1) exhibit noticeable deviations from the red reference line, particularly in the tails. This suggests that the residuals are heavy-tailed and that the normality assumption may not fully capture the empirical distribution of IBM stock returns. The presence of excess kurtosis in the standardized residuals across all models indicates that a distribution with heavier tails, such as the Student's t-distribution, might provide a better fit. However, due to time limitations, the estimation of models with non-normal error distributions was not pursued in this thesis. Incorporating such alternative specifications could enhance model performance and represents a promising direction for future research. Turning to the news impact curves, the differences in how the models respond to shocks become even more apparent. The GARCH(1,1) model assumes a symmetric response, meaning that positive and negative shocks of the same magnitude affect volatility equally. In contrast, the EGARCH(1,1) model clearly captures asymmetry, with negative shocks having a more pronounced effect on volatility a behavior consistent with the leverage effect observed in financial markets. The AR(1)-EGARCH(1,1) model retains this asymmetry while also accounting for autocorrelation in returns. This makes it more flexible and better suited to the empirical properties of IBM monthly stock returns, particularly when both volatility clustering and return autocorrelation are present. Together, these graphical diagnostics reinforce earlier statistical findings and suggest that while both EGARCH(1,1) and AR(1)-EGARCH(1,1) improve upon the GARCH(1,1) model, the AR(1)-EGARCH(1,1) specification offers a more nuanced representation of return dynamics by addressing both asymmetric volatility and temporal dependencies.

3.4 The multivariate model

The Multivariate GARCH models were estimated using the multivariate normal distribution, with DCC(1,1) and BEKK(1,1) specifications. The estimated parameters are summarized in Table 3.5, while the log-likelihood and model selection criteria are presented in Table 3.6.

Table 3.5: Estimated Parameters for DCC-GARCH and BEKK-GARCH Models Based on GE and IBM Monthly Stock Returns. Estimation was carried out using Quasi-Maximum Likelihood Estimation (QMLE). The standard errors (in parentheses) are calculated using the asymptotic covariance matrix of the QMLE, obtained directly from the Hessian matrix evaluated at the parameter estimates. Note that the full parameter vector of the BEKK-GARCH model includes additional elements specifically the matrices C, A and B which are not reported here for brevity.

Parameter	**DCC-GARCH**	**BEKK-GARCH**
μ_{GE}	0.013 (0.002)	0.013 (0.002)
ω_{GE}	0.00016 (0.00006)	0.00016 (0.00006)
$\alpha_{1,\text{GE}}$	0.112 (0.019)	0.112 (0.019)
$\beta_{1,\text{GE}}$	0.863 (0.019)	0.863 (0.019)
μ_{IBM}	0.014 (0.002)	0.014 (0.002)
ω_{IBM}	0.0003 (0.0001)	0.0003 (0.0001)
$\alpha_{1,\text{IBM}}$	0.110 (0.030)	0.110 (0.030)
$\beta_{1,\text{IBM}}$	0.824 (0.047)	0.824 (0.047)
a_1^{DCC}	0.043 (0.013)	—
b_1^{DCC}	0.920 (0.025)	—

Table 3.6: Log-likelihood, AIC, BIC for DCC-GARCH and BEKK-GARCH

Criterion	DCC-GARCH	BEKK-GARCH
Log-Likelihood	2681	2667
AIC	-5.36	-5.34
BIC	-5.31	-5.30
Hannan-Quinn (HQIC)	-5.34	-5.32
Shibata	-5.36	-5.34

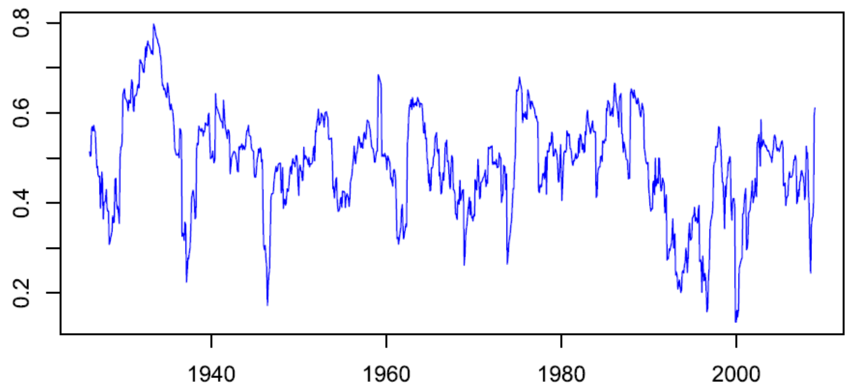


Figure 3.10: Time-Varying Conditional Correlation between GE and IBM stock returns estimated using the DCC-GARCH Model.

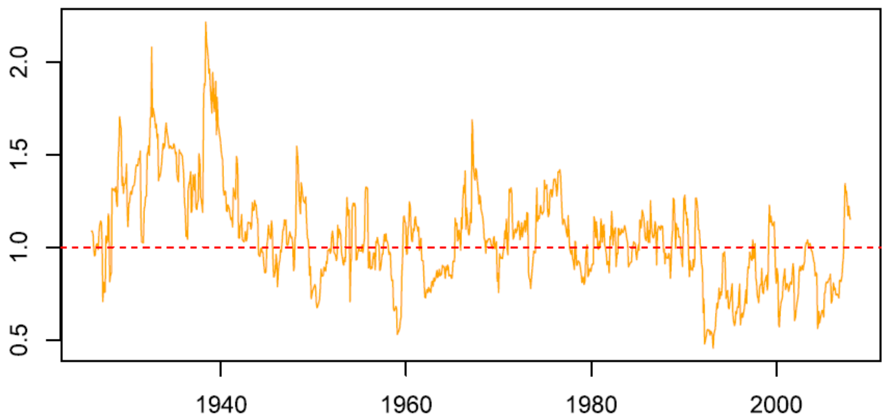


Figure 3.11: Volatility Spillover from GE to IBM Estimated Using the BEKK-GARCH Model. The red dashed line represents the baseline level of spillover (equal to 1), allowing deviations above and below to indicate heightened or reduced spillover effects, respectively.

The DCC and BEKK-GARCH models provide valuable insights into the volatility and correlation dynamics of GE and IBM stock returns (200; Engle, 2002; 215; Engle and Kroner, 1995). One key feature of the DCC-GARCH model is its ability to capture time-varying correlations, as shown in Figure 3.10. The time-varying conditional correlation between the two stocks evolves over time, strengthening during periods of high market volatility and weakening during more stable periods. This adaptability is a major advantage of DCC-GARCH, as it allows for time-dependent co-movement rather than assuming a constant correlation. The model's time-varying correlations are governed by the parameters a_1^{DCC} and b_1^{DCC}, which control the responsiveness to past shocks and the persistence of past correlations, respectively. These parameters capture the dynamics of the correlation process, with a_1^{DCC} reflecting the sensitivity to past innovations in the returns and b_1^{DCC} determining how much weight is given to past correlations in the estimation of current correlations. The results indicate that the conditional correlation tends to increase during periods of heightened market volatility, illustrating how the model captures the dynamic relationship between GE and IBM returns over time (205; Engle, 2002).

Unlike the DCC-GARCH model, which captures time-varying correlations, the BEKK-GARCH model allows for direct volatility spillovers between assets, as illustrated in Figure 3.11. In the BEKK-GARCH framework, spillover effects are measured by the off-diagonal terms of the conditional covariance matrix $\mathbf{H_t}$, specifically the element $H_{GE,IBM,t}$, which represents the covariance between the returns of GE and IBM. The figure illustrates how periods of high volatility in GE lead to increased volatility in IBM, confirming risk transmission effects. A value of 1 represents the baseline level of spillover, with deviations above and below indicating heightened or reduced spillover effects, respectively. These results highlight how periods of high volatility in GE lead to increased volatility in IBM, aligning with the understanding that financial markets exhibit spillover effects, where shocks in one asset significantly influence another (225; Engle and Kroner, 1995; 230; Engle, 2002).

The performance of the two models was compared using log-likelihood values and information criteria, as summarized in Table 3.6. DCC-GARCH achieved a higher log-likelihood (-5.36 vs. -5.34), suggesting a better overall fit (225; Engle, 2002). Additionally, its lower AIC and BIC values reinforce its suitability for modeling dynamic correlations. However, BEKK-GARCH provides deeper insights into volatility transmission, making it particularly useful for understanding financial contagion effects (230; Engle and Kroner, 1995). Both models offer valuable perspectives, with DCC-GARCH excelling in tracking time-varying correlations, making it well-suited for portfolio allocation and risk management (See e.g.; Tsay, 2020), while BEKK-GARCH is particularly effective in capturing volatility spillovers, which is critical for understanding how financial shocks propagate across assets (245; Engle and Kroner, 1995). Thus, the choice between these models depends on the specific financial application, with DCC-GARCH being more appropriate for portfolio diversification strategies, while BEKK-GARCH is better suited for systemic risk analysis and spillover detection.

While both the DCC-GARCH and BEKK-GARCH models offer powerful insights into volatility dynamics and correlations, they come with notable challenges. One significant difficulty is model complexity. Both models, particularly BEKK-GARCH, can be computationally demanding and require careful parameter estimation, which might lead to issues such as overfitting, especially when the dataset is large (250; Bauwens, Laurent, and Rombouts, 2006). Furthermore, both models rely on certain assumptions, such as normality of returns and stationarity, which may not always hold in real financial data (See e.g. ; Tsay, 2020). Violations of these assumptions can lead to model misspecifications, affecting the reliability of the results. Additionally, the need for a sufficiently large dataset to estimate multivariate relationships presents another challenge, as small datasets can lead to unreliable or biased parameter estimates (260; Bauwens et al., 2006). These difficulties highlight the importance of model selection and robustness checks when applying multivariate GARCH models to real-world financial data.

4. Conclution

This thesis investigated the volatility and correlation dynamics of GE and IBM monthly stock returns using both univariate and multivariate GARCH models. The univariate analysis showed that volatility is persistent and asymmetric, with negative shocks having a stronger impact than positive ones. Adding an AR(1) term to the EGARCH model effectively addressed autocorrelation in the residuals and improved model fit. In the multivariate setting, the DCC-GARCH model captured time-varying correlations that intensified during periods of financial stress, while the BEKK-GARCH model revealed clear volatility spillovers from GE to IBM. These results highlight the presence of dynamic co-movements and risk transmission between the two stocks.

Although both models have limitations, such as computational demands and distributional assumptions, they offer valuable insights. The AR(1)-EGARCH(1,1) model was the best fit for the univariate data, while DCC-GARCH was more suitable for capturing changing correlations and BEKK-GARCH for identifying spillover effects. Together, these models provide a comprehensive view of volatility behavior, which is crucial for informed decision-making in risk management and portfolio construction.

Bibliography

[1] Teräsvirta, T., Tjøstheim, D., & Granger, C. W. J. (2010). *Modeling Nonlinear Economic Time Series*. Oxford University Press.

[2] Tsay, R. S. (2005). *Analysis of Financial Time Series* (2nd ed.). John Wiley & Sons.

[3] Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). ARCH models. In Engle, R. F. & McFadden, D. L. (Eds.), *Handbook of Econometrics* (Vol. 4, pp. 2959–3038). Elsevier.

[4] Engle, R. F. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4), 157–168.

[5] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

[6] Bekaert, G., & Harvey, C. R. (1997). Emerging equity market volatility. *Journal of Financial Economics*, 43(1), 29–77.

[7] Christensen, B. J., & Diebold, F. X. (2006). A generalization of the autoregressive conditional heteroskedasticity model. *Econometric Theory*, 22(3), 351–376.

[8] Alexander, C. (2001). *Market Models: A Guide to Financial Data Analysis*. John Wiley & Sons.

[9] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

[10] Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.

[11] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347–370.

[12] Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36(4), 394–419.

[13] Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34–105.

[14] Baba, Y., Engle, R. F., Kraft, D. F., & Kroner, K. F. (1991). Multivariate simultaneous generalized ARCH. Unpublished manuscript, Department of Economics, University of California, San Diego.

[15] Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed.). John Wiley & Sons.

[16] [16] Meitz, M., & Saikkonen, P. (2011). Parameter estimation in nonlinear AR-GARCH models. *Econometric Theory*, *27*(6), 1236–1278.

[17] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* (Vol. 2, 2nd ed.). John Wiley & Sons.

[18] Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Duxbury Press.

[19] Johnson, N. L., & Kotz, S. (2000). *Breakthroughs in Statistics: Methodology and Distribution*. Springer.

[20] Cont, R. (2007). Volatility clustering in financial markets: Empirical facts and agent-based models. In Long memory in economics (pp. 289–309). Springer.

[21] Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579–625.

[22] Cappiello, L., Engle, R. F., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4), 537–572.

[23] Hafner, C. M., & Herwartz, H. (2008). Analytical quasi-maximum likelihood inference in multivariate volatility models. *Econometric Theory*, 24(5), 1172–1200.

[24] Comte, F., & Lieberman, O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis*, 84(1), 61–84.

[25] Engle, R. F., & Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper No. 8554.

[26] Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5), 1779–1801.

[27] Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21(1), 79–109.

[28] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

[29] R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

[30] Ghalanos, A. (2022). *rugarch: Univariate GARCH Models*. R package version 1.4-8.

[31] CRSP. (2021). *Center for Research in Security Prices (CRSP) Monthly Stock File*. University of Chicago Booth School of Business.

[32] Romer, C. D. (2003). Great Depression. In D. R. Henderson (Ed.), *The Concise Encyclopedia of Economics*. Liberty Fund.

[33] Tsay, R. S. (2020). *Multivariate Time Series Analysis: With R and Financial Applications*. John Wiley & Sons.

[34] Reinhart, C. M., & Rogoff, K. S. (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press.

[35] Hamilton, J. D. (1983). Oil and the macroeconomy since World War II. *Journal of Political Economy*, 91(2), 228–248.

[36] Shiller, R. J. (2005). *Irrational Exuberance* (2nd ed.). Princeton University Press.

[37] Hosking, J. R. M. (1980). The Multivariate Portmanteau Statistic. *Journal of the American Statistical Association*, 75(371), 602–608.

[38] McLeod, A. I., & Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4(4), 269–273.

[39] Doornik, J. A., & Hansen, H. (2008). An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics*, 70, 927–939.

[40] Henze, N., & Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics – Theory and Methods*, 19(10), 3595–3617.