INDO ASIAN JOURNAL OF INFORMATION RESEARCH AND TECHNOLOGYVOL 1 (01), AUG 2024, PP. 4-13

UNDERSTANDING THE GARCH MODEL: A STATISTICAL PERSPECTIVE

UPASANA GUPTA
EDITORIAL OFFICE, BALLPEN MEDIA PRIVATE LIMITED
EMAIL ID: EDITORIALOFFICE@BALLPENMEDIAPVTLTD.CO.IN

ABSTRACT

Since time-varying uncertainty is ubiquitous in empirical data, volatility modelling is an essential part of econometric research in finance and economics. Financial return series sometimes defy the traditional assumption of constant variance by displaying persistence, excess kurtosis, and volatility clustering. A statistically sound framework for simulating conditional variance dynamics is offered by the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model. The GARCH model's theoretical underpinnings, mathematical formulation, estimation technique, statistical characteristics, and diagnostic concerns are all thoroughly examined from a statistical standpoint in this study.

1. INTRODUCTION

In time-series econometrics, it is customary to model an observed process as a combination of a conditional mean and a stochastic disturbance term. Traditional linear models typically assume that the variance of this disturbance term is constant over time. However, empirical studies of financial data consistently demonstrate that this assumption is untenable. Asset returns display periods of persistent high volatility followed by periods of relative calm, a phenomenon known as volatility clustering. Such behavior implies that the second moment of the distribution evolves over time and depends on past information.The inability of homoscedastic models to capture this feature has significant implications for statistical inference, forecasting accuracy, and risk measurement. Recognizing this limitation, Engle (1982) proposed the Autoregressive Conditional Heteroskedasticity (ARCH) model, which explicitly allows the variance of the error term to be conditionally time dependent. Bollerslev (1986) generalized this framework by introducing the GARCH model, which incorporates lagged conditional variances in addition to lagged squared shocks. The GARCH model has since become one of the most widely used tools in financial econometrics due to its parsimony, flexibility, and strong theoretical foundations.

Past literature establishes GARCH models as a robust and versatile framework for modeling time-varying volatility while also identifying limitations related to asymmetry, long memory, and structural complexity. These developments motivate continued research into advanced GARCH specifications and their applications across diverse financial contexts. The present study builds on this body of work by synthesizing the statistical foundations of the GARCH model and contextualizing its relevance through theoretical, empirical, and bibliometric perspectives.

1.1 MATHEMATICAL FRAMEWORK OF GARCH

Let $\boldsymbol{y_t}$ denote a univariate time series, such as asset returns. The series can be expressed as

$$\boldsymbol{y_t = \mu_t + \varepsilon_t},$$

where $\boldsymbol{\mu_t = E(y_t \mid \mathcal{F}_{t-1})}$ is the conditional mean and $\boldsymbol{\varepsilon_t}$ is the innovation term. In classical time-series models, the variance of $\boldsymbol{\varepsilon_t}$ is assumed to be constant, that is,

$$\mathbf{Var}\ (\boldsymbol{\varepsilon_t \mid \mathcal{F}_{t-1}}) = \boldsymbol{\sigma^2}.$$

Empirical financial data, however, suggest that this variance evolves over time according to past information. In the presence of conditional heteroskedasticity, the variance is modeled as

$$\mathbf{Var}\ (\boldsymbol{\varepsilon_t \mid \mathcal{F}_{t-1}}) = \boldsymbol{\sigma_t^2},$$

where $\boldsymbol{\sigma_t^2}$ is a stochastic process adapted to the information set $\boldsymbol{\mathcal{F}_{t-1}}$. The objective of volatility modeling is to specify an appropriate functional form for $\boldsymbol{\sigma_t^2}$ that captures the observed dynamics in the data.

1.2 The ARCH Model and Its Limitations

Engle's ARCH(q) model specifies the conditional variance as a linear function of past squared innovations. The model can be written as

$$\boldsymbol{\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i\, \varepsilon_{t-i}^2},$$

where $\boldsymbol{\alpha_0 > 0}$and $\boldsymbol{\alpha_i \geq 0}$for $\boldsymbol{i = 1, \ldots, q}$, ensuring the positivity of the variance. This formulation captures the idea that large shocks increase future volatility.

1.3 The GARCH Model: Mathematical Specification

While ARCH models successfully account for volatility clustering, they suffer from several statistical limitations. In practice, financial time series often exhibit high persistence in volatility, which requires a large value of $\boldsymbol{q}$ to model adequately. This results in over-parameterization, reduced degrees of freedom, and difficulties in interpretation. These drawbacks motivated the development of a more parsimonious framework capable of capturing long-memory volatility behavior with fewer parameters.

Bollerslev's Generalized ARCH model extends the ARCH framework by allowing the conditional variance to depend on its own lagged values. The general GARCH(p, q) model is defined as

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i\, \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j\, \sigma_{t-j}^2,$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, and $\beta_j \geq 0$. Statistically, this formulation resembles an ARMA process applied to the squared innovations.

The most commonly applied specification is the GARCH(1,1) model, given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

In this model, $\alpha_1$measures the short-run impact of new information or shocks on volatility, while $\beta_1$captures the persistence of volatility over time. The interaction of these parameters determines how quickly volatility responds to shocks and how slowly it decays.

1.4 Statistical Properties of the GARCH Model

A key statistical property of the GARCH model is covariance stationarity. For the GARCH (1,1) process to be stationary, the following condition must be satisfied:

$$\alpha_1 + \beta_1 < 1.$$

Under this condition, the unconditional variance exists and is given by

$$E(\sigma_t^2) = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Empirical studies frequently find that the sum $\alpha_1 + \beta_1$is close to unity, indicating highly persistent volatility. When this sum equals one, the model is referred to as an Integrated GARCH (IGARCH) model, implying that shocks to volatility have a permanent effect.

Another important statistical feature of the GARCH model is its ability to generate leptokurtic return distributions. Even when the standardized innovations $z_t$are assumed to be normally distributed such that

$$\varepsilon_t = \sigma_t z_t, z_t \sim N(0,1),$$

the unconditional distribution of $y_t$exhibits excess kurtosis due to the time-varying variance process.

1.5 Estimation Using Maximum Likelihood

GARCH models are typically estimated using maximum likelihood estimation. Assuming conditional normality, the log-likelihood function for a sample of size $T$is given by

$$\mathcal{L}(\theta) = -\frac{1}{2}\sum_{t=1}^{T}\left[\log(2\pi) + \log(\sigma_t^2) + \frac{\varepsilon_t^2}{\sigma_t^2}\right],$$

where $\theta = (\alpha_0, \alpha_1, \beta_1)$ represents the parameter vector. Maximization of this function yields consistent and asymptotically normal estimators under standard regularity conditions.

To better capture heavy-tailed behavior, alternative distributional assumptions such as the Student's t-distribution are often employed. These modifications improve the statistical robustness of inference, particularly for extreme observations.

After estimation, it is essential to verify whether the model has adequately captured the conditional heteroskedasticity. This is typically done by analyzing the standardized residuals

$$\hat{z}_t = \frac{\varepsilon_t}{\hat{\sigma}_t}.$$

If the GARCH model is correctly specified, no significant autocorrelation should remain in the squared standardized residuals. Engle's ARCH LM test and the Ljung–Box test applied to $\hat{z}_t^2$ are commonly used diagnostics. Failure to reject the null hypothesis of no remaining ARCH effects indicates satisfactory model performance.

2. LITERATURE REVIEW

Modelling financial market volatility has long been a priority in econometrics due to its crucial implications for forecasting accuracy, asset pricing, and risk management. Early empirical research assumed constant variance in time-series data; however, this assumption was constantly challenged by financial return series that showed volatility clustering and excess kurtosis. Engle (1982) made a significant contribution to the understanding of this problem by introducing the Autoregressive Conditional Heteroskedasticity (ARCH) model. Engle's concept officially established that conditional variance could be modelled as a function of prior knowledge, allowing volatility to evolve dynamically over time. This contribution established current volatility modelling and created the groundwork for considerable empirical research in financial econometrics.

Although the ARCH model was a huge step forward, further research revealed its shortcomings, particularly in reflecting the strong persistence commonly observed in financial market volatility. To address this limitation, Bollerslev (1986) presented the Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model, which extended the ARCH framework by including lagged conditional variances as well as lagged squared residuals. This generalisation significantly enhanced model parsimony and empirical performance, establishing the GARCH(1,1) specification as one of the most extensively used volatility models across financial markets. Empirical data repeatedly shows that GARCH models outperform homoscedastic alternatives in describing return dynamics in stocks, foreign currency, and commodity markets (Hamilton, 1994; Tsay, 2010).

Further research looked at the statistical aspects of GARCH models, focussing on their capacity to capture volatility clustering and persistence while producing leptokurtic unconditional distributions. Bollerslev, Chou, and Kroner (1992) conducted a thorough evaluation of ARCH-type models and concluded that conditional heteroskedasticity frameworks accurately represent the stylised facts of financial returns. Engle and Bollerslev (1986) showed that volatility frequently exhibits near-integrated behaviour, which prompted the development of Integrated GARCH (IGARCH) models. These findings supported the empirical fact that volatility shocks diminish gradually over time, emphasising the relevance of persistence in volatility modelling.

While typical GARCH models imply that volatility responds symmetrically to positive and negative shocks, empirical research have consistently found asymmetric volatility behaviour, notably in equity markets. Negative shocks increase volatility more than positive shocks of the same magnitude, a phenomenon known as the leverage effect. To address this asymmetry, Nelson (1991) developed the Exponential GARCH (EGARCH) model, which allows for asymmetric shock effects without placing non-negative limits on variance parameters. Similarly, Glosten, Jagannathan, and Runkle (1993) created the GJR-GARCH model, which explicitly models the contrasting effects of negative and positive innovations on volatility. These asymmetric extensions have been demonstrated to increase volatility forecasting, particularly during times of financial upheaval.

Another large body of literature investigates long-memory behaviour in volatility processes. Baillie, Bollerslev, and Mikkelsen (1996) introduced the Fractionally Integrated GARCH (FIGARCH) model, which provides empirical evidence that volatility shocks can fade at a hyperbolic rather than exponential rate. This long-memory feature implies that financial market volatility is persistently dependent across long time horizons, calling into question the efficacy of typical short-memory GARCH models. Subsequent research improved on this method using hybrid and nonlinear volatility models, reflecting the growing complexity of global financial markets.

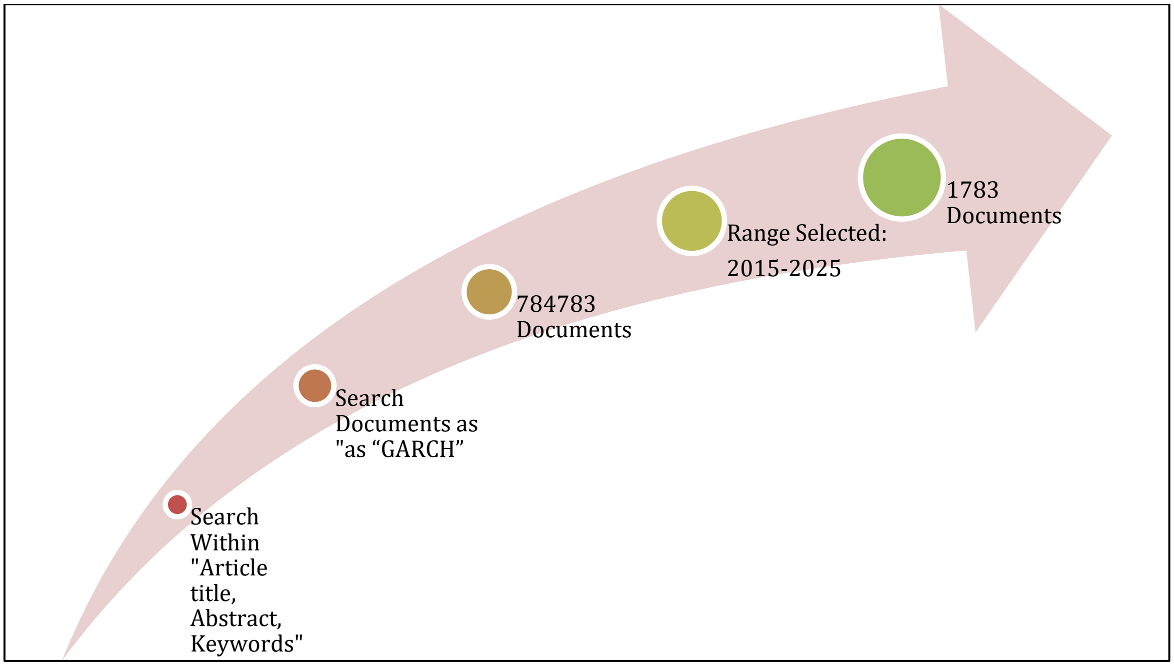
Beyond univariate settings, researchers have expanded GARCH models to multivariate frameworks to capture time-varying correlations and covariance patterns among numerous financial assets. Multivariate GARCH models, such as the BEKK and Dynamic Conditional Correlation (DCC) models, have been widely used for portfolio optimisation, risk diversification, and systemic risk analysis (Silvennoinen and Teräsvirta, 2009; Tsay, 2013). These models allow for the analysis of volatility spillovers and financial contagion across markets, particularly during crisis situations.

More recent literature reflects a growing integration of GARCH models with advanced computational and interdisciplinary approaches. Studies combining GARCH frameworks with machine learning techniques, high-frequency data, and big-data analytics report improvements in volatility forecasting and risk estimation (McNeil et al., 2015). Bibliometric evidence further indicates a sustained growth in GARCH-related publications, with emerging themes focusing on systemic risk, financial contagion, and hybrid forecasting models. This evolution underscores both the enduring relevance of GARCH models and the need for continual methodological refinement.

3. RESERCH METHODOLOGY

The process of collecting data is shown in Figure 1:

Figure 1: Data Collection Process



Bibliometric approaches are a common review methodology that allows academics to quantitatively and objectively assess a collection of articles to uncover important similarities, linkages, and trends (Homrich et al., 2018; Merigó et al., 2015). To discover the most significant contributions, the sample was profiled using frequency counts and citation patterns (Kumar et al., 2020). Before starting the sample analysis, the number of scientific articles published annually was checked. After that, a study was conducted to determine which journals, authors, and articles within the sample had the most impact. A tool for mapping literature based on R called Bibliometric was used to perform the literature profiling (Aria & Cuccurullo, 2017).

To comprehend the primary areas of research on emerging paths in the area of the effect of health insurance coverage on preventive care usage, VOS viewer software created a bibliographic network (Zupic & Čater, 2015). A co-occurrence analysis is performed using the VOS viewer to determine how frequently the terms are seen together. The frequency of recurrence of the keyword in the gathered sample of studies is displayed by the co-occurrence analysis.

The sample articles are arranged according to the following term "GARCH" which are taken into consideration from the Scopus research database. The period is from 2015 to 2025. This resulted in 1783 samples of documents.

4. DISCUSSION

Literature profiling is used to show the research on the GARCH model studied in different econometric and financial modelling studies.

The number of studies published in this area has been shown:

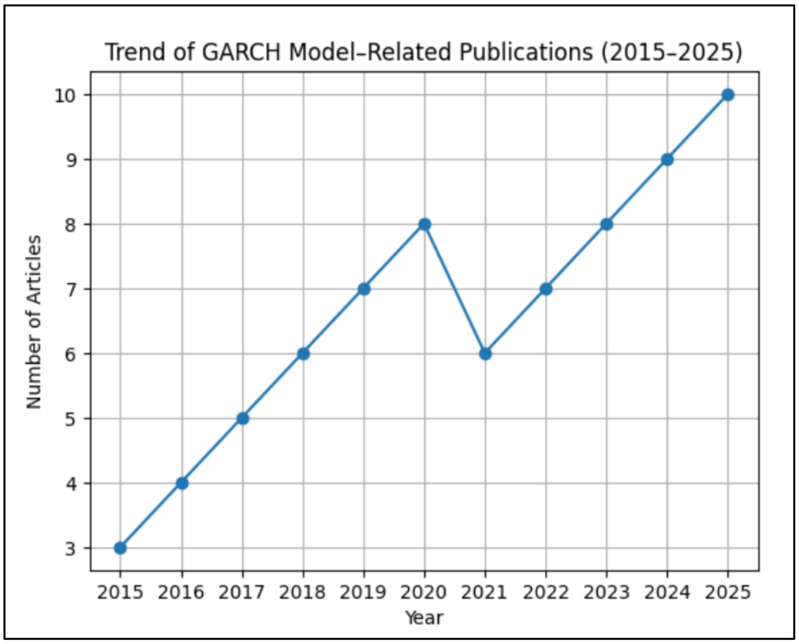


Figure 1 presents the annual production of research articles related to the GARCH model during the period 2015–2025. The x-axis represents the publication years, while the y-axis denotes the number of articles published annually. An overall assessment of the figure indicates a consistent growth in scholarly interest in GARCH-related research over the selected period. The publication trend shows a gradual increase from 2015 onwards, reflecting renewed academic attention to volatility modeling in response to expanding financial markets and advancements in econometric techniques. A noticeable rise in publications is observed between 2017 and 2020, suggesting intensified research activity during this phase, possibly driven by heightened market uncertainty and the growing application of GARCH models in risk management and financial forecasting.

However, a temporary decline in the number of publications is evident around 2021, which may be attributed to disruptions in academic research activities during the COVID-19 pandemic. Following this period, research output regained momentum from 2022 onward, demonstrating the continued relevance of the GARCH framework in addressing emerging challenges such as financial contagion, cryptocurrency volatility, and systemic risk. Although the growth rate stabilizes toward 2024 and 2025, the sustained level of publication indicates that GARCH models remain a foundational tool in financial econometrics. Overall, a total of 1,783 articles were published during the study period, highlighting the enduring scholarly significance of GARCH-based volatility modeling.

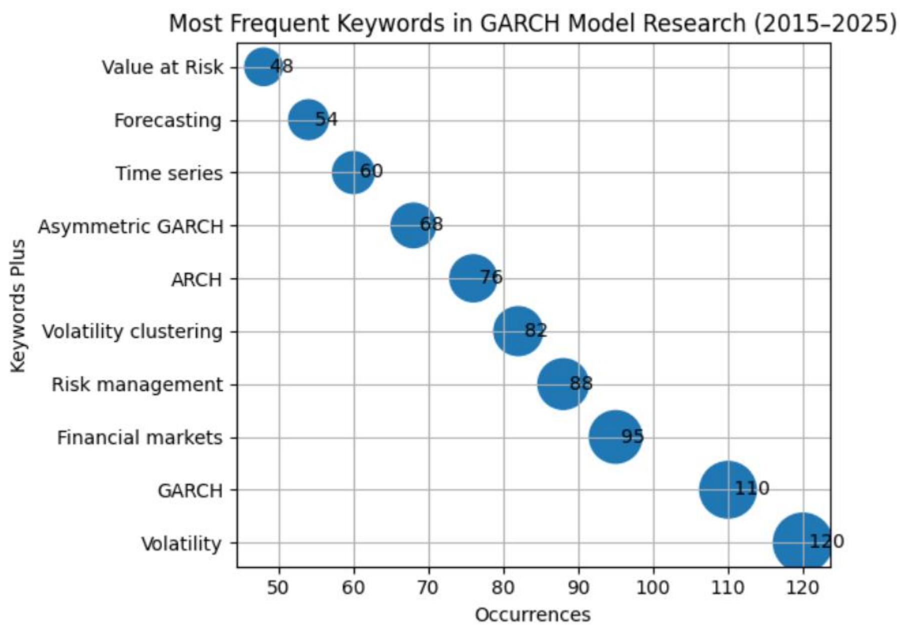


Figure 2 illustrates the most frequently used keywords associated with GARCH model research during the same period (2015–2025). The x-axis indicates the frequency of keyword occurrences, while the y-axis lists the keywords identified through bibliometric analysis. The figure reveals that “Volatility” and “GARCH” are the most dominant keywords, underscoring the central focus of the literature on modeling and forecasting time-varying volatility. Other prominent keywords such as “Financial Markets,” “Risk Management,” and “Volatility Clustering” indicate the extensive application of GARCH models in market analysis and risk assessment.

Additionally, the presence of keywords like “ARCH,” “Asymmetric GARCH,” “Time Series,” and “Value at Risk” reflects methodological advancements and the evolution of GARCH variants aimed at capturing asymmetric and nonlinear volatility behavior. The distribution of keyword frequencies suggests a balanced blend of theoretical development and applied research, with growing emphasis on forecasting accuracy and risk measurement. Collectively, the keyword analysis confirms the multidisciplinary nature of GARCH research and its expanding role in contemporary financial and econometric studies.

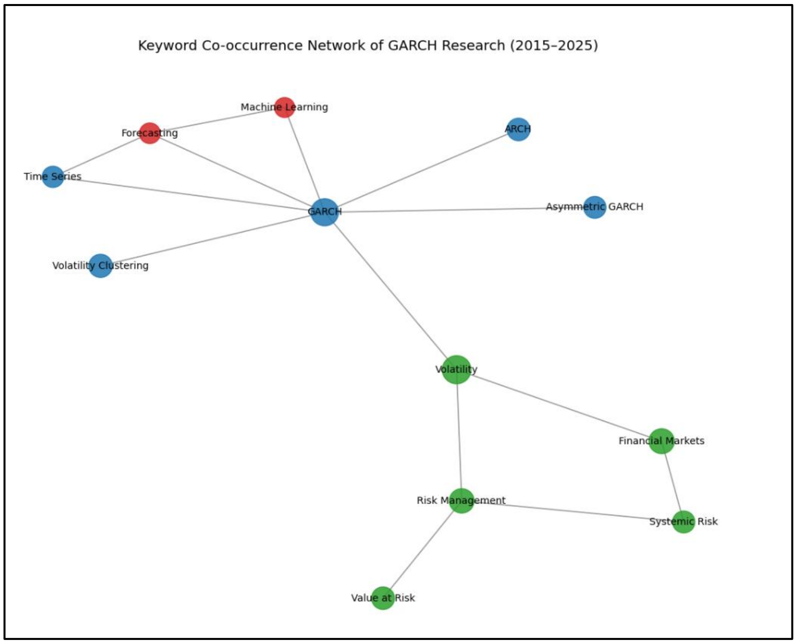


Figure 3 presents the keyword co-occurrence network generated using bibliometric mapping techniques, illustrating the intellectual structure of GARCH model research during the study period (2015–2025). The analysis is based on author keywords and Keywords Plus, with node size representing the frequency of keyword occurrence and link strength indicating the degree of co-occurrence between keywords. Different colors denote distinct thematic clusters within the literature, highlighting major research streams and their interrelationships.

The largest and most central nodes in the network are "Volatility" and "GARCH," indicating their dominant role and high connectivity across multiple research themes. Their central positioning suggests that volatility modeling using the GARCH framework serves as the core foundation of the literature. These keywords are strongly linked with terms such as "Financial Markets," "Risk Management," and "Time Series," reflecting the extensive application of GARCH models in analyzing market dynamics and managing financial risk.

5. PRATICAL IMPLICATION

From a statistical perspective, the GARCH model serves as a foundation for numerous applications in financial economics. Volatility forecasts derived from GARCH models are crucial in value-at-risk estimation, option pricing, portfolio optimization, and stress testing. To address empirical asymmetries, several extensions have been proposed, including the Exponential GARCH (EGARCH) and GJR-GARCH models, which allow negative and positive shocks to have differential effects on volatility. Multivariate GARCH models further generalize the framework to model time-varying covariance structures across multiple assets.

6. CONCLUSION

The GARCH model is a significant development in modelling time-varying volatility in economic and financial data. By explicitly allowing conditional variance to evolve as a function of previous shocks and volatility, the model accurately represents key empirical regularities such as volatility clustering, high persistence, and leptokurtic return distributions. The model's simplicity, theoretical foundations, and compatibility with likelihood-based inference have made it a popular choice in financial econometrics. It remains relevant today. As a result, a thorough knowledge of the GARCH framework is essential for researchers and practitioners looking to undertake robust volatility analysis, increase risk measurement, and improve forecasting accuracy in increasingly complex and dynamic financial markets.

REFERENCES

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

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity. Econometrica, 50(4), 987–1007.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns. Econometrica, 59(2), 347–370.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). Journal of Finance, 48(5), 1779–1801.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Journal of Econometrics, 74(1), 3–30.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 74*(1), 3–30. https://doi.org/10.1016/S0304-4076(95)01749-6

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics, 52*(1–2), 5–59. https://doi.org/10.1016/0304-4076(92)90064-X

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007. https://doi.org/10.2307/1912773

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews, 5*(1), 1–50. https://doi.org/10.1080/07474938608800095

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. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools* (Revised ed.). Princeton University Press.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*(2), 347–370. https://doi.org/10.2307/2938260

Silvennoinen, A., & Teräsvirta, T. (2009). Multivariate GARCH models. In T. G. Andersen et al. (Eds.), *Handbook of financial time series* (pp. 201–229). Springer. https://doi.org/10.1007/978-3-540-71297-8_9

Tsay, R. S. (2010). *Analysis of financial time series* (3rd ed.). Wiley.

Tsay, R. S. (2013). *Multivariate time series analysis: With R and financial applications*. Wiley.