








Statistics of Financial Markets

Chapter 9: ARCH and GARCH models



Daniel Traian PELE

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

Acronyms used in this material (1/2)

- **ACF**: Autocorrelation Function
- **AI**: Artificial Intelligence
- **AIC**: Akaike Information Criterion
- **AR**: AutoRegressive (model); Abnormal Return
- **ARCH**: AutoRegressive Conditional Heteroskedasticity
- **ARCH-LM**: ARCH Lagrange Multiplier (test)
- **ARMA**: AutoRegressive Moving Average (model)
- **BET**: Bucharest Exchange Trading (index)
- **BIC**: Bayesian Information Criterion
- **BVB**: Bursa de Valori București (the Bucharest Stock Exchange)
- **COVID-19**: COronaVIrus Disease 2019
- **DM**: Diebold–Mariano (test)
- **EGARCH**: Exponential GARCH
- **EODHD**: EOD Historical Data (market-data provider)
- **ES**: Expected Shortfall
- **ETF**: Exchange-Traded Fund
- **EWMA**: Exponentially Weighted Moving Average
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity

Acronyms used in this material (2/2)

- **GJR**: Glosten–Jagannathan–Runkle (asymmetric GARCH)
- **GJR-GARCH**: Glosten–Jagannathan–Runkle GARCH
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **IGARCH**: Integrated GARCH
- **LB**: Ljung–Box (test)
- **LR**: Likelihood Ratio (test)
- **MLE**: Maximum Likelihood Estimation
- **NIC**: News Impact Curve
- **OMV**: Österreichische Mineralölverwaltung — Austrian Mineral Oil Administration (OMV group)
- **QLIKE**: Quasi-LIKElihood (loss function)
- **QMLE**: Quasi-Maximum Likelihood Estimation
- **QQ**: Quantile–Quantile (plot)
- **RW1**: Random Walk 1 (i.i.d. increments)
- **RW3**: Random Walk 3 (uncorrelated increments)
- **SE**: Standard Error
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook)

Today's Question and Route

- **Question**: if tomorrow's return cannot be predicted, can tomorrow's **risk** be predicted?
  - Chapter 7: daily returns are almost uncorrelated; Chapter 8: their squares are strongly correlated (volatility clustering)
  - this chapter turns volatility clustering into a model for the conditional variance
- **Route** of the chapter
  - conditional mean and conditional variance; ARCH(q) and GARCH(1,1)
  - estimation by maximum likelihood; Student-t and skewed-t innovations
  - six markets; asymmetry: GJR-GARCH, EGARCH, the news impact curve
  - diagnostics, model choice, forecasts, VaR 1%

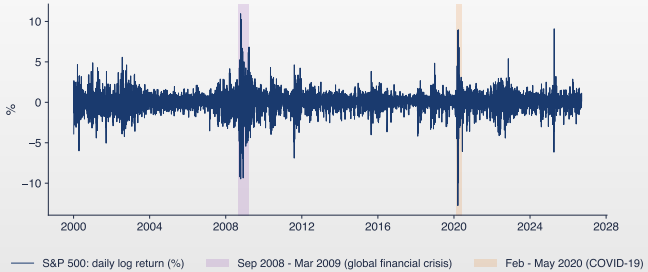
Learning Outcomes

- Distinguish the conditional variance from the unconditional variance and explain why returns can be uncorrelated yet dependent
- Write the ARCH(q) and GARCH(1,1) models and compute the persistence, the unconditional variance and the half-life
- Estimate GARCH models by maximum likelihood, read the standard errors and choose the innovation distribution
- Measure the leverage effect with GJR-GARCH, EGARCH and the news impact curve
- Check a fitted model with standardised residuals and choose among models with AIC and BIC
- Forecast volatility over several horizons, compare forecasts with QLIKE and compute a one-day VaR 1%

Reading and Tools

- Textbook: Franke, Härdle & Hafner (2019), Ch. 13 (13.1 ARCH and GARCH models, 13.2 extensions of the GARCH model)
  - exercises with solutions: Borak, Härdle & López-Cabrera (2013), Ch. 13
- Conditional heteroskedastic models: Tsay (2010), Ch. 3
- Python Quantlets of this chapter: Quantlets/Ch_09
  - ported from the SFE Quantlets SFEtimegarch, SFElikarch1, SFElikgarch, SFEgarchest, SFEvolgarchest and SFENewsImpactCurve
  - estimation with the Python package `arch`
- Lecture notebook: open in Google Colab
- Video course: Applied Time Series Analysis with Python

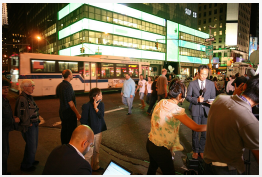
S&P 500: Calm Years and Turbulent Periods



- Daily log returns $r_t = 100(\ln P_t - \ln P_{t-1})$, in %, 6 714 days to 18 September 2026; the largest fall: −12.8% on 16 March 2020; the largest rise: 11.0% on 13 October 2008
- Standard deviation: 0.42% in 2017, 3.52% from September 2008 to March 2009, 3.70% from February to May 2020
- Large changes follow large changes, of either sign: volatility clustering Engle (1982)

SFM_ch9_volatility_markets

The Crises of 2008 and 2020



Lehman Brothers headquarters, New York, 15 September 2008

Photo: Robert Scoble (2008); CC BY 2.0; Wikimedia Commons



The Financial District of New York, 25 March 2020

Photo: Billie Grace Ward (2020); CC BY 2.0; Wikimedia Commons

- 15 September 2008: Lehman Brothers files for bankruptcy; March 2020: the COVID-19 pandemic closes economies
- In both episodes the daily standard deviation of the S&P 500 was 8.3 and 8.8 times that of 2017; calm returned only after months

Motivation for Modelling the Conditional Variance

- **Risk management**: VaR (value at risk) and ES (expected shortfall) need tomorrow's volatility, not the average volatility of the last 20 years
  - a constant-variance VaR is too large in calm years and too small in crises (Chapter 10)
- **Portfolios**: weights depend on variances and covariances, which change over time
- **Derivatives**: option prices depend on the volatility expected until maturity
- **Inference**: tests and confidence intervals that assume a constant variance can be wrong
  - this is why Chapter 7 needed tests that remain valid under heteroskedasticity

Stylised Facts a Volatility Model Must Reproduce

- **Volatility clustering**: the ACF (autocorrelation function) of r_t^2 is positive and decays slowly (Chapter 8)
  - S&P 500: autocorrelation of r_t^2 at lag 1: 0.31; at lag 50: still 0.053
- **Heavy tails**: kurtosis above 3, the value of the Normal distribution (Chapter 2)
  - S&P 500 daily returns: kurtosis 13.7
- **Leverage effect**: falls raise volatility more than rises of the same size Christie (1982)
- **Mean reversion of volatility**: after a storm, volatility returns to a long-run level
- Almost no autocorrelation in r_t itself (Chapter 7)

Conditioning on the Past

- $\mathcal{F}_{t-1}$: the information available at the end of day $t-1$ (past returns)
  - conditional expectation $E[X \mid \mathcal{F}_{t-1}]$: the best forecast of X given the past (Chapter 4)
- **Conditional mean**: $\mu_t = E[r_t \mid \mathcal{F}_{t-1}]$
  - Chapter 7: for daily returns, μ_t is almost constant: $\mu_t = \mu$
- **Conditional variance**: $\sigma_t^2 = \mathrm{Var}(r_t \mid \mathcal{F}_{t-1}) = E[(r_t - \mu_t)^2 \mid \mathcal{F}_{t-1}]$
  - σ_t is the **conditional volatility**: known at $t-1$, it changes from day to day
  - **heteroskedasticity**: a variance that is not constant

The Basic Decomposition

- Every model of this chapter has the form $r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t$
  - z_t: the **innovations**, i.i.d. with mean 0 and variance 1 (Normal, Student-t, ...)
  - ε_t: the **shock** (the return minus its mean); σ_t depends only on the past
- Consequences
  - $E[\varepsilon_t \mid \mathcal{F}_{t-1}] = \sigma_t E[z_t] = 0$: the shocks are a martingale difference, hence uncorrelated
  - $\mathrm{Var}(\varepsilon_t \mid \mathcal{F}_{t-1}) = \sigma_t^2$: the variance is predictable
  - the ε_t are uncorrelated but not independent: ε_t^2 is correlated with ε_{t-1}^2
- In the terms of Chapter 7: RW1 (i.i.d. increments) is rejected, RW3 (uncorrelated increments) is kept

Worked Example: Conditional and Unconditional Variance

- A market has calm days ($\sigma_t^2 = 0.5$) and stormy days ($\sigma_t^2 = 4.5$), each with probability $1/2$; the mean is 0
  - law of total variance: $\mathrm{Var}(r_t) = E[\mathrm{Var}(r_t \mid \mathcal{F}_{t-1})] + \mathrm{Var}(E[r_t \mid \mathcal{F}_{t-1}])$
  - step by step: $0.5 \times 0.5 + 0.5 \times 4.5 + 0 = 2.50$, so the unconditional standard deviation is 1.58%
- The unconditional variance is an average; the conditional variance tells which day we are in
  - a calm day: $\sigma_t = 0.71\%$; a stormy day: $\sigma_t = 2.12\%$; the average 1.58% fits neither
- A mixture of Normal distributions with different variances has heavy tails: conditional heteroskedasticity creates kurtosis

1982: Robert Engle and ARCH

- Engle (1982): **ARCH** (autoregressive conditional heteroskedasticity): today's variance depends on yesterday's squared shocks
  - first application: the variance of inflation in the United Kingdom; written while Engle was visiting the London School of Economics
- Nobel Prize 2003: "for methods of analyzing economic time series with time-varying volatility (ARCH)"
  - shared with Clive Granger (cointegration); Nobel lecture: Engle (2004)



Robert F. Engle (2022)

Photo: Jon Demske (2022); CC BY-SA 4.0; Wikimedia Commons



Stockholm Concert Hall, venue of the Nobel Prize ceremony

Photo: Karen Zhou (2025); CC BY-SA 4.0; Wikimedia Commons

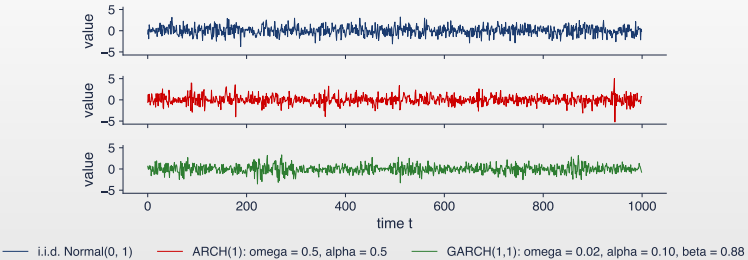
ARCH(1): Definition

- $\varepsilon_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2, \quad \omega > 0,\ \alpha \geq 0$
  - in words: today's variance = a constant plus a share α of yesterday's squared shock
  - ω: the variance after a zero shock; α: the reaction to yesterday's shock; both keep the variance positive
  - a large shock yesterday, of either sign, raises today's variance
- "Autoregressive": ε_t^2 follows an AR(1) (autoregressive) model
  - write $\varepsilon_t^2 = \sigma_t^2 + v_t$ with $v_t = \sigma_t^2(z_t^2 - 1)$, $E[v_t \mid \mathcal{F}_{t-1}] = 0$
  - then $\varepsilon_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + v_t$: an AR(1) for the squared shocks
- Hence the ACF of ε_t^2 is $\rho(k) = \alpha^k$ when the fourth moment exists: clustering, but a fast decay

ARCH(1): Properties

- Mean and autocorrelation: $E[\varepsilon_t] = 0$, $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$ for $k \neq 0$
- **Unconditional variance** (if $\alpha < 1$): take expectations in $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2$
  - $\bar{\sigma}^2 = E[\varepsilon_t^2] = \omega + \alpha\bar{\sigma}^2$, so $\bar{\sigma}^2 = \omega/(1-\alpha)$
- **Kurtosis** with Normal z_t (if $3\alpha^2 < 1$): $K = \dfrac{E[\varepsilon_t^4]}{(E[\varepsilon_t^2])^2} = 3\,\dfrac{1-\alpha^2}{1-3\alpha^2} > 3$
  - Normal innovations, but heavy-tailed returns Franke, Härdle & Hafner (2019)
  - if $\alpha \geq 1/\sqrt{3} = 0.577$, the fourth moment is infinite
- **Worked example**: $\omega = 0.5$, $\alpha = 0.5$
  - $\bar{\sigma}^2 = 0.5/0.5 = 1.00$; $K = 3 \times 0.75/0.25 = 9.0$; after a shock $\varepsilon_{t-1} = 2$: $\sigma_t^2 = 0.5 + 0.5 \times 4 = 2.50$, $\sigma_t = 1.58$

Simulated Paths: i.i.d., ARCH(1) and GARCH(1,1)



- Three series of 1000 values with unconditional variance 1; ARCH(1) with $\omega = \alpha = 0.5$, GARCH(1,1) with $\omega = 0.02$, $\alpha = 0.10$, $\beta = 0.88$ (next section)
- Sample kurtosis: i.i.d. 3.12, ARCH 4.91, GARCH 3.74; theory: 3, 9.0, 6.06; the fourth moment converges slowly
- ARCH(1) gives isolated spikes; GARCH(1,1) gives long calm and long agitated periods, as in real returns

SFM_ch9_simulation_likelihood

ARCH(q) and Its Limits

- **ARCH(q)**: $\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2$, $\alpha_i \geq 0$; stationary if $\sum_i \alpha_i < 1$
  - q: the number of lags; α_i: the reaction to the shock of i days ago
  - the effect of a shock lasts exactly q days; real clustering lasts for months
- Testing for ARCH effects: the ARCH-LM test of Engle (regression of $\hat{\varepsilon}_t^2$ on its lags), Chapter 8

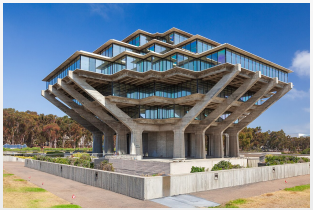
S&P 500, Normal z_t	parameters	log-likelihood	AIC	BIC	$\sum \alpha_i$ $(+\beta)$
ARCH(1)	3	−10305.5	20617.0	20637.5	0.394
ARCH(5)	7	−9340.5	18695.1	18742.7	0.816
ARCH(10)	12	−9227.6	18479.3	18561.0	0.849
GARCH(1,1)	4	−9218.0	**18444.0**	**18471.2**	0.980

- Interpretation: ARCH needs ten lags and still loses to GARCH(1,1), which has four parameters (AIC and BIC: smaller is better, Section 10)

SFM_ch9_garch_estimation

1986: Tim Bollerslev and GARCH

- Bollerslev (1986), a doctoral student of Engle at the University of California, San Diego: add yesterday's variance to the equation
  - **GARCH** (generalised ARCH): a few parameters replace a long ARCH(q)
- Student-t innovations for returns: Bollerslev (1987)
- GARCH(1,1) is still the benchmark: Hansen & Lunde (2005) compared 330 models and found it hard to beat on exchange rates



Geisel Library, University of California, San Diego

Photo: Stephen Bay (2010); CC BY 4.0; Wikimedia Commons

GARCH(1,1): Definition

- $r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha\,\varepsilon_{t-1}^2 + \beta\,\sigma_{t-1}^2$
  - $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$: the variance stays positive
- Reading the parameters
  - α: the **reaction** to yesterday's news (size of the jump after a shock)
  - β: the **memory** of the variance (how much of yesterday's level is kept)
  - $\alpha + \beta$: the **persistence**; ω sets the long-run level
- By recursion, an ARCH(∞) with geometric weights
  - $\sigma_t^2 = \dfrac{\omega}{1-\beta} + \alpha \sum_{j=0}^{\infty} \beta^j\, \varepsilon_{t-1-j}^2$: all past shocks matter, with decaying weights
  - the shock of $j+1$ days ago gets the weight $\alpha\beta^j$

Stationarity, Long-Run Variance and Kurtosis

- As for ARCH: $\varepsilon_t^2 = \omega + (\alpha+\beta)\varepsilon_{t-1}^2 + v_t - \beta v_{t-1}$, an ARMA(1,1) for the squared shocks
  - $v_t = \varepsilon_t^2 - \sigma_t^2$: the surprise in the squared shock, as for ARCH(1)
  - ARMA: autoregressive moving average; the ACF of ε_t^2 decays like $(\alpha+\beta)^k$ at lag k
- **Covariance stationary** if $\alpha+\beta<1$; then the **unconditional (long-run) variance** is
  - $\bar{\sigma}^2 = E[\sigma_t^2] = \dfrac{\omega}{1-\alpha-\beta}$, annualised volatility $\sqrt{252\,\bar{\sigma}^2}$ for 252 trading days
- Kurtosis with Normal z_t, if $(\alpha+\beta)^2 + 2\alpha^2 < 1$: $K = 3\,\dfrac{1-(\alpha+\beta)^2}{1-(\alpha+\beta)^2-2\alpha^2} > 3$
  - $\alpha = 0.10$, $\beta = 0.88$: $(\alpha+\beta)^2 + 2\alpha^2 = 0.9804$, $K = 6.06$

Persistence and Half-Life

- Forecast of the variance h days ahead (derived in Section 11): $E[\sigma^2_{t+h} \mid \mathcal{F}_t] - \bar{\sigma}^2 = (\alpha+\beta)^{h-1}(\sigma^2_{t+1} - \bar{\sigma}^2)$
  - $E[\cdot \mid \mathcal{F}_t]$: the forecast made at the end of day t; $\bar{\sigma}^2$: the long-run variance
  - a deviation from the long-run variance shrinks by the factor $\alpha+\beta$ each day
- **Half-life**: the number of days after which half of a variance shock is gone
  - $(\alpha+\beta)^h = 1/2 \;\Rightarrow\; h_{1/2} = \dfrac{\ln 0.5}{\ln(\alpha+\beta)}$
  - $\alpha+\beta = 0.90$: 6.6 days; 0.98: 34.3 days; 0.99: 69.0 days: small changes near 1 matter a lot
- Daily equity returns: α around 0.05–0.15, β around 0.80–0.93, $\alpha+\beta$ close to 1 (Section 7)

Worked Example: One GARCH(1,1) Step

- Daily returns in %: $\omega = 0.02$, $\alpha = 0.10$, $\beta = 0.88$, $\mu = 0$
- Long-run quantities
  - persistence $\alpha + \beta = 0.98$; $\bar{\sigma}^2 = 0.02/(1 - 0.98) = 1.00$; annualised volatility $\sqrt{252 \times 1.00} = 15.9\%$
  - half-life $\ln 0.5 / \ln 0.98 = 34.3$ days
- Today: $\sigma_t^2 = 1.2$ and the return is $r_t = -3\%$
  - tomorrow: $\sigma_{t+1}^2 = 0.02 + 0.10 \times 9 + 0.88 \times 1.2 = 1.976$, $\sigma_{t+1} = 1.41\%$
  - the shock raised the variance from 1.2 to almost 2; it then decays towards 1.00 with the factor 0.98 per day

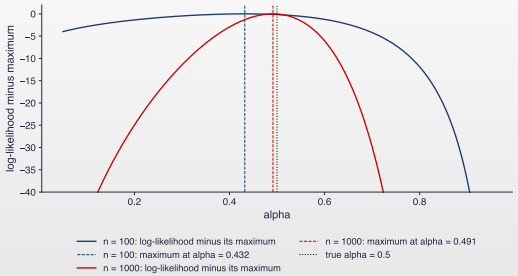
IGARCH and the EWMA Link

- **IGARCH** (integrated GARCH) Engle & Bollerslev (1986): $\alpha + \beta = 1$
  - shocks to the variance never die out: no half-life, no finite unconditional variance
  - with $\omega = 0$ the forecast of the variance is the same for every horizon; with $\omega > 0$ it grows by ω per day
- **EWMA** (exponentially weighted moving average, Chapter 8): $\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2$
  - this is IGARCH with $\omega = 0$, $\alpha = 1 - \lambda$, $\beta = \lambda$; RiskMetrics uses $\lambda = 0.94$ for daily data
  - EWMA: nothing to estimate, but no mean reversion; GARCH: three parameters, and volatility returns to $\bar{\sigma}$
- **Question for the room**: an analyst forecasts the volatility one year ahead with EWMA right after a crash. What goes wrong?
  - **Answer**: EWMA keeps today's crisis level for the whole year; GARCH lets it decay towards the long-run level

The Likelihood of a GARCH Model

- The returns are dependent, so we factor the joint density into conditional densities
  - f: a probability density; T: the number of returns; θ: the vector of parameters; ℓ: the log-likelihood
  - $f(r_1, \ldots, r_T) = f(r_1) \prod_{t=2}^{T} f(r_t \mid \mathcal{F}_{t-1})$, and $r_t \mid \mathcal{F}_{t-1} \sim N(\mu, \sigma_t^2)$ for Normal innovations
- **Conditional log-likelihood** with $\theta = (\mu, \omega, \alpha, \beta)$:
  - $\ell(\theta) = -\frac{1}{2} \sum_{t=2}^{T} \left[\ln(2\pi) + \ln \sigma_t^2(\theta) + \frac{(r_t - \mu)^2}{\sigma_t^2(\theta)} \right]$
  - in words: each day adds a penalty for a large σ_t^2 and a penalty for a squared error that is large relative to σ_t^2
  - $\sigma_t^2(\theta)$ is computed recursively from a starting value σ_1^2 (for example the sample variance)
- **MLE** (maximum likelihood estimation): $\hat{\theta} = \arg\max_\theta \ell(\theta)$, under $\omega > 0$, $\alpha, \beta \geq 0$, $\alpha + \beta < 1$
  - no closed form: a numerical optimiser searches for the maximum Franke, Härdle & Hafner (2019)

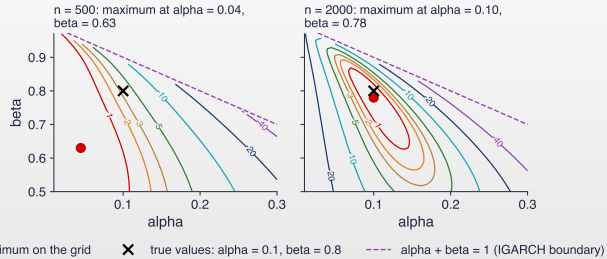
The Likelihood of ARCH(1)



- Simulated ARCH(1) with $\omega = \alpha = 0.5$; $\ell(\alpha)$ with $\omega = 1 - \alpha$, as in SFElikarch1; each curve minus its maximum
- $n = 100$: $\hat{\alpha} = 0.432$ (standard error 0.127); $n = 1000$: $\hat{\alpha} = 0.491$ (0.035)
- Interpretation: more data make the curve sharper; its curvature at the maximum gives the standard error

SFM_ch9_simulation_likelihood

The Likelihood of GARCH(1,1)



- Simulated GARCH(1,1), $\omega = 0.1$, $\alpha = 0.1$, $\beta = 0.8$; ω set so that the unconditional variance equals the sample variance, as in SFElikgarch; contours: log-likelihood minus its maximum
- $n = 500$: maximum at $(0.045, 0.63)$, far from the truth; $n = 2000$: $(0.100, 0.78)$
- Interpretation: a long, flat ridge along which α and β trade off: GARCH needs long samples

SFM_ch9_simulation_likelihood

Estimation Step by Step and with a Package

S&P 500, GARCH(1,1)-N	step by step (SE)	`arch`	classic SE	robust SE
μ	0.0614 (0.0098)	0.0615	0.0098	0.0100
ω	0.0253 (0.0030)	0.0254	0.0030	0.0048
α	0.1205 (0.0086)	0.1206	0.0086	0.0118
β	0.8600 (0.0091)	0.8599	0.0091	0.0125

- Step by step: write $-\ell(\theta)$ in Python and minimise it with `scipy.optimize.minimize` under $\alpha + \beta < 1$
  - the two routes differ only in the start-up of σ_t^2 and in the optimiser; log-likelihood -9219.9 against -9218.0
- SE (standard error), classic: from the inverse Hessian (the curvature of ℓ); robust: valid when z_t is not Normal Bollerslev & Wooldridge (1992)
- Interpretation: the robust SE of α and β are about 1.4 times larger (1.6 for ω, 1.02 for μ): with heavy-tailed z_t the classic SE overstate the precision of the variance parameters

SFM_ch9_garch_estimation

Quasi-Maximum Likelihood

- **QMLE** (quasi-maximum likelihood estimation): maximise the Normal likelihood even if z_t is not Normal
  - the estimates are still consistent if μ and σ_t^2 are correctly specified Bollerslev & Wooldridge (1992)
  - but the classic SE are wrong: use the robust ("sandwich") SE
- Practical rules
  - report robust SE (the default of `arch`)
  - use at least 1000–2000 daily observations; returns in % avoid numerical problems with tiny ω
  - an estimate on the boundary ($\alpha + \beta = 1$) has no usual standard error: report it as IGARCH
- A better innovation distribution gives more efficient estimates: next section

Student-t and Skewed-t Innovations

- GARCH with Normal z_t explains only part of the kurtosis: S&P 500 returns 13.7, standardised residuals $\hat{z}_t = \hat{\varepsilon}_t/\hat{\sigma}_t$: 4.7
  - the remaining heavy tails must come from the distribution of z_t
- **Standardised Student-t** with $\nu > 2$ degrees of freedom (Chapter 2) Bollerslev (1987)
  - t_ν: a Student-t variable with ν degrees of freedom; $z_t = t_\nu \sqrt{(\nu-2)/\nu}$ has variance 1
  - a small ν means heavy tails: kurtosis $3 + 6/(\nu - 4)$ for $\nu > 4$; $\nu \to \infty$: the Normal distribution
  - ν is estimated together with $(\mu, \omega, \alpha, \beta)$
- **Skewed t** Hansen (1994): one more parameter $\lambda \in (-1, 1)$; $\lambda < 0$: a longer left tail
  - nested models: Normal $\subset$ t $\subset$ skewed t; compare them with likelihood-ratio tests or AIC/BIC (Chapter 6)

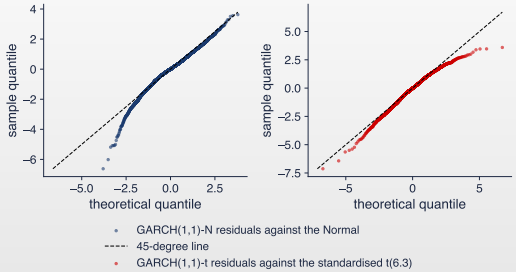
S&P 500: Three Innovation Distributions

GARCH(1,1)	ω	α	β	ν	λ	log-lik.	$\alpha+\beta$
Normal	0.0254	0.121	0.860	–	–	−9218.0	0.980
Student-t	0.0157	0.123	0.872	6.26	–	−9062.5	0.995
skewed t	0.0153	0.120	0.872	6.92	−0.114	−9038.4	0.993

- Likelihood-ratio statistic $2(\ell_1 - \ell_0)$ (ℓ_1, ℓ_0: maximised log-likelihoods of the larger and the smaller model): t against Normal 311.0 (1 restriction, $\chi^2_{0.95}(1) = 3.84$); skewed t against t 48.1: both reject the smaller model
- $\hat{\nu} = 6.26$: tails far heavier than Normal; $\hat{\lambda} = -0.114$: large falls more frequent than large rises
- Interpretation: α and β hardly change; the innovation distribution matters for tail quantiles (VaR), less for σ_t

SFM_ch9_garch_estimation

QQ Plots of the Standardised Residuals



- Left: residuals of the Normal GARCH against the Normal distribution: the 0.1% quantile is -4.81, against -3.09 in theory
- Right: residuals of GARCH-t against the standardised t with $\hat{\nu} = 6.3$: -4.79 against -4.19: much closer; the largest falls are still deeper
- Interpretation: the left tail is the problem: a reason for skewed innovations and for asymmetric models (Section 8)

SFM_ch9_garch_estimation

GARCH(1,1)-t Estimates

	from	α	β	ν	$\alpha+\beta$	half-life	long-run vol.	sample vol.
S&P 500	2000	0.123	0.872	6.26	0.9947	131	27.3	19.2
DAX	2000	0.097	0.895	7.08	0.9922	88	25.9	22.4
BET	2000	0.192	0.799	5.01	0.9907	74	34.7	22.1
Bitcoin	2014	0.109	0.891	3.18	1.000	–	–	66.6
TLV	2010	0.139	0.824	4.91	0.9630	18	28.2	26.5
SNP	2010	0.135	0.807	4.81	0.9425	12	25.6	24.5

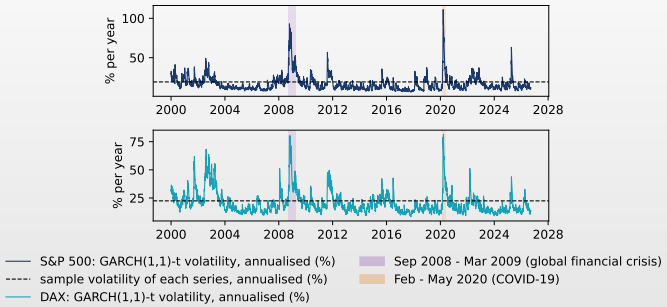
- Daily log returns in % to 18 September 2026; robust SE of $\hat{\alpha}$ and $\hat{\beta}$ between 0.010 and 0.041; half-life in days; volatilities annualised with the actual number of observations per year, in %
- Indices: $\alpha+\beta$ between 0.9907 and 0.9947, half-lives of 74–131 trading days; BVB stocks: shorter memory (TLV 18 days, SNP 12 days)
- BET: the largest α (0.192): a stronger reaction to news; all $\hat{\nu}$ between 3.2 and 7.1: heavy tails everywhere

SFM_ch9_volatility_markets

Interpretation: Bitcoin and the Long-Run Volatility

- Bitcoin: $\hat{\alpha} + \hat{\beta} = 1.0000$, on the boundary $\alpha + \beta < 1$: an IGARCH
  - no half-life and no long-run volatility: the model behaves like an EWMA with $\lambda \approx 0.891$
  - $\hat{\nu} = 3.18$: extremely heavy tails (the fourth moment of z_t does not exist)
- **Question for the room**: S&P 500: long-run volatility 27.3%, sample volatility 19.2%. Is the model wrong?
  - **Answer**: not necessarily: $\bar{\sigma}^2 = \omega/(1 - \alpha - \beta)$ divides by $1 - \alpha - \beta = 0.0053$, a small and imprecise number
  - the Normal GARCH gives 18.1%; the long-run level is the least reliable output of a persistent GARCH

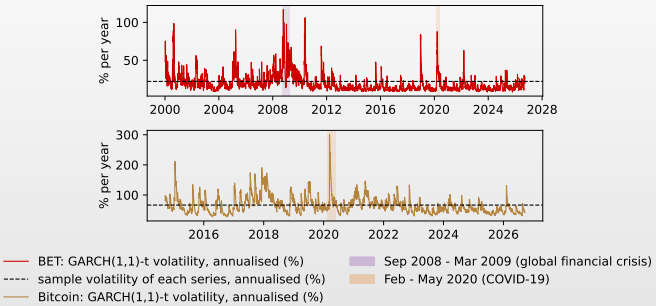
Conditional Volatility: S&P 500 and DAX



- GARCH(1,1)-t volatility $\hat{\sigma}_t$, annualised, as in SFEvolgarchest; dashed: the sample volatility
- S&P 500: peaks 93% (16 October 2008) and 111% (17 March 2020); median 14%; on 18 September 2026: 12%

SFM_ch9_volatility_markets

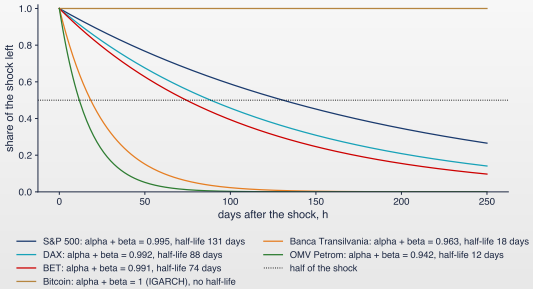
Conditional Volatility: BET and Bitcoin



- BET: the highest peak in 2008 (117%, 15 October 2008), lower in 2020 (88%); Bitcoin: 301% on 13 March 2020, median 59%
- Interpretation: the 2020 shock hit all markets in the same week; the median volatility of Bitcoin is 4.1 times that of the S&P 500 (14%)

SFM_ch9_volatility_markets

Duration of a Volatility Shock



- Share of a variance shock left after h days: $(\alpha + \beta)^h$; the half-life is where the curve crosses 1/2
- Interpretation: after a crisis, the S&P 500 needs about half a year to halve the excess variance; OMV Petrom about two weeks; Bitcoin never

SFM_ch9_volatility_markets

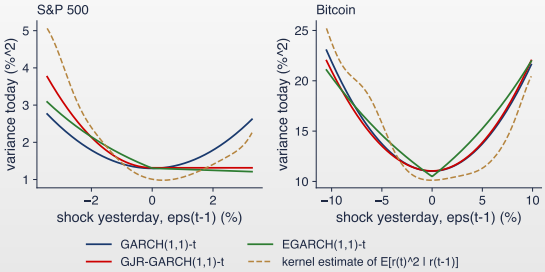
The Leverage Effect

- The leverage effect of Section 1 Christie (1982) has two explanations
  - **leverage** explanation: a lower share price raises the debt-to-equity ratio, so the equity becomes riskier
  - **volatility feedback**: news of higher risk lowers prices at once
- GARCH(1,1) cannot see it: σ_t^2 depends on ε_{t-1}^2, not on the sign of ε_{t-1}
  - the QQ plot of Section 6 and the skewed t ($\hat{\lambda} < 0$) already pointed to the left tail
- **NIC** (news impact curve) Engle & Ng (1993): σ_t^2 as a function of ε_{t-1}, with σ_{t-1}^2 fixed at the unconditional variance

GJR-GARCH and EGARCH

- **GJR-GARCH** Glosten, Jagannathan & Runkle (1993): $\sigma_t^2 = \omega + (\alpha + \gamma\, I_{t-1})\, \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$, $I_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, else 0
  - I_{t-1}: an indicator of a negative shock yesterday; γ: the extra reaction to bad news
  - good news: slope α; bad news: $\alpha + \gamma$; leverage effect: $\gamma > 0$
  - persistence $\alpha + \beta + \gamma/2$ for symmetric z_t (half of the shocks are negative)
- **EGARCH** (exponential GARCH) Nelson (1991): $\ln \sigma_t^2 = \omega + \alpha\big(|z_{t-1}| - E|z_{t-1}|\big) + \gamma z_{t-1} + \beta \ln \sigma_{t-1}^2$
  - $|z_{t-1}| - E|z_{t-1}|$: the size of yesterday's standardised shock above its average ($E|z| = \sqrt{2/\pi}$ for Normal z)
  - the logarithm keeps $\sigma_t^2 > 0$ without sign restrictions on the parameters
  - leverage effect: $\gamma < 0$ (a negative z_{t-1} raises $\ln \sigma_t^2$); persistence: β
- Textbook treatment of both, and of the threshold GARCH: Franke, Härdle & Hafner (2019), Section 13.2

News Impact Curves: S&P 500 and Bitcoin



- Model curves: σ_t^2 against ε_{t-1}, σ_{t-1}^2 fixed at the sample variance; dashed: a kernel (Nadaraya–Watson) estimate of $E[r_t^2 \mid r_{t-1}]$, the idea of SFENewsImpactCurve
- S&P 500, GJR: after a fall of 2%, tomorrow's variance is 1.62 times the variance after a rise of 2%; Bitcoin: ratio 1.00, a symmetric curve

SFM_ch9_asymmetry

Asymmetry in Six Markets

	GJR α	GJR γ	$t(\gamma)$	LR	p	EGARCH γ	$t(\gamma)$
S&P 500	0.000	0.205	10.0	235.7	<0.001	-0.165	-15.2
DAX	0.000	0.163	8.7	204.5	<0.001	-0.133	-11.8
BET	0.172	0.043	2.1	4.9	0.026	-0.028	-2.6
Bitcoin	0.113	-0.013	-0.7	0.7	0.416	0.015	1.2
TLV	0.101	0.065	2.7	7.8	0.005	-0.040	-3.0
SNP	0.117	0.033	1.4	1.7	0.188	-0.027	-1.9

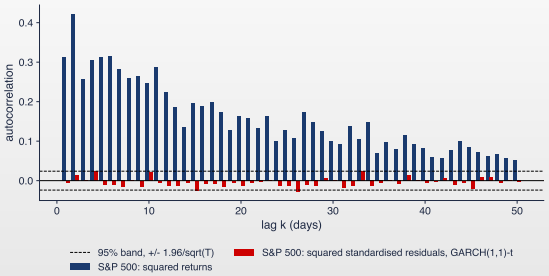
- Student-t innovations; t statistics with robust SE; LR (likelihood ratio) $= 2(\ell_{GJR} - \ell_{GARCH}) \sim \chi^2(1)$
- S&P 500 and DAX: GJR $\hat{\alpha} = 0$: only bad news raises volatility; a very strong leverage effect
- BET and TLV: a weak effect, significant by LR, yet BIC prefers the symmetric GARCH; SNP: not significant at 5%; Bitcoin: none (EGARCH γ slightly positive, not significant)
- Interpretation: leverage is a property of mature equity markets; in crypto, large rises are as "frightening" as large falls

SFM_ch9_asymmetry

Checking a Fitted Model

- If the model is right, the **standardised residuals** $\hat{z}_t = (r_t - \hat{\mu})/\hat{\sigma}_t$ are close to i.i.d. with mean 0 and variance 1
  - **LB** (Ljung–Box) $Q(10)$ on $\hat{z}_t$: no autocorrelation left in the mean (Chapter 7) Ljung & Box (1978)
  - $Q(10)$ on $\hat{z}_t^2$ and the ARCH-LM test: no ARCH effects left (Chapter 8)
  - QQ plot of $\hat{z}_t$: is the assumed distribution right? (Section 6)
- **Sign-bias test** Engle & Ng (1993): regress $\hat{z}_t^2$ on I_{t-1}, $I_{t-1}\hat{\varepsilon}_{t-1}$ and $(1 - I_{t-1})\hat{\varepsilon}_{t-1}$
  - joint test $TR^2 \sim \chi^2(3)$: a rejection means the sign of past shocks still predicts the variance (asymmetry is missing)
- A passed diagnostic does not prove the model right; a failed one shows where it is wrong

Autocorrelations Removed by GARCH



- ACF of r_t^2: 0.31 at lag 1 and still 0.053 at lag 50; ACF of $\hat{z}_t^2$ (GARCH(1,1)-t): -0.005 at lag 1, at most 0.028 in absolute value; band ± 0.024
- Interpretation: three parameters absorb almost all the volatility clustering in the daily data since 2000

SFM_ch9_diagnostics

Diagnostics on Real Data

S&P 500	$Q(10)$ (p)	$Q(10)$ of squares (p)	ARCH-LM(5) (p)	sign bias $\chi^2(3)$ (p)
returns r_t	96.3 (<0.001)	6133 (<0.001)	1679 (<0.001)	–
$\hat{z}_t$, Normal GARCH	21.2 (0.020)	16.4 (0.089)	8.0 (0.154)	–
$\hat{z}_t$, GARCH-t	21.9 (0.016)	13.6 (0.192)	5.8 (0.326)	40.7 (<0.001)

- GARCH removes the ARCH effects ($Q(10)$ of squares: from 6133 to 13.6); a little autocorrelation in the mean remains (constant-mean model)
- The sign-bias test rejects for the symmetric GARCH (sign $t = 3.56$): the asymmetry of Section 8 is missing; similar for the DAX ($p < 0.001$) and the BET ($p = 0.009$)
- Other markets, $Q(10)$ of $\hat{z}_t^2$ (GARCH-t): p between 0.115 (DAX) and 0.993 (TLV): no ARCH effect left anywhere

SFM_ch9_diagnostics

Nine Models for the S&P 500

- **AIC** (Akaike information criterion) $= -2\ell + 2k$ Akaike (1974); **BIC** (Bayesian information criterion) $= -2\ell + k \ln T$ Schwarz (1978); ℓ: maximised log-likelihood, k parameters, T observations; smaller is better (Chapter 6)

	Normal		Student-t		skewed t	
	log-lik.	ΔBIC	log-lik.	ΔBIC	log-lik.	ΔBIC
GARCH(1,1)	−9218.0	648.9	−9062.5	346.7	−9038.4	307.4
GJR-GARCH(1,1)	−9086.5	394.7	−8944.6	119.9	−8905.1	49.6
EGARCH(1,1)	−9075.6	372.9	−8922.1	74.8	−8880.3	0.0

- ΔBIC: the difference from the best model; AIC gives the same ranking
- Both choices matter: starting from the Normal GARCH, asymmetry alone (Normal EGARCH) lowers BIC by 276, heavy tails alone (GARCH-skewed t) by 341, both together by 649
- Interpretation: EGARCH with skewed-t innovations wins; with 6 714 observations even small improvements are "significant"; the forecast comparison in Section 11 is the real test

SFM_ch9_diagnostics

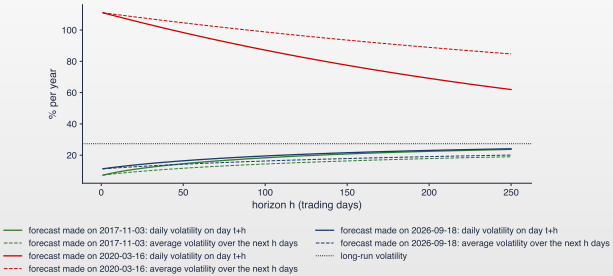
Multi-Step Forecasts

- One step: $\sigma_{t+1}^2 = \omega + \alpha\varepsilon_t^2 + \beta\sigma_t^2$ is known at the end of day t
  - two steps: $E_t[\sigma_{t+2}^2] = \omega + \alpha E_t[\varepsilon_{t+1}^2] + \beta\sigma_{t+1}^2 = \omega + (\alpha+\beta)\sigma_{t+1}^2$, since $E_t[\varepsilon_{t+1}^2] = \sigma_{t+1}^2$
- By induction: $E_t[\sigma_{t+h}^2] = \bar{\sigma}^2 + (\alpha+\beta)^{h-1}(\sigma_{t+1}^2 - \bar{\sigma}^2)$
  - E_t: the expectation given the information at the end of day t; h: the horizon in days
  - the forecast converges to the long-run variance; the speed is set by $\alpha+\beta$
- Variance of the H-day return: the sum of the daily forecasts (returns are uncorrelated)
  - $\mathrm{Var}_t(r_{t+1} + \cdots + r_{t+H}) = \sum_{h=1}^{H} E_t[\sigma_{t+h}^2]$, not $H\sigma_{t+1}^2$
  - the square-root-of-time rule $\sigma_{t+1}\sqrt{H}$ is too high in a crisis and too low in a calm period

Worked Example: the S&P 500 on 18 September 2026

- GARCH(1,1)-t: $\alpha + \beta = 0.9947$, $\bar{\sigma}^2 = 2.972$; forecast for the next day $\sigma^2_{t+1} = 0.509$ (below the long-run level)
- Step by step
  - day 10: $2.972 + 0.9947^9(0.509 - 2.972) = 2.972 + 0.953 \times (0.509 - 2.972) = 0.624$
  - 10-day variance: the sum of the ten daily forecasts $= 5.67$, so the 10-day volatility is 2.38%
  - square-root-of-time rule: $\sqrt{10 \times 0.509} = 2.26\%$: too low, because volatility is expected to rise towards its long-run level

The Term Structure of Volatility



- S&P 500, GARCH(1,1)-t estimated on all data; forecasts made on a calm day (3 November 2017), at the COVID-19 peak (16 March 2020) and on 18 September 2026
- Calm day: 7.3% tomorrow, average 19.0% over a year; COVID-19 peak: 111.0% tomorrow, still 84.7% on average over a year
- Interpretation: all curves move towards the long-run level (27.3%) at the same slow speed: the slope of the term structure tells whether we are in a calm or a storm

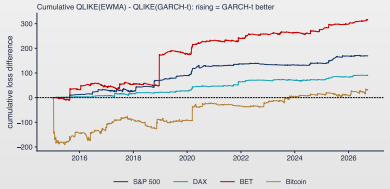
SFM_ch9_forecasts

Evaluating a Volatility Forecast

- The true σ_t^2 is never observed: we compare the forecast h_t with a noisy **proxy**, here r_t^2
  - $E_{t-1}[r_t^2] = \sigma_t^2$ (with $\mu \approx 0$), but r_t^2 is very noisy: low R^2 even for a perfect forecast Andersen & Bollerslev (1998)
- **QLIKE** loss Patton (2011): $L(r_t^2, h_t) = \dfrac{r_t^2}{h_t} + \ln h_t$ (up to a constant); smaller is better
  - h_t: the variance forecast for day t, made at the end of day $t-1$
  - ranks forecasts correctly even with a noisy proxy; it is minus the Normal log-likelihood
  - penalises forecasts that are too low more than forecasts that are too high
- **Out of sample**: from 2015, each forecast uses only data up to the day before; the models are re-estimated every 250 days
  - the **DM** (Diebold–Mariano) test Diebold & Mariano (1995): is the mean loss difference zero? t with HAC (heteroskedasticity and autocorrelation consistent) standard errors

GARCH against EWMA, Out of Sample

2015–2026	days	QLIKE GARCH-t	QLIKE GJR-t	QLIKE EWMA	DM t, GARCH-t vs EWMA (p)
S&P 500	2 944	0.761	0.733	0.819	-3.69 (<0.001)
DAX	2 973	1.090	1.064	1.121	-3.19 (0.001)
BET	2 926	0.680	0.674	0.788	-2.22 (0.027)
Bitcoin	4 279	3.383	3.401	3.390	-0.20 (0.845)



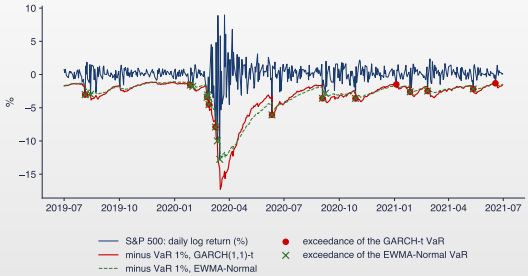
- Equity indices: GARCH-t beats EWMA ($t < -2$); GJR-t is better still for the S&P 500, DAX and BET
- Interpretation: Bitcoin: no difference ($t = -0.20$): its GARCH is an IGARCH, which is almost an EWMA (Section 7)

SFM_ch9_forecasts

Conditional VaR

- **VaR** at level α (value at risk): the loss exceeded with probability α: $\mathrm{VaR}_\alpha = -q_\alpha(r_{t+1})$ (Chapter 6)
  - here α is the VaR level, not the GARCH parameter; q_α: the α-quantile
  - here $\alpha = 1\%$: VaR 1%, a loss exceeded on one day in a hundred
- With $r_{t+1} = \mu + \sigma_{t+1} z_{t+1}$: $\mathrm{VaR}_{t+1} = -(\mu + \sigma_{t+1}\, q_\alpha(z))$
  - Normal: $q_{0.01} = -2.326$; standardised t: $q_{0.01} = t_\nu^{-1}(0.01)\sqrt{(\nu-2)/\nu}$
  - $t_\nu^{-1}(0.01)$: the 1% quantile of the Student-t with ν degrees of freedom; the square root rescales it to variance 1
  - the VaR moves every day with σ_{t+1}: high in storms, low in calm periods
- **Worked example**: S&P 500 on 18 September 2026, GARCH(1,1)-t: $\hat{\mu} = 0.079$, $\sigma_{t+1} = 0.714$, $\hat{\nu} = 6.26$
  - $q_{0.01}(z) = -3.100 \times 0.825 = -2.557$; VaR 1% $= -(0.079 + 0.714 \times (-2.557)) = 1.75\%$
  - with Normal z_t: 1.58%; over 10 days (sum of variances, μ ignored): $2.557 \times \sqrt{5.67}$, about 6.09%, an approximation: the 10-day return is not Student-t with the same ν, and its quantile needs simulation

VaR 1% through the COVID-19 Crash



- S&P 500, July 2019 – June 2021 (505 days): one-day VaR 1% from GARCH(1,1)-t (re-estimated every 250 days) and from EWMA with Normal z_t
- Exceedances (return below −VaR): GARCH-t 13, EWMA-Normal 17; expected about 5; both models react, but only after the first large falls of March 2020

SFM_ch9_forecasts

Exceedances 2015–2026: a Preview of Backtesting

VaR 1%	days	expected	GARCH-t	EWMA-Normal
S&P 500	2 944	29	49 (1.66%)	65 (2.21%)
DAX	2 973	30	45 (1.51%)	74 (2.49%)
BET	2 926	29	34 (1.16%)	67 (2.29%)
Bitcoin	4 279	43	64 (1.50%)	93 (2.17%)

- A correct VaR 1% is exceeded on 1% of the days, and the exceedances do not cluster
- EWMA-Normal: 2.2–2.5%: the Normal quantile is too small for heavy tails
- GARCH-t: 1.2–1.7%: much closer to 1%, but still above it in all four markets, most for the S&P 500, the DAX (the leverage effect is missing) and Bitcoin
- Are these differences significant? Chapter 10: the test of Kupiec (1995), ES 2.5%, backtesting
- Further reading on VaR forecasting for Bitcoin: Pele & Mazurencu-Marinescu-Pele (2019)

SFM_ch9_forecasts

An Open Research Question

- **Is Bitcoin's volatility becoming more like the volatility of an equity index?**
  - today: IGARCH, $\hat{\nu} = 3.18$, no leverage effect; the S&P 500: mean-reverting, strong leverage
  - since 2024, spot Bitcoin ETFs (exchange-traded funds) bring equity investors to the crypto market
- Earlier evidence on crypto assets as an asset class: Pele et al. (2023); volatility as information: Pele, Lazar & Dufour (2017)
- Why it is open: regime changes, the 2022 crypto crash and the ETF launch overlap; the answer depends on the window
- AI tools can speed up such a study, but its results must still be checked Wang et al. (2023)

Possible Uses of AI Tools

- **Literature**: a list of studies on GARCH models for crypto assets and the models they use
- **Code**: a first draft of rolling GJR-GARCH-t estimates for Bitcoin and the S&P 500 with the `arch` package
- **Robustness**: other windows, Ethereum, weekend days removed, skewed-t innovations
- Example prompt
  - `Write Python code that estimates a GJR-GARCH(1,1) model with Student-t innovations on daily Bitcoin and S&P 500 log returns in percent, on rolling windows of 750 days moved by 21 days, and plots the persistence, gamma and nu over time.`

What to Check

- Units: returns in % (not decimals), otherwise the optimiser may stop at a wrong ω
- The persistence of GJR is $\alpha + \beta + \gamma/2$, not $\alpha + \beta$; for EGARCH it is β
- Estimates on the boundary ($\alpha + \beta = 1$): no half-life, no long-run volatility; do not report infinite numbers
- Calendars: Bitcoin trades 7 days a week, the S&P 500 five; annualise each series with its own frequency
- Rolling windows overlap: consecutive estimates are not independent evidence
- References: every cited paper must exist; check the DOI

Project Idea

- **Question**: has Bitcoin's volatility moved closer to equity volatility since 2021?
  - data: Bitcoin since 2014, Ethereum since 2015, S&P 500 and BET (EODHD)
- Steps
  - GJR-GARCH-t on 2015–2019 and on 2021–2026: persistence, γ, ν
  - correlation of the Bitcoin and S&P 500 GARCH volatilities in each period
  - out-of-sample QLIKE of GARCH against EWMA in each period
- Deliverable: one table, one chart, and a paragraph on what the data can and cannot show
- Declare any AI use, and list the errors of the AI that you corrected

Key Takeaways

- Returns: unpredictable mean, predictable variance: $r_t = \mu + \sigma_t z_t$
- GARCH(1,1) with three parameters captures volatility clustering in all our series; $\alpha + \beta$ is close to 1
- Estimate by maximum likelihood, report robust SE; Student-t or skewed-t innovations for the tails
- The S&P 500 and the DAX need asymmetry (GJR, EGARCH); for the BET it is weak (BIC prefers the symmetric GARCH); Bitcoin does not need it, and is an IGARCH
- Forecasts revert to the long-run level; GARCH beats EWMA out of sample where volatility mean-reverts
- VaR 1% $= -(\mu + \sigma_{t+1} q_{0.01}(z))$: the volatility model and the quantile both matter

Key Formulas

Quantity	Formula
ARCH(q)	$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \quad \bar{\sigma}^2 = \omega/(1 - \sum \alpha_i)$
GARCH(1,1)	$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad \bar{\sigma}^2 = \omega/(1 - \alpha - \beta)$
Half-life	$h_{1/2} = \ln 0.5 / \ln(\alpha + \beta)$
GJR-GARCH	$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1}) \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$
EGARCH	$\ln \sigma_t^2 = \omega + \alpha(\lvert z_{t-1} \rvert - E\lvert z_{t-1} \rvert) + \gamma z_{t-1} + \beta \ln \sigma_{t-1}^2$
Log-likelihood (Normal)	$\ell = -\frac{1}{2} \sum_t [\ln 2\pi + \ln \sigma_t^2 + (r_t - \mu)^2 / \sigma_t^2]$
Forecast	$E_t[\sigma_{t+h}^2] = \bar{\sigma}^2 + (\alpha + \beta)^{h-1} (\sigma_{t+1}^2 - \bar{\sigma}^2)$
QLIKE	$L = r_t^2 / h_t + \ln h_t$
VaR 1%	$-(\mu + \sigma_{t+1}\, q_{0.01}(z))$

Self-Assessment

- **Question**: $\omega = 0.05$, $\alpha = 0.05$, $\beta = 0.90$. What are the long-run variance and the half-life?
  - **Answer**: $0.05/0.05 = 1$; $\ln 0.5 / \ln 0.95 = 13.5$ days
- **Question**: GJR gives $\hat{\alpha} = 0$, $\hat{\gamma} = 0.2$. What happens after a rise of 3%?
  - **Answer**: the shock adds nothing: $\sigma^2_{t+1} = \omega + \beta\sigma^2_t$, as after a day with $\varepsilon_t = 0$ (if $\varepsilon_t = r_t - \mu > 0$); only negative shocks raise the variance
- **Question**: $Q(10)$ of $\hat{z}_t^2$ has p = 0.40. Is the model correct?
  - **Answer**: we only know that no ARCH effect is left; asymmetry or the wrong distribution may remain
- Next: Chapter 10, VaR, ES and backtesting

References (1/2)

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Andersen, T.G., Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885–905.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics*, 69(3), 542–547.
- Bollerslev, T., Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2), 143–172.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Christie, A.A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics*, 10(4), 407–432.
- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Engle, R.F. (2004). Risk and volatility: econometric models and financial practice. *American Economic Review*, 94(3), 405–420.
- Engle, R.F., Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5(1), 1–50.
- Engle, R.F., Ng, V.K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance*, 48(5), 1749–1778.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.

References (2/2)

- **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. ***Journal of Finance***, **48(5), 1779–1801.**
- **Hansen, B.E. (1994).** Autoregressive conditional density estimation. ***International Economic Review***, **35(3), 705–730.**
- **Hansen, P.R., Lunde, A. (2005).** A forecast comparison of volatility models: does anything beat a GARCH(1,1)? ***Journal of Applied Econometrics***, **20(7), 873–889.**
- **Kupiec, P.H. (1995).** Techniques for verifying the accuracy of risk measurement models. ***Journal of Derivatives***, **3(2), 73–84.**
- **Ljung, G.M., Box, G.E.P. (1978).** On a measure of lack of fit in time series models. ***Biometrika***, **65(2), 297–303.**
- **Nelson, D.B. (1991).** Conditional heteroskedasticity in asset returns: a new approach. ***Econometrica***, **59(2), 347–370.**
- **Patton, A.J. (2011).** Volatility forecast comparison using imperfect volatility proxies. ***Journal of Econometrics***, **160(1), 246–256.**
- **Pele, D.T., Lazar, E., Dufour, A. (2017).** Information entropy and measures of market risk. ***Entropy***, **19(5), 226.**
- **Pele, D.T., Mazurencu-Marinescu-Pele, M. (2019).** Using high-frequency entropy to forecast Bitcoin's daily value at risk. ***Entropy***, **21(2), 102.**
- **Pele, D.T., Wesselhöfft, N., Härdle, W.K., Kolossiatis, M., Yatracos, Y.G. (2023).** Are cryptos becoming alternative assets? ***European Journal of Finance***, **29(10), 1064–1105.**
- **Schwarz, G. (1978).** Estimating the dimension of a model. ***Annals of Statistics***, **6(2), 461–464.**
- **Tsay, R.S. (2010).** ***Analysis of Financial Time Series***, **3rd ed. Wiley.**
- **Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023).** Scientific discovery in the age of artificial intelligence. ***Nature***, **620, 47–60.**