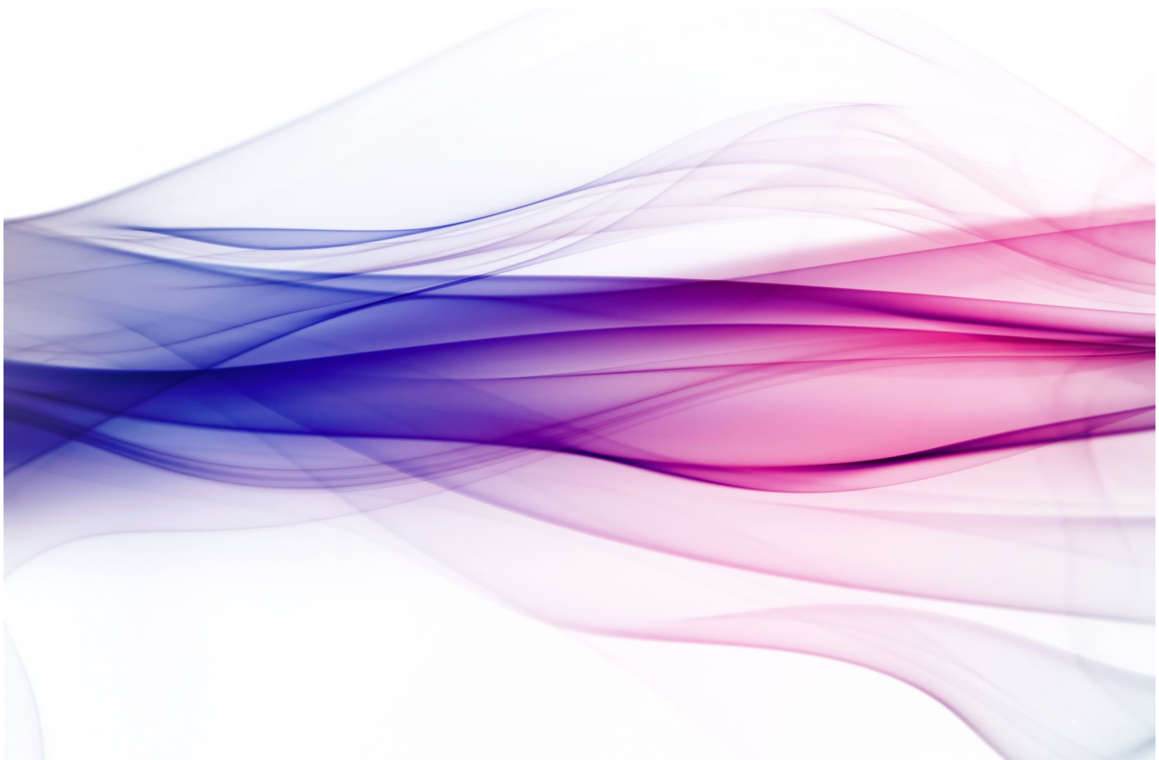
NITH INVESTMENTS
TECHNOLOGY • INSIGHT • CLARITY

GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY

Capturing the Memory of Volatility in Interest Rates and Asset Price Returns



Jimisi Rajasombat
Investment Research

NITH INVESTMENTS
TECHNOLOGY • INSIGHT • CLARITY

TABLE OF CONTENTS

INTRODUCTION

The landscape of financial markets is characterized by turbulence, calm, crisis, and recovery. Within this ever-shifting terrain, the idea that volatility exhibits memory to a noticeable degree has become quite clear. Large price movements (up or down) tend to cluster together. Periods of market turmoil are followed by continued turmoil, and periods of tranquility are followed by continued calm. This phenomenon, known as volatility clustering, represents the financial market's memory of past shocks. Capturing this memory via statistics and mathematical formulas is essential for practical financial applications.

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, introduced by Tim Bollerslev in 1986, provides an efficient and rigorous framework for modeling this time-varying volatility and has become the dominant family of models in the financial industry for volatility forecasting. This paper explores the GARCH framework in depth, explaining its mathematical structure, its advantages over traditional alternatives, its parameter interpretation, and its practical applications in forecasting.

THE PROBLEM: HETEROSKEDASTICITY AND CLASSICAL MODELS

Traditional time series models such as ARMA and ARIMA assume that the variances of the error terms in time series movements remain constant over time. This assumption is known as homoskedasticity. However, financial returns exhibit heteroskedasticity, where the variance changes systematically. More specifically, they display the type of volatility clustering described above.

This clustering behavior violates the constant-variance homoskedasticity assumption underlying classical statistical models. If we apply standard linear models to data exhibiting volatility clustering, our statistical inferences would become unreliable. We need a model that allows variance itself to be time-varying and predictable based on past information. Indeed, before GARCH there were other popular attempts (such as EWMA and ARCH) to capture heteroskedasticity.

PREVIOUS MODELS OF HETEROSKEDASTICITY: EWMA AND ARCH

Attempts at capturing heteroskedasticity relies on the idea of conditional variance. Conditional variance is defined as the variance of returns at time t given all information available up to time $t-1$, and it is modeled as a time-varying quantity rather than a constant. This is important because in essence it measures the memory of past clustering.

Previous models of heteroskedasticity have incorporated this idea to some degree.

Exponentially Weighted Moving Average (EWMA)

Practitioners often used the Exponentially Weighted Moving Average (EWMA) model to estimate time-varying volatility before GARCH became the industry standard. The EWMA variance forecast is given by:

$$\sigma_t^2 = (1-\lambda)\epsilon_{t-1}^2 + \lambda\sigma_{t-1}^2$$

where:

- σ_t^2: Current conditional variance estimate at time t. This is the EWMA volatility being updated.
- ε_{t-1}^2: Squared return innovation (or residual) from the previous period, usually $(r_{t-1} - \mu)^2$, representing the most recent shock to returns.
- σ_{t-1}^2: Previous period's conditional variance estimate, the volatility EWMA calculated at time $t-1$.
-
- λ: Decay factor or smoothing parameter set between 0 and 1 (typically somewhere between 0.94 and 0.97), controlling how quickly older information is down-weighted (larger λ= slower decay).

While EWMA captures volatility clustering by giving more weight to recent observations, it has significant limitations. The decay parameter λ is typically fixed rather than estimated from data, the model lacks a formal statistical foundation for hypothesis testing, and it provides no explicit long-run variance level or mean reversion target.

ARCH Models

Robert Engle's ARCH(q) model was the first to provide a rigorous statistical framework for modeling conditional heteroskedasticity.

In an ARCH(q) model, the conditional variance depends on past squared returns:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \epsilon_{t-2}^2 + \cdots + \alpha_q \epsilon_{t-q}^2$$

The new variables introduced in this formula are:

- ω: Constant term. A baseline or long-run component of the variance, required to keep σ_t^2 positive.
- α_i: ARCH coefficients that weight each lagged squared shock. They must be non-negative, and a larger α_i means that the corresponding past shock has a stronger effect on current volatility.
- q: Order of the ARCH model, indicating how many past squared shocks enter the variance equation (how far back volatility remembers).

While ARCH represented a major breakthrough, empirical applications revealed a significant drawback. To adequately capture the long memory in volatility observed in financial data, very high orders (large values of q) were often required. This resulted in models with many parameters, leading to estimation difficulties and reduced forecasting efficiency.

GARCH MODEL: A PARSIMONIOUS SOLUTION OVER THE ARCH MODEL

The GARCH(1,1) Model

Bollerslev's GARCH model addressed the limitations of ARCH by incorporating lagged conditional variances into the variance equation, creating an autoregressive moving average structure for variance.

The standard GARCH(1,1) model is defined with:

$$r_t = \mu + \epsilon_t$$

$$\epsilon_t = \sigma_t z_t \text{ where } z_t \sim N(0,1)$$

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

where:

- r_t: The return at time t.
- μ: The mean return.
- ϵ_t: The innovation or shock, defined similarly as in the previous equations.
- σ_t^2: The conditional variance, conditional on information available at time $t-1$.
- z_t: A standardized innovation (white noise with zero mean and unit variance).
- ω, α, and β are model parameters to be estimated.

Understanding the (1,1) Notation

The notation GARCH(p,q) indicates the order of the model, where $\mathbf{p}$ represents the number of lagged conditional variance terms (the GARCH terms) and $\mathbf{q}$ represents the number of lagged squared error terms (the ARCH terms). Thus, GARCH(1,1) means:

- One lagged conditional variance term: $\beta\sigma_{t-1}^2$
- One lagged squared error term: $\alpha\epsilon_{t-1}^2$

Higher-order models such as GARCH(2,1) or GARCH(1,2) include additional lags. However, empirical studies have shown that GARCH(1,1) is consistently effective at capturing volatility dynamics in financial time series, often outperforming more complex specifications. This parsimony, the ability to capture complex volatility patterns with just three parameters, is one of GARCH's greatest strengths.

WHY GARCH IS PREFERRED OVER ALTERNATIVES

GARCH models offer several advantages over EWMA and ARCH:

- Statistical Foundation: GARCH provides a complete probabilistic framework with formal likelihood-based estimation and hypothesis testing capabilities, enabling rigorous model comparison and inference.
- Parsimony: GARCH(1,1) achieves what might require ARCH(20) or higher with just three parameters, dramatically improving estimation efficiency.
- Flexibility: The separate α and β parameters allow independent control over shock sensitivity and persistence, providing more realistic modeling of actual volatility dynamics.
- Mean Reversion: GARCH naturally incorporates mean reversion to a long-run variance level, a stylized fact observed in all financial markets.
- Forecasting Capability: GARCH provides explicit multi-period-ahead volatility forecasts with well-defined convergence properties.

DEEP DIVE: THE GARCH PARAMETERS IN PLAIN ENGLISH

Below is what each of the new parameters (α, β and ω) introduced by GARCH captures about market behavior.

Alpha (α): The News Coefficient or Shock Sensitivity

The parameter α captures how sensitive volatility is to recent market shocks. Specifically, it measures how much yesterday's squared return ϵ_{t-1}^2 contributes to today's expected volatility.

A large α (e.g., 0.15-0.20) means that markets react strongly to new information. A large price movement yesterday will cause a substantial increase in today's expected volatility. A small α (e.g., 0.05) indicates that markets are relatively unresponsive to individual shocks.

The constraint $\alpha > 0$ ensures that shocks always increase (never decrease) volatility, which is intuitive since unexpected news, whether good or bad, creates uncertainty and raises volatility.

Therefore, α is the reaction coefficient. It tells us how much today's volatility spiked in response to yesterday's surprise. A high α means that the markets are nervous and reactive, while a low α means that markets can take news in stride and are less jumpy.

Beta (β): The Persistence Coefficient or Memory

The parameter β captures how of much yesterday's volatility level σ_{t-1}^2 carries forward into today. This is sometimes called the persistence of volatility. This is the true memory parameter of the model.

A high β (e.g., 0.85-0.95) indicates that volatility is highly persistent. Once volatility rises, it stays elevated for many periods. A low β (e.g., 0.50-0.70) suggests volatility shocks dissipate quickly.

The constraint $\beta > 0$ ensures positive persistence, and the stationarity requirement $\alpha + \beta < 1$ ensures that volatility shocks eventually decay.

In plain English, β measures stickiness or inertia in volatility. It answers the question: once the market becomes volatile, how long does that volatility persist? High β means volatility has a long memory and low β means markets quickly forget yesterday's stress.

Omega (ω): The Baseline Volatility Component

The parameter ω is a positive constant that represents the baseline component of volatility. It is not the long-run variance itself, but rather an irreducible minimum that ensures volatility never reaches zero. Think of ω as the foundation upon which market volatility is built. Without this positive constant, the model could theoretically produce zero or negative variance, which is economically meaningless.

Mathematically, ω must satisfy $\omega > 0$. Typical values in financial applications are very small (e.g., 0.000002 to 0.0001).

The Interaction: How GARCH Captures Volatility Clustering

The key mechanism of GARCH lies in how α and β work together in the recursive equation:

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

Let's trace through the mechanism step by step:

1. Initial Shock: Suppose at time $t-1$ there is a large market shock, producing a large squared return ϵ_{t-1}^2.

2. Immediate Response: Through the α term, this large shock directly increases the variance forecast for time t and σ_t^2 rises.

3. Persistence: The elevated σ_t^2 now feeds into the next period through the β term: $\sigma_{t+1}^2 = \omega + \alpha\epsilon_t^2 + \beta\sigma_t^2$. Even if ϵ_t^2 is moderate, the large σ_t^2 keeps σ_{t+1}^2 elevated.

4. Clustering Emerges: This process continues. High volatility today leads to high expected volatility tomorrow, which leads to high expected volatility the day after, and so on. This creates clusters of high volatility periods.

5. Eventually Dissipates: Because $\alpha + \beta < 1$, the effect of the original shock gradually decays. Each period, volatility moves a bit closer to its long-run level.

The critical insight is that the α parameter creates the immediate spike in volatility after a shock, while the β parameter creates the prolonged elevation. This is the clustering pattern we observe in real financial data.

PERSISTENCE, MEAN REVERSION, AND LONG-RUN VARIANCE

Volatility Persistence

The sum $\alpha + \beta$ therefore makes up a persistence parameter and is central to understanding GARCH dynamics. This sum measures how close the volatility process is to a random walk:

- If $\alpha + \beta \approx 1$ (e.g., 0.98-0.995), then volatility is highly persistent with very slow mean reversion.
- If $\alpha + \beta \approx 0.85$, then volatility mean reverts relatively quickly.
- If $\alpha + \beta = 1$, then the model exhibits no mean reversion, and the shocks have permanent effects.
- If $\alpha + \beta > 1$, then the process is explosive and non-stationary.

For the model to be covariance stationary, we require $\alpha + \beta < 1$.

Mean Reversion and the Long-Run Variance

When $\alpha + \beta < 1$, the GARCH process is mean-reverting, meaning volatility tends to revert toward a long-run average level. This unconditional variance $\overline{\sigma^2}$ (also called the long-run variance) is given by:

$$\overline{\sigma^2} = \frac{\omega}{1 - \alpha - \beta}$$

This formula reveals an insightful relationship. While ω anchors the equation, the actual long-run variance level depends on all three parameters working together. The term $\gamma = 1 - \alpha - \beta$ is sometimes called the mean reversion rate.

Example: If $\omega = 0.000003$, $\alpha = 0.12$, and $\beta = 0.87$, then:

- Persistence: $\alpha + \beta = 0.99$ (highly persistent)
- Mean reversion rate: $\gamma = 1 - 0.99 = 0.01$

- Long-run variance: $\overline{\sigma^2} = 0.000003/0.01 = 0.0003$
- Long-run volatility (standard deviation): $\bar{\sigma} = \sqrt{0.0003} \approx 0.0173$ or 1.73%

Half-Life of Volatility Shocks

The half-life measures how many periods it takes for volatility to revert halfway back to its long-run level after a shock. It is calculated as:

$$\text{Half-Life} = \frac{\ln(0.5)}{\ln(\alpha + \beta)}$$

Example calculations:

- If $\alpha + \beta = 0.95$: Half-life ≈ 13 periods
- If $\alpha + \beta = 0.98$: Half-life ≈ 34 periods
- If $\alpha + \beta = 0.99$: Half-life ≈ 69 periods

For daily data, a persistence of 0.98 means it takes approximately 34 days for half the effect of a volatility shock to dissipate. We can interpret this as over a month of memory is embedded in financial volatility, so the effect of a major market event can linger for weeks and not just days.

ESTIMATING GARCH PARAMETERS VIA MAXIMUM LIKELIHOOD

GARCH parameters are typically estimated using Maximum Likelihood Estimation (MLE). Under the assumption of normally distributed innovations where $z_t \sim N(0,1)$, the log-likelihood function for a sample of T observations is:

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

where $\theta = (\omega, \alpha, \beta)$ and σ_t^2 is computed recursively using the GARCH equation.

The MLE Procedure

1. **Initialization**: Start with an initial variance estimate (often the sample variance or an EWMA estimate).

2. **Recursive Calculation**: For each observation t, calculate $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$.

3. **Likelihood Evaluation**: Compute the log-likelihood contribution for each observation.

4. **Optimization**: Use numerical optimization algorithms (e.g., BFGS, Nelder-Mead, L-BFGS-B) to find the parameter values that maximize the log-likelihood, subject to constraints: $\omega > 0$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta < 1$.

Modern software packages implement these procedures efficiently, often using analytical gradients to speed convergence.

Alternative Distributional Assumptions

While the Gaussian distribution is computationally convenient, financial returns often exhibit fat tails. Alternative distributions commonly used include:

- **Student's t-distribution**: Captures heavy tails with an additional degrees-of-freedom parameter.

- **Generalized Error Distribution (GED)**: Provides flexible tail behavior.

- **Skewed distributions**: Capture asymmetry in return distributions.

These are estimated using the same MLE framework with modified likelihood functions.

GARCH VARIANTS AND EXTENSIONS

The basic GARCH(1,1) model has spawned numerous extensions to capture additional stylized facts of financial volatility.

GJR-GARCH: Capturing Asymmetric Volatility

The GJR-GARCH model, developed by Glosten, Jagannathan, and Runkle, incorporates the so-called leverage effect. This is the empirical observation that negative returns (bad news) tend to increase volatility more than positive returns (good news) of the same magnitude.

The GJR-GARCH(1,1) variance equation is:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma I_{[\epsilon_{t-1}<0]}\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

where $I_{[\epsilon_{t-1}<0]}$ is an indicator function that equals 1 when $\epsilon_{t-1} < 0$ (negative shock) and 0 otherwise. The parameter γ captures the additional impact of negative shocks.

If $\gamma > 0$, negative shocks have a total effect of $(\alpha + \gamma)$ while positive shocks have effect α, capturing the asymmetry observed in equity markets.

EGARCH: Exponential GARCH

The Exponential GARCH (EGARCH) model uses a logarithmic specification for conditional variance, ensuring positivity without parameter constraints:

$$\ln(\sigma_t^2) = \omega + \alpha\frac{|\epsilon_{t-1}|}{\sigma_{t-1}} + \gamma\frac{\epsilon_{t-1}}{\sigma_{t-1}} + \beta\ln(\sigma_{t-1}^2)$$

This specification allows for asymmetric effects through the γ parameter and has different persistence dynamics.

ARCH-M: ARCH in Mean

In the ARCH-M (ARCH in Mean) model, the conditional variance directly enters the mean equation, allowing risk premia to vary with volatility:

$$r_t = \mu + \lambda\sigma_t + \epsilon_t$$

where λ captures the risk-return tradeoff. This is particularly useful in asset pricing applications where higher risk should command higher expected returns.

Other Extensions

Additional variants include:

- IGARCH: Integrated GARCH where $\alpha + \beta = 1$, implying infinite persistence.

- TARCH (Threshold GARCH): Similar to GJR-GARCH with threshold effects.

- CGARCH (Component GARCH): Separates volatility into permanent and transitory components.

- FIGARCH: Fractionally Integrated GARCH for long memory processes.

- Multivariate GARCH: Modeling covariance matrices for multiple assets simultaneously.

FORECASTING WITH GARCH: FROM THEORY TO PRACTICE

A key advantage of GARCH models is their ability to generate multi-period volatility forecasts with well-defined properties.

One-Step-Ahead Forecasting

The one-step-ahead variance forecast from time t to time $t+1$ is straightforward:

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

This forecast uses the most recent squared return ϵ_t^2 and the current variance estimate σ_t^2, both of which are known at time t.

Note that here, ϵ_t represents the realized innovation from the previous period (e.g., the actual return net of the mean) rather than the theoretical decomposition $\sigma_t z_t$ presented in the model specification. In other words, ϵ_t^2 incorporates both the volatility and the randomness that occurred at time t, and is directly observable.

Multi-Step-Ahead Forecasting

For forecasts further into the future, we use the recursive structure of GARCH. The h-step-ahead forecast is:

$$\sigma_{t+h}^2 = \overline{\sigma^2} + (\alpha+\beta)^{h-1}\left[\sigma_{t+1}^2 - \overline{\sigma^2}\right]$$

where $\overline{\sigma^2} = \omega/(1-\alpha-\beta)$ is the long-run variance.

This formula reveals several important properties:

- As $h \to \infty$, the forecast converges to the long-run variance: $\sigma_{t+h}^2 \to \overline{\sigma^2}$
- The rate of convergence depends on the persistence $(\alpha+\beta)$
- Higher persistence means slower convergence to the long-run level

A Numerical Example

Let's work through a concrete forecasting example with actual numbers.

Given Parameters:

- $\omega = 0.000002$
- $\alpha = 0.10$
- $\beta = 0.85$
- Current volatility: $\sigma_t = 2.0\%$ (so $\sigma_t^2 = 0.0004$)
- Yesterday's return: $r_t = -1.5\%$ (so $\epsilon_t = -0.015$, assuming $\mu \approx 0$)

Step 1: Calculate long-run variance

$$\overline{\sigma^2} = \frac{\omega}{1 - \alpha - \beta} = \frac{0.000002}{1 - 0.10 - 0.85} = \frac{0.000002}{0.05} = 0.00004$$

Long-run volatility: $\bar{\sigma} = \sqrt{0.00004} = 0.0063$ or 0.63%

Step 2: One-day-ahead forecast

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

$$\sigma_{t+1}^2 = 0.000002 + 0.10 \times (0.015)^2 + 0.85 \times 0.0004$$

$$\sigma_{t+1}^2 = 0.000002 + 0.0000225 + 0.00034 = 0.0003645$$

One-day-ahead volatility: $\sigma_{t+1} = \sqrt{0.0003645} = 0.0191$ or 1.91%

Step 3: Two-day-ahead forecast

For the two-day forecast, we need $E_t[\epsilon_{t+1}^2]$. Since we don't know the actual return, we use the expected value: $E_t[\epsilon_{t+1}^2] = \sigma_{t+1}^2$.

$$\sigma_{t+2}^2 = \omega + \alpha\sigma_{t+1}^2 + \beta\sigma_{t+1}^2 = \omega + (\alpha + \beta)\sigma_{t+1}^2$$

$$\begin{aligned}\sigma_{t+2}^2 &= 0.000002 + 0.95 \times 0.0003645 \\ &= 0.000002 + 0.00034628 \\ &= 0.00034828\end{aligned}$$

Two-day-ahead volatility: $\boldsymbol{\sigma_{t+2} = 0.0187}$ or 1.87%

Step 4: Using the multi-step formula

Alternatively, we can use the closed-form formula:

$$\sigma_{t+h}^2 = 0.00004 + (0.95)^{h-1} \times (0.0003645 - 0.00004)$$

For $\boldsymbol{h = 10}$ days:

$$\begin{aligned}\sigma_{t+10}^2 &= 0.00004 + (0.95)^9 \times 0.0003245 \\ &= 0.00004 + 0.6302 \times 0.0003245 \\ &= 0.00024457\end{aligned}$$

Ten-day-ahead volatility: $\boldsymbol{\sigma_{t+10} = 0.0156}$ or 1.56%

For $\boldsymbol{h = 30}$ days:

$$\begin{aligned}\sigma_{t+30}^2 &= 0.00004 + (0.95)^{29} \times 0.0003245 \\ &= 0.00004 + 0.2141 \times 0.0003245 \\ &= 0.00010948\end{aligned}$$

Thirty-day-ahead volatility: $\boldsymbol{\sigma_{t+30} = 0.0105}$ or 1.05%

Some observations from this exercise:

- The volatility forecast starts at 1.91% (elevated due to yesterday's large return).
- It gradually decays toward the long-run level of 0.63%.
- The high persistence (0.95) means this decay is slow.
- After 30 days, volatility is still above the long-run level but closer to it.
- Eventually, as $h \rightarrow \infty$, all forecasts converge to 0.63%.

Forecasting Returns and Prices

While GARCH directly forecasts volatility, it can be used to generate scenarios for future returns through simulation following these steps:

1. Estimate the GARCH model on historical data to obtain $\hat{\omega}$, $\hat{\alpha}$, $\hat{\beta}$.
2. Generate volatility forecasts using the recursive formula.
3. Simulate future returns. Draw random standardized innovations $z_{t+h} \sim N(0,1)$ and compute:

$$r_{t+h} = \mu + \sigma_{t+h} z_{t+h}$$

4. Construct price paths. If P_t is the current price, future prices can be simulated as:

$$P_{t+h} = P_t \exp\left(\sum_{i=1}^{h} r_{t+i}\right)$$

This Monte Carlo approach generates distributions of future returns and prices that reflect the time-varying volatility structure captured by GARCH.

WHY GARCH DOMINATES THE INDUSTRY

GARCH models have become the backbone of volatility modeling in finance for several reasons:

- Empirical Success: GARCH(1,1) consistently provides accurate out-of-sample volatility forecasts across asset classes, time periods, and markets.
- Parsimony: The ability to capture complex volatility dynamics with just three parameters makes GARCH computationally efficient and reduces overfitting risks.
- Statistical Rigor: The formal likelihood framework enables rigorous model testing, diagnostics, and inference.
- Flexibility: The numerous GARCH variants (GJR, EGARCH, CGARCH, etc.) allow practitioners to tailor models to specific market characteristics while maintaining the core framework.
- Industry Adoption: GARCH is embedded in risk management systems, regulatory frameworks (Basel accords), and trading platforms worldwide.
- Integration with Other Models: GARCH volatility estimates feed into option pricing models, Value-at-Risk calculations, portfolio optimization, and derivative hedging strategies.
- Captures Stylized Facts: GARCH naturally generates the key empirical features observed in financial data: volatility clustering, mean reversion, and leptokurtosis.

The subtitle of this paper "Capturing the Memory of Volatility in Interest Rates and Asset Price Returns" was chosen because GARCH models are fundamentally about memory. The β parameter explicitly models how volatility remembers its past values, while the α parameter captures how volatility remembers past shocks. Together, they create a model that chronicles the market's emotional memory, indicating its tendency to remain anxious after stress and calm after tranquility.

CONCLUSION

The GARCH family of models represents a successful application of econometric theory to practical financial problems. By providing a parsimonious, statistically rigorous framework for modeling time-varying volatility, GARCH has become indispensable for many market practitioners.

The elegance of GARCH lies in its simplicity. The three parameters (ω, α, and β) work together to capture the persistent and mean-reverting nature of financial volatility. Through their interaction in the recursive GARCH equation, these parameters generate the volatility clustering phenomenon that defines financial markets.

The model's success stems not only from its theoretical elegance but also from its empirical performance. GARCH consistently outperforms simpler alternatives like EWMA while remaining computationally tractable. Its extensibility to handle asymmetric effects, regime changes, and multivariate dynamics ensures its continued relevance as markets evolve.

As financial markets grow more complex and interconnected, the need for accurate volatility forecasting becomes more critical. GARCH models, having captured the memories of interest rates, equity returns, exchange rates, and commodity prices for nearly four decades, remain at the forefront of this challenge. They are not merely statistical tools but rather sophisticated memory systems that distill the market's accumulated experience into forward-looking risk assessments.