

Modelling Financial Markets

Chapter 7: Value-at-Risk and Expected Shortfall



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

How Large Are Daily Tail Losses?



Freddy Delbaen. **Photo: ETH-Bibliothek Zürich (1997); CC BY-SA 4.0; Wikimedia Commons**

"We present and justify a set of four desirable properties for measures of risk, and call the measures satisfying these properties 'coherent.'"

— *Artzner, Delbaen, Eber & Heath (1999)*

- S&P 500 (Standard & Poor's), 2000–2026: historical VaR 1% (Value-at-Risk, the loss exceeded on 1% of days) is 3.45%
  - ES 1% (Expected Shortfall, the mean loss on the worst 1% of days) is 5.03%; the largest loss is 12.77% (16 March 2020)
- Two independent bonds, each losing 100 with probability 0.9%: VaR 1% is 0 for each bond and 100 for the pair
  - ES 1% is 90 for each bond and 100.81 for the pair, below the sum of 180
- Over the last 500 days, ES 2.5% rests on only $n\alpha = 12.5$ losses
  - its HAC (heteroskedasticity and autocorrelation consistent) interval is 2.55 percentage points wide
- Days with a loss above the one-day VaR 1%, 2006–2026 (target 1%)
  - 1.67% for 500-day historical simulation, 1.54% for filtered historical simulation (FHS), 2.34% for GARCH (Generalised AutoRegressive Conditional Heteroskedasticity) with Normal shocks

▸ Central chart: Historical vs Filtered Historical Simulation: S&P 500 (1/2)

How Large Are Daily Tail Losses, and How Precisely Can We Estimate Them?

What we will explain

- Two tail measures: VaR and ES
  - VaR (Value-at-Risk): a loss threshold
  - ES (Expected Shortfall): the mean loss beyond it
- Part I: definitions, coherence, unconditional estimation methods and their precision
- Part II: conditional forecasts, extreme value theory, horizon and portfolio
  - conditional forecasts: filtered historical simulation (FHS) and direct quantile models (CAViaR)
  - regulation, two case studies and one pass of the research loop assisted by AI (artificial intelligence)
- Solved exercises, Seminar 7: A1, A2, A3, A5, B1, B4

What you should be able to do

- derive Normal and Student-t VaR and ES
- interpret why ES is coherent and VaR is not
- compute both, each with a standard error or an interval
  - historically, parametrically, by filtered simulation and by extreme value theory
- compute 10-day risk and Euler contributions
- interpret the Basel rules, including FRTB
  - FRTB (Fundamental Review of the Trading Book): the 2019 market-risk standard
- Further Topics and Appendix: reading for familiarity

Which Losses and Units Will We Use?

- Running example: S&P 500 daily losses, 2000–2026
  - comparisons on BET (main index of the Bucharest Stock Exchange, BVB), Bitcoin, EUR (euro)/RON and gold
  - EUR/RON: the reference rate of the National Bank of Romania (BNR), in Romanian lei (RON) per euro (EUR)
  - data up to 18 September 2026 (full inventory in the Appendix)
- Units: daily log returns in %, $r_t = 100\,(\ln P_t - \ln P_{t-1})$, with P_t the closing price on day t
- Why 2000–2026: the sample spans three crises (2000–2002, 2008, 2020)
  - so the tail holds enough extreme losses

How Much Can We Lose? RiskMetrics and VaR



J.P. Morgan & Co. building, 23 Wall Street, New York

Photo: Beyond My Ken (2012); CC BY-SA 4.0; Wikimedia Commons

- Management and regulators need **one number**: a loss exceeded only rarely
  - to set limits and **capital**
- J.P. Morgan published **RiskMetrics** in October 1994
  - a free methodology and data set for market risk
  - its Technical Document (4th ed., 1996) became the reference for VaR J.P. Morgan/Reuters (1996)
- Textbooks: Franke, Härdle & Hafner (2019), McNeil, Frey & Embrechts (2015), Daníelsson (2011)

Losses and the Sign Convention

- Return over one day: $X_{t+1} = r_{t+1}$, the log return in %
- **Loss**: $L_{t+1} = -X_{t+1}$, a **log loss**
  - $L > 0$ means the position lost money
- In money, with V_t the value of the position: loss $= V_t\,(1 - e^{-L_{t+1}/100}) \approx V_t L_{t+1}/100$ for small L
- Why log: log returns add over days, simple returns add across assets
  - over days: the horizon h; across assets: the portfolio
- The map $L \mapsto 1 - e^{-L/100}$ is increasing
  - so VaR in money is $V_t(1 - e^{-\mathrm{VaR}_\alpha/100})$
  - ES in money is computed scenario by scenario, then averaged
- **Sign convention**: VaR and ES are **positive** numbers when they are losses

Horizon and Level: Why 1 Day, 10 Days, 1% and 2.5%

- **Horizon** h: the number of days over which the loss is measured
- $h = 1$ day: the horizon of trading limits and of backtesting (comparing forecasts with realised losses), with a new observation every day
- $h = 10$ days: the time Basel assumes for liquidating a position Basel Committee (1996)
- **Level** α: the **tail probability**
  - we write "VaR 1%", "ES 2.5%"
- $\alpha = 1\%$: the VaR level of the 1996 Basel amendment Basel Committee (1996), one exceedance every 100 days on average
- $\alpha = 2.5\%$: the ES level of FRTB Basel Committee (2019)
  - chosen because under the Normal distribution ES 2.5% is close to VaR 1%

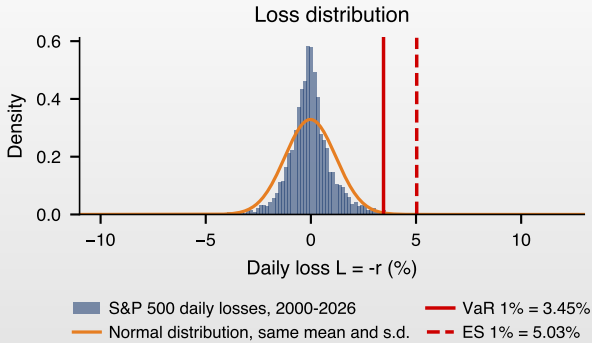
Definitions (1/2): Value-at-Risk

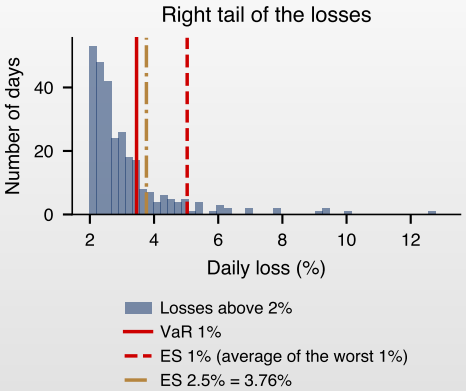
- The question of this section: which loss is exceeded on only a fraction α of days?
  - and how much do we lose, on average, beyond it?
- **VaR** at level α: the loss exceeded with probability α (exactly for continuous L, at most α in general)
  - $\mathrm{VaR}_\alpha(X) = -\inf\{x \in \mathbb{R} : P(X \le x) > \alpha\} = -q_\alpha(X)$
  - X: the return of the position over the horizon; $q_\alpha(X)$: its α-quantile, negative for small α, so VaR is a positive loss
  - For a continuous loss $L = -X$: $\mathrm{VaR}_\alpha = q_{1-\alpha}(L)$, the upper quantile of the loss
- Reading of VaR 1%
  - "With probability 1%, tomorrow's loss will exceed $\mathrm{VaR}_{1\%}$"
  - VaR does **not** say how large the loss is beyond this threshold

Definitions (2/2): Expected Shortfall

- **Expected Shortfall**: the average of all quantiles in the tail of probability α
  - $\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q_u(X)\,du$
  - equivalently, $\mathrm{ES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(X)\,du$
  - u: the integration variable, a tail probability between 0 and α
- For a continuous distribution: the mean loss on the days beyond VaR
  - $\mathrm{ES}_\alpha(X) = -\mathbb{E}[X \mid X \le q_\alpha(X)]$
  - Always $\mathrm{ES}_\alpha \ge \mathrm{VaR}_\alpha$: ES looks **into** the tail, VaR only at its entrance; two portfolios with the same VaR can have very different ES
- Other names: Conditional VaR (CVaR), Tail VaR, Average VaR Rockafellar & Uryasev (2002), Acerbi & Tasche (2002)

VaR and ES on the S&P 500 (1/2)





MFM_ch7_loss_distribution ▶ Interactive

VaR and ES on the S&P 500 (2/2)

- Daily losses 2000–2026 against the Normal distribution with the same mean and standard deviation
  - left: the whole distribution; right: the tail
- Historical $\mathrm{VaR}_{1\%} = 3.45\%$: 68 of 6 717 days (1.01%) lost more
  - $\mathrm{ES}_{1\%} = 5.03\%$ is the average loss over the worst 1% of days
- At the 2.5% level: historical VaR is 2.51% and ES is 3.76%
  - largest loss 12.77% on 2020-03-16 (COVID-19 crash), about ten standard deviations
- The Normal distribution fits the centre but is far too thin in the tail (Chapter 1)
- Type of evidence: description (it shows what happened in the data, without testing a cause)

MFM_ch7_loss_distribution

Closed-Form ES under the Normal Distribution (1/2)

- Setting: $X \sim N(\mu, \sigma^2)$, a Normal return with mean μ and standard deviation σ
  - Φ, φ: the distribution function and the density of the standard Normal distribution $N(0,1)$
  - $z_\alpha = \Phi^{-1}(\alpha) < 0$: the standard Normal α-quantile, e.g. $z_{1\%} = -2.326$
- VaR: the mean minus $|z_\alpha|$ standard deviations, sign changed
  - $\mathrm{VaR}_\alpha = -(\mu + \sigma z_\alpha)$
- ES: the same form, with the larger factor $\varphi(z_\alpha)/\alpha$ in place of $-z_\alpha$
  - $\mathrm{ES}_\alpha = -\mu - \dfrac{\sigma}{\alpha} \displaystyle\int_{-\infty}^{z_\alpha} u\varphi(u)\, du$
  - The integral equals $-\varphi(z_\alpha)$ because $\varphi'(u) = -u\varphi(u)$
  - Hence $\mathrm{ES}_\alpha = -\mu + \sigma \dfrac{\varphi(z_\alpha)}{\alpha}$

Closed-Form ES under the Normal Distribution (2/2)

- Constants to remember
  - $-z_{1\%} = 2.326$, $\varphi(z_{1\%})/0.01 = 2.665$
  - $-z_{2.5\%} = 1.960$, $\varphi(z_{2.5\%})/0.025 = 2.338$
  - So, under the Normal distribution, ES 2.5% is almost equal to VaR 1% (2.338 against 2.326 standard deviations)
- S&P 500 ($\hat{\mu} = 0.025\%$, $\hat{\sigma} = 1.213\%$): Normal $\mathrm{VaR}_{1\%} = 2.80\%$ vs historical 3.45%
  - Formula: $\mathrm{VaR}_{1\%} = -\hat{\mu} + \hat{\sigma}\,(-z_{1\%})$
  - Arithmetic: $-0.025 + 1.213 \times 2.326 \approx 2.80\%$
  - The historical VaR is 1.24 times the Normal VaR
- Under the Normal distribution, for a given μ, VaR and ES carry no information beyond σ

The Student-t Distribution: VaR and ES

- Model: the return is a shifted and rescaled Student-t variable, $X = m + s\,T$ with $T \sim t_\nu$
  - m: location; $s > 0$: scale; ν: degrees of freedom, smaller ν = heavier tails
  - $t_\alpha = t_\nu^{-1}(\alpha) < 0$: the α-quantile of t_ν; g_ν: the density of t_ν
- VaR as for the Normal distribution, with t_α in place of z_α
  - ES gets an extra factor $(\nu + t_\alpha^2)/(\nu - 1) > 1$ for the heavy tail
  - $\mathrm{VaR}_\alpha = -(m + s\,t_\alpha), \qquad \mathrm{ES}_\alpha = -m + s\,\dfrac{g_\nu(t_\alpha)}{\alpha}\,\dfrac{\nu + t_\alpha^2}{\nu - 1}$ (for $\nu > 1$)
  - Variance $s^2\nu/(\nu - 2)$ exists only for $\nu > 2$, so s is not the standard deviation

Coherent Risk Measures (1/2): the Four Axioms

- The question: which properties must a number used as capital have?
  - so that diversification is never penalised
- A **risk measure** ρ maps a loss L (a random variable) to the capital to hold
  - it is **coherent** if it satisfies four axioms Artzner, Delbaen, Eber & Heath (1999)
  - L_1, L_2: two losses; c: a sure amount (constant); $\lambda \geq 0$: a scaling factor (position size)
- **Monotone**: $L_1 \leq L_2 \Rightarrow \rho(L_1) \leq \rho(L_2)$
  - A position that always loses more needs more capital
- **Translation invariant**: $\rho(L + c) = \rho(L) + c$
  - Adding a sure loss c raises the capital by exactly c
- **Positively homogeneous**: $\rho(\lambda L) = \lambda \rho(L)$
  - Doubling the position doubles the capital
- **Subadditive**: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$
  - Merging two positions never needs more capital than holding them apart: diversification does not add risk

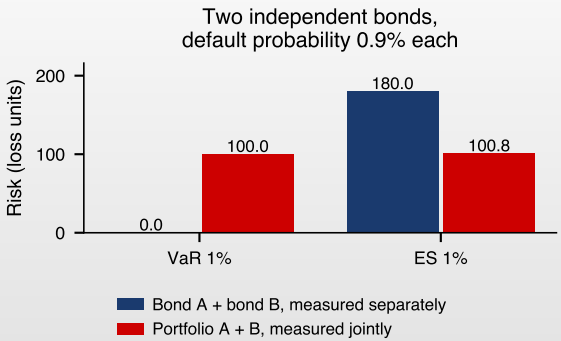
Coherent Risk Measures (2/2): Why the Axioms Matter

- **Limits**
  - By subadditivity, bank-wide risk does not exceed the sum of the desk risk limits, $\sum_i \rho_i$
  - **No regulatory arbitrage** by splitting a firm
- **Optimisation**: homogeneity + subadditivity $\Rightarrow$ convexity in the portfolio weights
  - Every local minimum is global; uniqueness would need **strict** convexity, which scenario-based ES lacks (it is piecewise linear)
- VaR satisfies the first three axioms, but **not** subadditivity in general
  - Exception: jointly elliptical losses (multivariate Normal or Student-t) and $\alpha < 50\%$, where VaR and ES are both $\mu_p + \sigma_p \times$ constant McNeil, Frey & Embrechts (2015)
  - μ_p, σ_p: mean and standard deviation of the portfolio loss (heavy-tailed refinements: Appendix)

Counterexample: Two Bonds, Step by Step

- Bonds A and B, independent: each loses 100 with probability $p = 0.9\%$ (default)
  - and 0 otherwise
  - X: the value change, $X = -L$
- General formula for a discrete loss $L = -X$: $\mathrm{ES}_\alpha = \frac{1}{\alpha}\left(\mathbb{E}[L\,1\{L > \mathrm{VaR}_\alpha\}] + \mathrm{VaR}_\alpha\left(\alpha - P(L > \mathrm{VaR}_\alpha)\right)\right)$
  - $1\{\cdot\}$: indicator, 1 if the condition holds and 0 otherwise; first term: losses strictly beyond VaR; second term: the rest of the tail probability α, placed at VaR itself
- **One bond**: $P(X \le -100) = p = 0.9\%$, not above 1%, so $q_{1\%} = 0$ and $\mathrm{VaR}_{1\%}(A) = \mathrm{VaR}_{1\%}(B) = 0$
  - $\mathrm{ES}_{1\%}(A) = (100 \cdot 0.009 + 0 \cdot 0.001)/0.01 = 90$
- **Portfolio A + B**: $P(X \le -100) = 1 - (1-p)^2 = 1.79\% > 1\%$
  - $P(X \le -200) = p^2 = 0.0081\%$, so $q_{1\%} = -100$
  - $\mathrm{VaR}_{1\%}(A+B) = 100 > 0 + 0 = \mathrm{VaR}_{1\%}(A) + \mathrm{VaR}_{1\%}(B)$
- $\mathrm{ES}_{1\%}(A+B) = 100.81 \le 180 = \mathrm{ES}_{1\%}(A) + \mathrm{ES}_{1\%}(B)$

Subadditivity of VaR and ES: a Counterexample



- Blue: sum of the two stand-alone measures
  - red: the measure of the combined portfolio
- VaR jumps from $0 + 0$ to 100 when the bonds are combined (the counterexample of the previous slide)
- A VaR limit would reward **concentration**: holding only A looks riskless
  - ES sees the size of the tail and is coherent Acerbi & Tasche (2002)
- Credit portfolios, options and other skewed pay-offs are where VaR fails in practice

MFM_ch7_subadditivity ▶ Interactive

Quick Check: Which Position Is Riskier?

- Two positions, losses in %, level $\alpha = 10\%$
- **A**: loss L_A uniform on $[0, 10]$, every value equally likely
  - Find $\mathrm{VaR}_{10\%}(A)$ and $\mathrm{ES}_{10\%}(A)$ from the definitions
- **B**: loss 0 with probability 0.95, loss 20 with probability 0.05
  - Find $\mathrm{VaR}_{10\%}(B)$ and $\mathrm{ES}_{10\%}(B)$ with the discrete ES formula of the two-bond slide
- Which position does each measure call riskier?
- What does VaR miss?
- Answers in the ▸ Appendix

Coherence versus Robustness

- **Qualitative robustness** Cont, Deguest & Scandolo (2010)
  - a small change in the data distribution changes the law of the estimator only a little
  - Historical VaR is qualitatively robust at distributions with a unique α-quantile; historical ES is **not**: one extreme loss moves it without bound (the precision section)
- S&P 500, last 500 days: drop the single largest loss (6.16%, 2025-04-04)
  - $\mathrm{ES}_{2.5\%}$: 2.98% $\to$ 2.64%; $\mathrm{VaR}_{1\%}$: 2.73% $\to$ 2.68%
- A trade-off, not a free lunch
  - subadditivity needs the tail beyond the quantile
  - this is exactly what makes ES sensitive to single observations
- FRTB chose coherence for capital (ES 2.5%) and kept VaR for backtesting
  - VaR is robust
  - VaR forecasts can be compared directly through a score (elicitability, next slide)

Elicitability: Why ES Is Harder to Evaluate

- A **scoring function** $S(v, X)$: a penalty computed from the forecast v and the outcome X
  - lower means a better forecast
- A risk measure is **elicitable** if it is the unique minimiser of $\mathbb{E}[S(v, X)]$ for some S
  - such an S is a strictly consistent score
  - VaR: the quantile $q_\alpha(X)$ minimises the pinball loss $\mathbb{E}[(1\{X \le v\} - \alpha)(v - X)]$
  - The pinball loss charges $1 - \alpha$ per unit when X falls below v and α per unit when it stays above
- ES alone is **not** elicitable Gneiting (2011)
  - no strictly consistent scoring function exists for ES on its own
- The pair (VaR, ES) is jointly elicitable Fissler & Ziegel (2016)
  - joint scores such as FZ0 rank (VaR, ES) forecasts (Appendix; Chapter 8)
  - FZ0: the Fissler–Ziegel score of Patton, Ziegel & Chen (2019)
  - $S(cv, ce; cY) = S(v, e; Y) + \ln c$: score differences, not the score itself, are homogeneous of degree zero

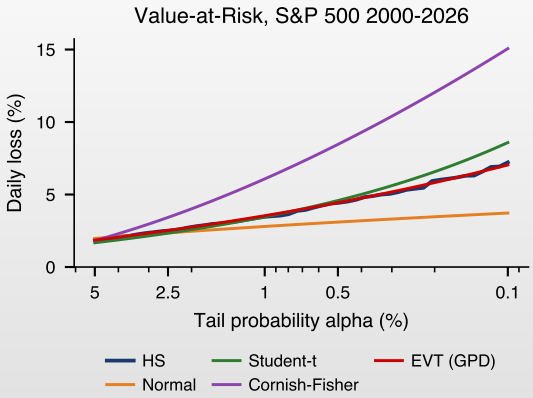
Historical Simulation

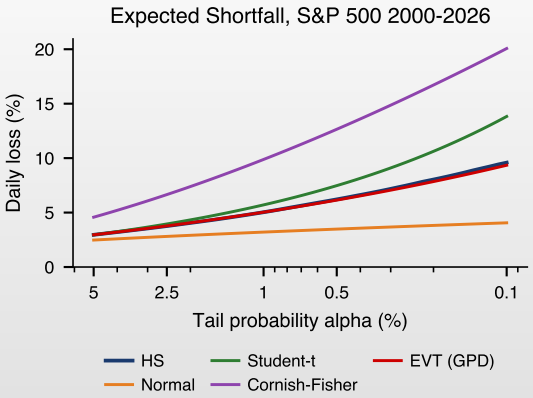
- The question of this section: how do we estimate VaR and ES from data?
  - and how much does the answer depend on the tail assumption?
- Take the last n daily losses $L_{t-n+1}, \ldots, L_t$ and sort them: $L_{(1)} \leq \cdots \leq L_{(n)}$
  - $L_{(i)}$: the i-th smallest loss
  - $\widehat{\mathrm{VaR}}_\alpha = L_{(j)} + (\ell - j)\,(L_{(j+1)} - L_{(j)})$, with $\ell = 1 + (n-1)(1-\alpha)$ and $j = \lfloor \ell \rfloor$: the empirical $(1-\alpha)$-quantile, interpolated linearly
  - $\widehat{\mathrm{ES}}_\alpha = \frac{1}{n\alpha}\left[\sum_{i=1}^{m} L_{(n-i+1)} + (n\alpha - m)\, L_{(n-m)}\right]$, with $m = \lfloor n\alpha \rfloor$: the integral of the empirical quantile function over the worst fraction α of the losses
  - The m largest losses get weight 1; the next one, $L_{(n-m)}$, gets the fractional weight $n\alpha - m$
  - Alternative, the mean of the selected exceedances: the simple mean of the k_α losses $\geq \widehat{\mathrm{VaR}}_\alpha$; it differs from $\widehat{\mathrm{ES}}_\alpha$ when $n\alpha$ is not an integer
- Choice of window n: 250 days reacts faster, 1000 days is more precise
  - With $n = 500$: 5 losses lie beyond VaR 1%; ES 2.5% gives weight 1 to the 12 largest losses and weight 0.5 to the 13th ($n\alpha = 12.5$)
- No loss beyond the worst one in the window can appear in the estimate
- **Ghost effect**: a crash day enters the window, stays for n days, then drops out and VaR falls abruptly

Cornish–Fisher: Correcting the Normal Quantile, and Where It Breaks

- Adjust the Normal quantile for **skewness** $S = \mathbb{E}[(X-\mu)^3]/\sigma^3$ and **excess kurtosis** $K = \mathbb{E}[(X-\mu)^4]/\sigma^4 - 3$ Cornish & Fisher (1938)
  - $\tilde{z}_\alpha = z_\alpha + \frac{(z_\alpha^2-1)S}{6} + \frac{(z_\alpha^3-3z_\alpha)K}{24} - \frac{(2z_\alpha^3-5z_\alpha)S^2}{36}, \quad \mathrm{VaR}_\alpha = -(\mu+\sigma\tilde{z}_\alpha)$
  - $\tilde{z}_\alpha$: the corrected quantile; $S = K = 0$ gives back the Normal VaR
  - For α below about 4% ($|z_\alpha| > \sqrt{3}$): $S < 0$ (long left tail of returns) or $K > 0$ (heavy tails) push $\tilde{z}_\alpha$ further below z_α and raise VaR; the S^2 term pushes the other way. At $\alpha = 5\%$ the K term has the opposite sign
- ES: average of $-(\mu+\sigma\tilde{z}_u)$ over $u \in (0,\alpha)$, numerically
- An expansion around the Normal distribution: accurate only for **small** S and K
  - for large K the map $z \mapsto \tilde{z}$ is not even monotone (Seminar 7, A6)
- S&P 500 returns: $S = -0.35$, $K = 10.7$: far outside that range
  - the sample K requires a finite fourth moment, which Chapter 1 questions
  - $\mathrm{VaR}_{1\%}^{CF} = 6.08\%$ versus historical 3.45%: the kurtosis term overshoots

Five Methods Across Tail Probabilities (1/2)





MFM_ch7_var_levels ▶ Interactive

Five Methods Across Tail Probabilities (2/2)

- Horizontal axis: α from 5% down to 0.1%, log scale
  - vertical axis: VaR (left) and ES (right), S&P 500 daily losses 2000–2026, in %
  - legend: HS = historical simulation
  - EVT (extreme value theory) = a GPD (generalised Pareto distribution) tail, a preview of the Extreme Value Theory section
- VaR 0.1% (one day in four years)
  - historical 7.22%, EVT 7.04%, Student-t 8.58%, Cornish–Fisher 15.06%
  - Normal 3.72%, about half
- The gap between methods **grows** as α falls, more for ES than for VaR
  - EVT and historical agree down to 0.5%, below that EVT extrapolates a fitted tail
- Type of evidence: comparison of procedures (it shows which method works better on the same data, not why)

MFM_ch7_var_levels

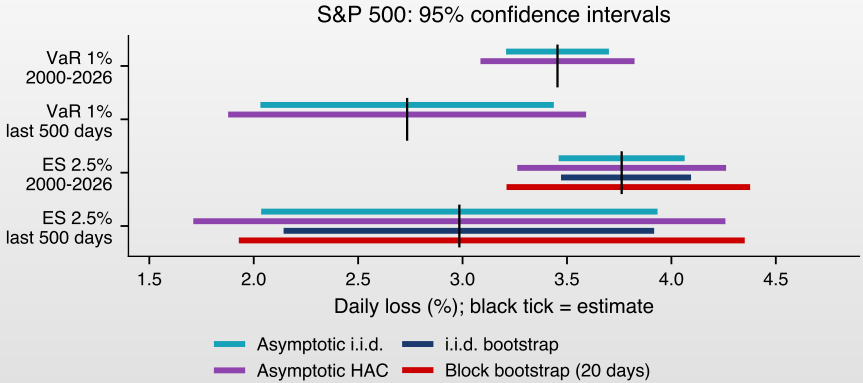
Two Usable Standard Errors (1/2): VaR

- The question of this section: how much would the estimated VaR and ES change on another sample of the same length?
- **VaR**: i.i.d. (independent and identically distributed) losses with density f continuous and positive at VaR_α
  - $\mathrm{SE}(\widehat{\mathrm{VaR}}_\alpha) \approx \dfrac{1}{f(\mathrm{VaR}_\alpha)} \sqrt{\dfrac{\alpha(1-\alpha)}{n}}$ (SE: standard error)
  - $\sqrt{\alpha(1-\alpha)/n}$: the SE of the share of days beyond VaR; dividing by the density $f(\mathrm{VaR}_\alpha)$ turns it into an error on the loss scale
  - The tail density is small, so the SE is large
  - $\sigma = 1.2\%$, $n = 500$: SE of $\mathrm{VaR}_{1\%} = 0.20$ percentage points (pp) under the Normal distribution, 0.43 pp under a t_4 (Student-t, 4 degrees of freedom) with the same σ

Two Usable Standard Errors (2/2): ES

- **ES** Chen (2008): influence function $\mathrm{IF}(x) = \mathrm{VaR}_\alpha + (x - \mathrm{VaR}_\alpha)_+/\alpha - \mathrm{ES}_\alpha$, unbounded in x
  - $\mathrm{IF}(x)$: the effect of one extra observation x on the estimate; it grows without bound with x, so one extreme loss can move ES arbitrarily far
- The standard error
  - If $\mathbb{E}[L^2] < \infty$: $\mathrm{SE}(\widehat{\mathrm{ES}}_\alpha) \approx \frac{1}{\alpha}\sqrt{\frac{\mathrm{Var}((L - \mathrm{VaR}_\alpha)_+)}{n}}$; Var: the variance; estimating VaR adds no first-order term (Rockafellar–Uryasev)
  - For Pareto-type tails with tail index $a > 2$ the **relative** error falls like $1/\sqrt{n\alpha}$: the tail count governs precision Yamai & Yoshiba (2005), Pele & Mazurencu-Marinescu-Pele (2026)
- Under volatility clustering
  - Replace the variances by long-run HAC (heteroskedasticity and autocorrelation consistent) variances Newey & West (1987)of $1\{L_t > \mathrm{VaR}_\alpha\}$ and $(L_t - \mathrm{VaR}_\alpha)_+$
  - Or use a block bootstrap, which resamples blocks of consecutive days (derivations: ▸ Appendix)

Precision on the S&P 500 (1/2)



MFM_ch7_precision ▶ Interactive

Precision on the S&P 500 (2/2)

- 95% intervals (bars), with estimates marked by black ticks
  - bootstrap computation: Seminar 7, B2
  - complete table in the Appendix
- Full sample: clustering widens the ES interval from 0.60 (i.i.d.) to 1.00 (HAC)
  - the bootstrap does the same: 0.62 i.i.d. vs 1.17 by blocks
  - last 500 days: ES 2.5% rests on only $n\alpha = 12.5$ losses, HAC interval width 2.55 pp
- These intervals cover the **unconditional** VaR and ES
  - tomorrow's risk is a conditional question (next section)
- The HAC interval for ES requires a finite second moment of the excess beyond VaR, $\mathbb{E}[(L - v)_+^2] < \infty$ (tail index above 2, estimated in Chapter 1), plus stationarity and weak dependence
  - GARCH-QML (GARCH Quasi-Maximum Likelihood) inference has its own conditions: a finite fourth moment of the standardised innovations z_t
- Type of evidence: description

MFM_ch7_precision

Quick Check: How Many Days Make a Precise VaR?

- i.i.d. Normal daily losses with mean 0 and $\sigma = 1\%$; level $\alpha = 5\%$
  - so $\mathrm{VaR}_{5\%} = 1.645\%$
  - the density there is $f(\mathrm{VaR}_{5\%}) = \varphi(1.645)/\sigma = 0.1031$
  - (a) With $n = 1\,000$ days, what is the SE of historical $\mathrm{VaR}_{5\%}$ and the half-width of its 95% interval?
  - (b) How many days halve that SE?
  - (c) With $n = 1\,000$, how many losses does historical $\mathrm{ES}_{1\%}$ average, and why is that a warning?
- Answers in the ▸ Appendix

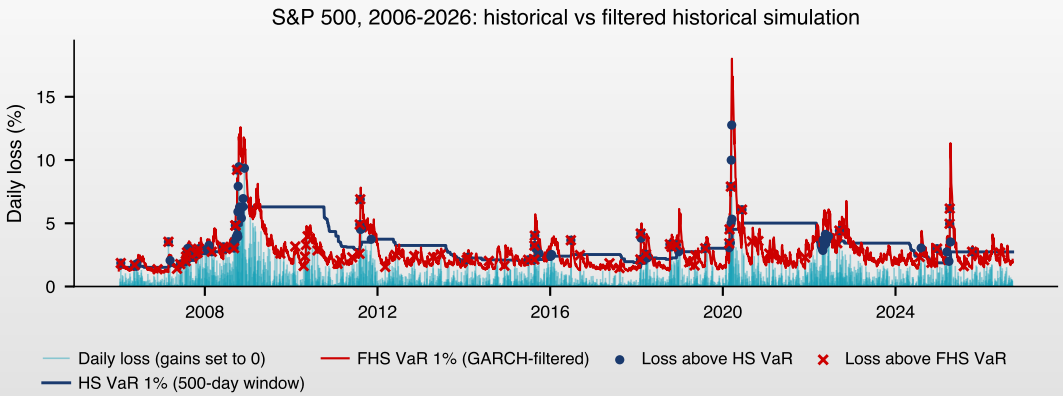
Filtered Historical Simulation (1/2): the Model

- The question of this section: how do we adapt VaR and ES to today's volatility?
  - rather than to the average of recent years
- Idea Barone-Adesi, Giannopoulos & Vosper (1999), Hull & White (1998)
  - history gives the **shape** of the shocks, a GARCH model gives today's **scale**
  - GARCH: generalised autoregressive conditional heteroskedasticity
  - $r_t = \mu_t + \sigma_t z_t$, $\mu_t = c + \phi r_{t-1}$, $\sigma_t^2 = \omega + \alpha_1 (r_{t-1} - \mu_{t-1})^2 + \beta_1 \sigma_{t-1}^2$ (Chapter 5, Bollerslev (1986); α_1 is a GARCH coefficient, not the level α)
- Notation
  - μ_t: the conditional mean, an AR(1) (autoregressive of order 1) in the previous return with constant c and coefficient ϕ
  - σ_t: the conditional volatility; z_t: i.i.d. shock with mean 0 and variance 1, distribution left unspecified
  - $\omega > 0$: constant; $\alpha_1 \geq 0$: reaction to yesterday's squared surprise; $\beta_1 \geq 0$: weight on yesterday's variance; $\alpha_1 + \beta_1$: persistence; $\alpha_1 + \beta_1 < 1$ ensures a finite unconditional variance (close to 1: shocks fade slowly)

Filtered Historical Simulation (2/2): Three Steps

- Step 1: estimate the AR(1)-GARCH(1,1) model by quasi-maximum likelihood
  - Normal likelihood, robust to other shock distributions
- Step 2: standardised residuals $\hat{z}_t = (r_t - \hat{\mu}_t)/\hat{\sigma}_t$: approximately i.i.d., still heavy-tailed
- Step 3: rescale by tomorrow's volatility
  - $\mathrm{VaR}_{\alpha,t+1} = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\,\hat{\xi}_\alpha$, $\hat{\xi}_\alpha = -q_\alpha(\hat{z})$; $\quad \mathrm{ES}_{\alpha,t+1} = -\hat{\mu}_{t+1} - \hat{\sigma}_{t+1}\,\bar{z}_\alpha$
  - $\hat{\mu}_{t+1}$, $\hat{\sigma}_{t+1}$: the model forecasts for tomorrow; $\hat{\xi}_\alpha$: minus the empirical α-quantile of the residuals (not the EVT shape ξ)
  - $\bar{z}_\alpha$: the average of the lowest fraction α of the residuals $\hat{z}_t$, with the fractional weight of historical simulation

Historical vs Filtered Historical Simulation: S&P 500 (1/2)



MFM_ch7_hs_fhs_sp500 ▶ Interactive

Historical vs Filtered Historical Simulation: S&P 500 (2/2)

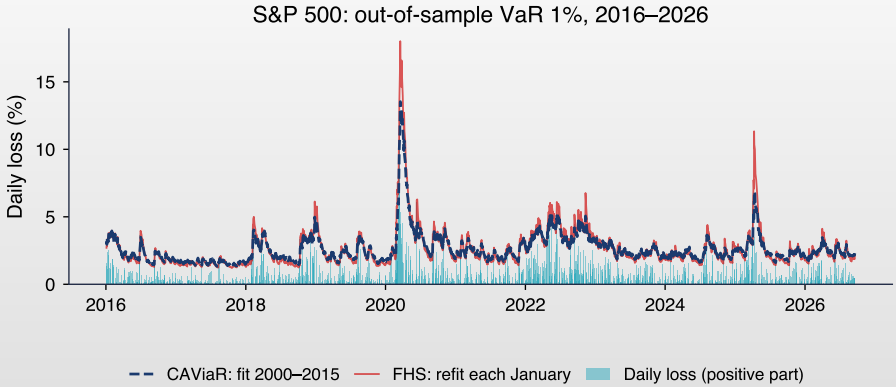
- Lines: one-day-ahead $\mathrm{VaR}_{\mathbf{1\%}}$; bars: the daily loss
  - dark blue for 500-day historical simulation, red for FHS
- Blue dots: losses above historical VaR; red crosses: losses above FHS VaR
  - GARCH re-estimated each January on all earlier data
  - so each forecast uses only information available at its origin
- Days with a loss above VaR 1%, 2006–2026 (5 210 days; target 1%)
  - historical 500 days 1.67%, filtered 1.54%
  - GARCH with Normal shocks 2.34%
- Historical exceedances **cluster** (2008, 2020): the window reacts months late
  - filtered VaR peaked at 18.01% on 2020-03-18, two days after the largest loss
  - FHS spreads exceedances more evenly over time (tests in Chapter 8)
- Type of evidence: comparison of procedures

MFM_ch7_hs_fhs_sp500

Modelling the Quantile Directly: CAViaR

- CAViaR Engle & Manganelli (2004): the quantile follows its own recursion; SAV specification
  - CAViaR: Conditional Autoregressive Value at Risk
  - $\mathrm{VaR}_t = \beta_0 + \beta_1 \mathrm{VaR}_{t-1} + \beta_2 |r_{t-1}|, \quad q_t = -\mathrm{VaR}_t$
  - SAV (symmetric absolute value): $\beta_1 \in (0,1)$ is the persistence, $\beta_2 > 0$ the response to yesterday's absolute return
  - A large move of either sign raises VaR; q_t: the return quantile
- **Quantile regression** Koenker & Bassett (1978): minimise the pinball loss
  - with no assumption on the shock distribution
  - $\hat{\beta} = \arg\min_\beta \sum_t (\alpha - 1\{r_t < q_t(\beta)\})(r_t - q_t(\beta))$
  - $q_t(\beta)$: the quantile path implied by $\beta = (\beta_0, \beta_1, \beta_2)$; elicitability turned into an estimator (optimiser settings in the Appendix)
- S&P 500, $\alpha = 1\%$, estimated on 2000–2015 (4 024 days): $\hat{\beta}_0 = 0.090$, $\hat{\beta}_1 = 0.897$, $\hat{\beta}_2 = 0.246$
  - Half-life of a shock to VaR: $\ln 0.5 / \ln \hat{\beta}_1 = 6.4$ days
- CAViaR reacts linearly to $|r_{t-1}|$, GARCH to r_{t-1}^2

CAViaR vs Filtered Historical Simulation, Out of Sample (1/2)



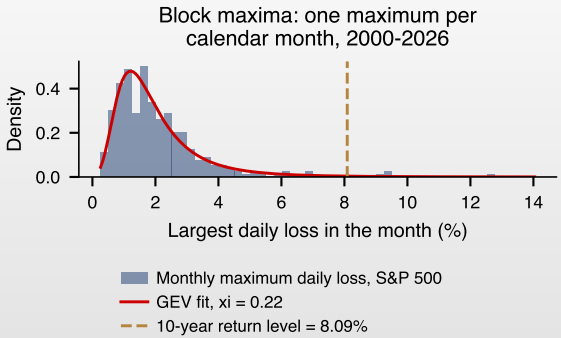
MFM_ch7_caviar ▶ Interactive

CAViaR vs Filtered Historical Simulation, Out of Sample (2/2)

- Dashed blue line: CAViaR; red line: FHS
  - bars: the daily loss (positive part)
  - historical simulation is reported, not drawn
- S&P 500, 2016–2026 (2 693 days not used in estimation): CAViaR with parameters fixed on 2000–2015
  - FHS re-estimated each January (a different refitting design)
- Exceedances of VaR 1%: CAViaR 1.30% (35), FHS 1.30% (35), historical 500 days 1.41% (38)
  - mean pinball loss (lower is better): 0.039, 0.037, 0.052
- COVID-19, 20 February – 30 April 2020 (50 days)
  - exceedances 6 / 3 / 9
  - peak VaR 13.5% / 18.0% / 4.5%: after a crash FHS jumps higher
- Counts alone do not establish that one model is better
  - significance is a Chapter 8 question
- Type of evidence: comparison of procedures

MFM_ch7_caviar

Monthly Maxima of S&P 500 Losses



- The question: how large is the daily S&P 500 loss exceeded once every ten years?
- Histogram: the largest daily loss in each calendar month (321 months, 2000–2026)
  - red curve: GEV (Generalised Extreme Value) distribution fitted by maximum likelihood
- $\hat{\xi} = 0.22$ (Fréchet), $\hat{\mu} = 1.37$, $\hat{\sigma} = 0.79$: a power-law tail
- Return level, exceeded on average once per period: 10 years 8.09% (dashed line), 30 years 10.95%
  - largest monthly maximum 12.77%
- Monthly maxima do not give a daily VaR directly
  - the conversion assumes independent days, which clustering violates (▸ Appendix)
- One value per month wastes data: the next method uses all large losses

MFM_ch7_gev ▶ Interactive

Block Maxima and the GEV Distribution (1/2)

- Split the sample into blocks of n days
  - keep each block maximum $M_n = \max(L_1, \ldots, L_n)$
  - Fisher–Tippett–Gnedenko theorem Fisher & Tippett (1928), Gnedenko (1943): if $(M_n - b_n)/a_n$ has a non-degenerate limit for i.i.d. losses, the limit is the generalised extreme value (GEV) distribution
  - $a_n > 0$, b_n: normalising constants that rescale and recentre the maximum
- The GEV distribution function of a block maximum
  - $H_\xi(x) = \exp\left\{-\left(1+\xi\,\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right\}$, $1+\xi(x-\mu)/\sigma > 0$ ($\xi = 0$: $\exp\{-e^{-(x-\mu)/\sigma}\}$)
  - μ: location; $\sigma > 0$: scale; ξ: shape (the tail type, on the slide GEV (2/2); reused symbols: ▸ Appendix)

Block Maxima and the GEV Distribution (2/2)

- The **shape** ξ decides the tail
  - $\xi > 0$: Fréchet, power-law tail with tail index $1/\xi$ (financial returns)
  - $\xi = 0$: Gumbel domain, the normalised excesses have an exponential limit (e.g. Normal, lognormal)
  - within the Gumbel domain the lognormal tail is still heavier than any exponential tail
  - $\xi < 0$: Weibull, bounded tail
- **Return level** for k blocks: exceeded on average once every k blocks
  - $x_k = H_\xi^{-1}(1 - 1/k) = \mu + \frac{\sigma}{\xi}\left[(-\ln(1 - 1/k))^{-\xi} - 1\right]$
  - 10 years: $k = 120$ months
- **Block size**: one month (about 21 trading days) is a compromise Embrechts, Klüppelberg & Mikosch (1997), Longin (2000)
  - longer blocks bring the maximum closer to its GEV limit but leave fewer maxima

Peaks over Threshold and the GPD (1/2)

- Fix a high threshold u
  - model the **excesses** $Y = L - u$ given $L > u$
  - Pickands–Balkema–de Haan theorem Pickands (1975), Balkema & de Haan (1974): for a wide class of distributions, as u approaches the right end-point of the distribution the excesses follow a generalised Pareto distribution
- The GPD distribution function of the excess
  - $G_{\xi,\beta}(y) = 1 - (1 + \xi y/\beta)^{-1/\xi}$ $(\xi \neq 0)$, $1 - e^{-y/\beta}$ $(\xi = 0)$
  - $y \geq 0$, and $y < -\beta/\xi$ if $\xi < 0$; scale $\beta > 0$: the typical size of an excess; shape ξ: the same as in the GEV, $\xi > 0$ = heavy tail

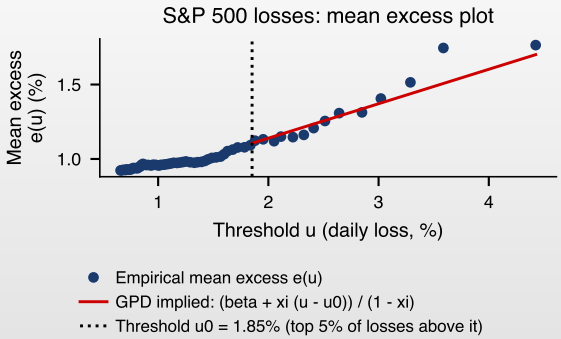
Peaks over Threshold and the GPD (2/2)

- Tail estimator with N_u exceedances among n observations McNeil, Frey & Embrechts (2015)
  - $\widehat{\mathrm{VaR}}_\alpha = u + \frac{\hat\beta}{\hat\xi}\left[\left(\frac{n}{N_u}\alpha\right)^{-\hat\xi} - 1\right]$, $\quad \widehat{\mathrm{ES}}_\alpha = \frac{\widehat{\mathrm{VaR}}_\alpha}{1-\hat\xi} + \frac{\hat\beta - \hat\xi u}{1-\hat\xi}$ ($\alpha < N_u/n$, $\hat\xi \neq 0$; ES needs $\hat\xi < 1$)
  - In words: VaR is the threshold u plus the GPD quantile of the excess that leaves probability α in the tail; N_u/n is the share of days above u
  - ES adds to VaR the mean excess beyond VaR, which grows with $\hat\xi$; $\hat\xi \geq 1$ means an infinite mean
- Bias–variance trade-off: a low u violates the limit theorem, a high u leaves few exceedances

Choosing the Threshold: the Mean Excess Function

- **Mean excess function**: $e(u) = \mathbb{E}[L - u \mid L > u]$
  - empirically, the average excess above u, plotted against u
- For a GPD tail $e(u)$ is **linear** in u
  - $e(u) = \dfrac{\beta + \xi(u - u_0)}{1 - \xi}$ for $u \geq u_0$, finite only if $\xi < 1$; slope $\xi/(1-\xi)$: upward for heavy tails, flat for exponential tails
  - u_0: the lowest threshold above which the GPD holds; β: the GPD scale at u_0
- Rule of thumb: choose u where the plot becomes roughly linear
  - check the stability of $\hat{\xi}$ across thresholds
- The Hill estimator Hill (1975), on the largest $k+1$ losses (here k does not count blocks), $L^{(1)} \geq \cdots \geq L^{(k+1)}$
  - $\hat{\xi}^H = \frac{1}{k} \sum_{i=1}^{k} \ln L^{(i)} - \ln L^{(k+1)}$
  - the tail index is $1/\hat{\xi}^H$ (Chapter 1)

Mean Excess Plot: S&P 500

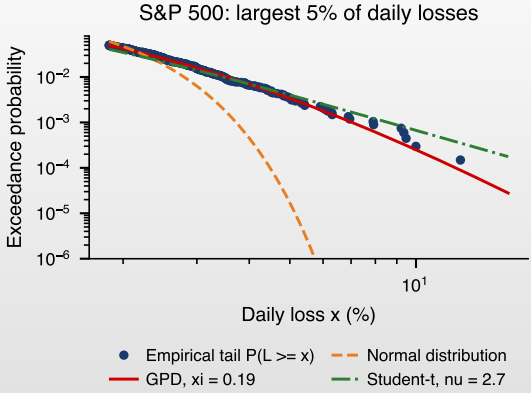


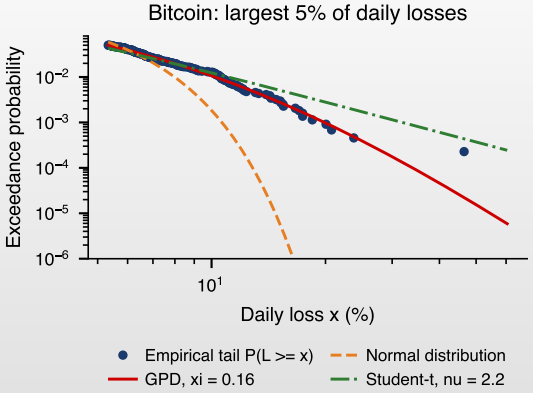
- Thresholds with 20% down to 0.5% of daily losses above them
  - red line: GPD fitted to the largest 5%
- The mean excess rises with u: compatible with a GPD tail with $\xi > 0$
  - not the flat line of an exponential tail
  - to be checked by parameter stability and other diagnostics
- The threshold $u_0 = 1.85\%$ is the 95% quantile of the losses
  - it leaves $N_u = 336$ exceedances, and above it the plot is roughly linear
  - empirical mean excess near u_0: 1.09% (maximum-likelihood GPD: $\hat{\beta}/(1-\hat{\xi}) = 1.11\%$)
  - GPD slope $\hat{\xi}/(1-\hat{\xi}) = 0.23$
- Beyond VaR 1% the points scatter: each averages very few losses

MFM_ch7_mean_excess ▶ Interactive

▸ Further Topics: GPD estimator, step by step

GPD Tails: S&P 500 and Bitcoin (1/2)





MFM_ch7_gpd_tail ▶ Interactive

GPD Tails: S&P 500 and Bitcoin (2/2)

- Log–log scale: probability of a daily loss above x, largest 5% of losses
- The Normal tail falls far below the data
  - the Student-t tail ($\nu = 2.7$) follows the GPD inside the data
  - but it is heavier beyond the data: tail index 2.7 vs 5.3
- S&P 500: $\hat{\xi} = 0.19$ (SE 0.065), tail index 5.3
  - Bitcoin: $\hat{\xi} = 0.16$ (SE 0.078), $\mathrm{VaR}_{1\%} = 10.23\%$ (all five series: Appendix)
- $\mathrm{SE}(\hat{\xi}) = (1 + \hat{\xi})/\sqrt{N_u}$ assumes independent exceedances
  - clustering makes it too small (Further Topics)
- Type of evidence: description

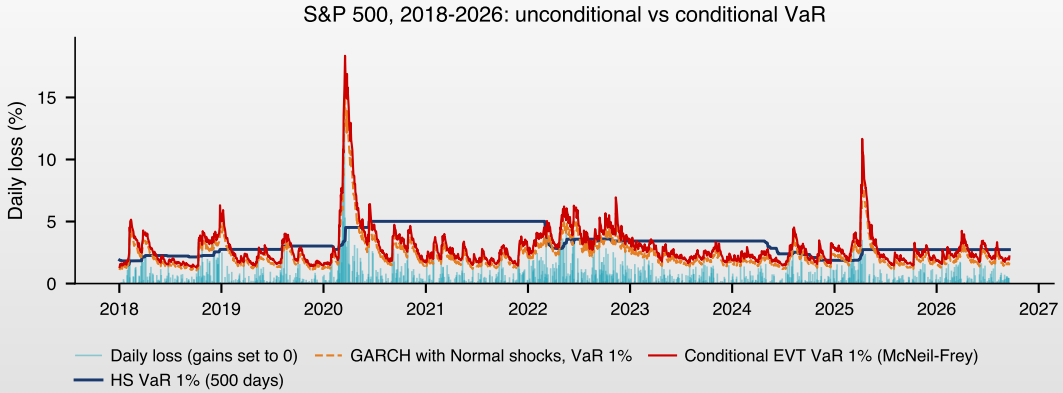
MFM_ch7_gpd_tail

Conditional EVT (McNeil and Frey)

- McNeil & Frey (2000): the FHS pipeline, with one step changed
- Filter: AR(1)-GARCH(1,1) by quasi-maximum likelihood, residuals $\hat{z}_t$, forecasts $\hat{\mu}_{t+1}$, $\hat{\sigma}_{t+1}$
- **Tail step**: a GPD instead of the empirical quantile $-q_\alpha(\hat{z})$
  - GPD fitted to the largest 10% of $-\hat{z}$ (threshold u_z, N_z exceedances out of n residuals)
  - The peaks-over-threshold (POT) formulas applied to the residuals: $\hat{z}_\alpha^{GPD} = u_z + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\frac{n}{N_z}\alpha\right)^{-\hat{\xi}} - 1\right]$,

    $\hat{e}_\alpha^{GPD} = \frac{\hat{z}_\alpha^{GPD} + \hat{\beta} - \hat{\xi}u_z}{1-\hat{\xi}}$
  - $\mathrm{VaR}_{\alpha,t+1} = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\hat{z}_\alpha^{GPD}, \qquad \mathrm{ES}_{\alpha,t+1} = -\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}\hat{e}_\alpha^{GPD}$
- The 10% threshold follows McNeil & Frey (2000) ($N_z = 100$ exceedances of 1 000 residuals)
  - filtered residuals have thinner tails and unclustered extremes, so the threshold can be lower
- A smooth parametric tail beyond the data
  - matters most far in the tail (VaR 0.5%, VaR 0.1%), where only a handful of residuals exist
- S&P 500, first trading day after 18 September 2026
  - VaR 1% = 2.03% with conditional EVT, against 1.97% with FHS
  - 90% FHS bootstrap interval [1.90, 2.09]%, which contains the conditional EVT value

Conditional EVT on the S&P 500 (1/2)



MFM_ch7_cond_evt ▶ Interactive

Conditional EVT on the S&P 500 (2/2)

- One-day-ahead $\mathrm{VaR}_{1\%}$, 2018–2026, parameters re-estimated each January
- Exceedance rate (2 190 days, target 1%)
  - conditional EVT 1.23%, GARCH with Normal shocks 2.42%
  - historical simulation over 500 days 1.60%
- GARCH with Normal shocks follows the rises and falls of risk correctly, but its VaR is too low
- Conditional EVT keeps the dynamics and corrects the level
  - the tail of the shocks matters, not only their scale
- Type of evidence: comparison of procedures
  - exceedance rates only; the tests are in Chapter 8

MFM _ch7_ cond _evt

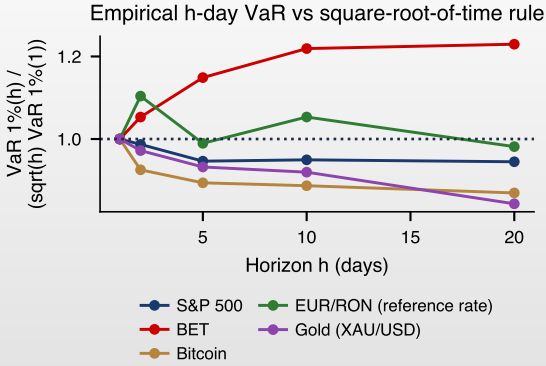
The Square-Root-of-Time Rule: Assumptions

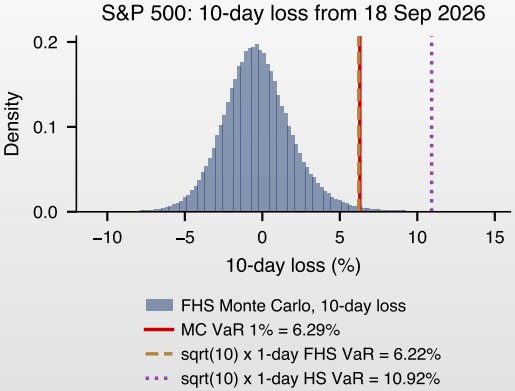
- The question: when can we get the h-day risk by multiplying the one-day risk by $\sqrt{h}$?
- i.i.d. $N(0, \sigma^2)$ daily losses: $\mathrm{VaR}_\alpha^{(h)} = \sqrt{h}\,\mathrm{VaR}_\alpha^{(1)}$, $\mathrm{ES}_\alpha^{(h)} = \sqrt{h}\,\mathrm{ES}_\alpha^{(1)}$
  - $\mathrm{VaR}_\alpha^{(h)}$, $\mathrm{ES}_\alpha^{(h)}$: the measures for the loss summed over h days; the standard deviation of a sum of h i.i.d. terms grows like $\sqrt{h}$
  - The rule needs zero mean, no autocorrelation, constant volatility and Normal shocks
- Basel 1996 allowed $\sqrt{10} = 3.162$ Basel Committee (1996)
- Each assumption fails in the data
  - Chapters 1, 2, 5
  - the next slide shows in which direction the rule errs

When the $\sqrt{h}$ Rule Fails

- **Autocorrelation**: $\mathrm{Var}(r_{t+1} + \cdots + r_{t+h}) = h\sigma^2\,\mathrm{VR}(h)$
  - variance ratio $\mathrm{VR}(h) = 1$ without autocorrelation (Chapter 2)
  - $\mathrm{VR}(h) > 1$ (positive autocorrelation): $\sqrt{h}$ **underestimates** the h-day variance
  - VaR and ES follow the variance only in a location-scale family, e.g. the Normal distribution
- **Volatility clustering**: volatility reverts to its mean over h days
  - Calm today: the h-day variance exceeds $h\,\sigma^2_{t+1}$, so $\sqrt{h}$ tends to give too low a risk
  - Turbulent today: $\sqrt{h}$ tends to give too high a risk
  - For the quantile, the direction is checked by filtered Monte Carlo, with paths simulated from the GARCH model (next slides)
- **Heavy tails**: for i.i.d. shocks, $\sqrt{h}$ can **overstate** the far-tail quantile Daníelsson (2011)
  - tail index $a > 2$: $P(L > x)$ decays like x^{-a} (regular variation)
  - with h fixed and $\alpha \to 0$, the quantile of the sum grows like $h^{1/a} < \sqrt{h}$
  - at the finite levels of the course (1%, 2.5%) this is an approximation, not a rule
- With jumps the rule can badly misstate multi-day risk Daníelsson & Zigrand (2006)

Square-Root-of-Time in the Data (1/2)





MFM_ch7_horizon ▶ Interactive

Square-Root-of-Time in the Data (2/2)

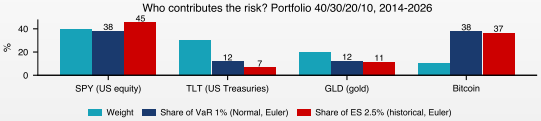
- Left: empirical $\mathrm{VaR}_{\mathbf{1}\%}$ of overlapping h-day losses divided by $\sqrt{h}$ times the one-day value
  - overlapping sums: far fewer independent observations
  - right: filtered Monte Carlo (MC in the legend), 200 000 paths
- Empirical ratio (left panel) at $h = 10$
  - BET 1.22 (lag-1 autocorrelation 0.109) and EUR/RON 1.05: $\sqrt{h}$ too low
  - S&P 500 0.95, gold 0.92, Bitcoin 0.89: conservative
- S&P 500 10-day $\mathrm{VaR}_{\mathbf{1}\%}$: filtered Monte Carlo 6.29%, $\sqrt{10}\times$ filtered one-day 6.22%
  - $\sqrt{10}\times$ historical one-day 10.92%
  - the gap to 10.92% comes from moving from unconditional (historical) to conditional (filtered) VaR
- Type of evidence: description

MFM_ch7_horizon

Quick Check: From One Day to Ten Days

- A desk reports a one-day $\mathrm{VaR}_{1\%} = 2\%$ from i.i.d. Normal losses with mean 0
  - (a) What is the 10-day $\mathrm{VaR}_{1\%}$ under the square-root-of-time rule?
  - (b) Returns turn out to be positively autocorrelated with $\mathrm{VR}(10) = 1.21$; returns stay Normal. What is the 10-day $\mathrm{VaR}_{1\%}$ now?
  - (c) No autocorrelation, but today's GARCH volatility is well below its long-run level. Is $\sqrt{10}\times$ the conditional one-day VaR too low or too high, and why?
- Answers in the ▸ Appendix

Who Contributes the Risk?



- The question: how much does diversification cut risk, and which asset produces the risk that remains?
- Weights: SPY 40%, TLT 30%, GLD 20%, Bitcoin 10%, rebalanced daily
  - SPY: S&P 500 ETF (ETF: exchange-traded fund)
  - TLT: long-term US (United States) Treasury ETF; GLD: gold ETF
- Simple returns on 3 017 common days, 2014–2026
  - teal: weights, blue: shares of Normal $\mathrm{VaR}_{1\%}$, red: shares of historical $\mathrm{ES}_{2.5\%}$ (the Euler contributions of the next slides)
- Portfolio $\mathrm{VaR}_{1\%} = 1.77\%$ vs 3.11% undiversified (-43%)
- Portfolio $\mathrm{ES}_{2.5\%} = 2.16\%$ vs 3.88% undiversified (-44%)
- Bitcoin: 10% of the money, 37.8% of VaR and 36.6% of ES
- TLT: 30% of the money, only 6.6% of ES
  - it lost 0.47% on average on the portfolio's worst 2.5% of days
- Type of evidence: description

MFM_ch7_components ▶ Interactive

Portfolio VaR and Euler Allocation (1/2)

- Setting: portfolio loss $L_p = -w^\top R$
  - w: vector of weights; R: vector of simple returns, with covariance matrix Σ; σ_i: standard deviation of asset i
- Normal returns, mean ignored: VaR is a multiple of the portfolio standard deviation
  - $\mathrm{VaR}_\alpha(L_p) = z_{1-\alpha}\sigma_p, \quad \sigma_p = \sqrt{w^\top \Sigma w}$; $z_{1-\alpha} = -z_\alpha > 0$
  - For long-only $w \geq 0$, the undiversified VaR $\sum_i z_{1-\alpha} w_i \sigma_i$ (the sum of stand-alone VaRs) is larger unless all correlations are 1

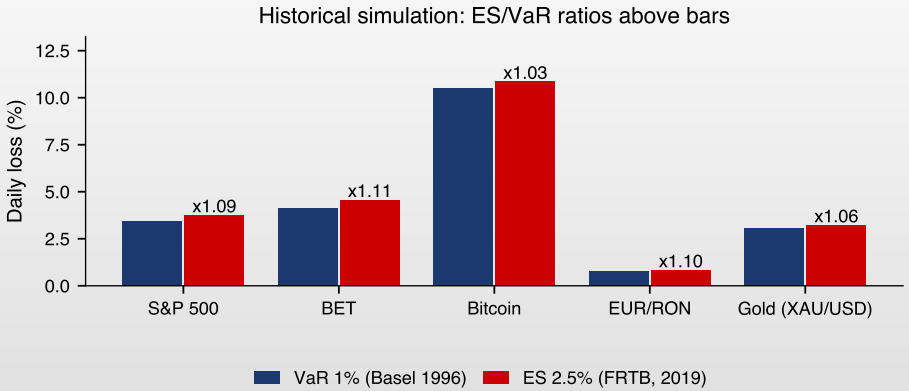
Portfolio VaR and Euler Allocation (2/2)

- **Marginal VaR**: the change in portfolio VaR per marginal increase of weight w_i
  - $\partial \mathrm{VaR}_\alpha / \partial w_i = z_{1-\alpha} (\Sigma w)_i / \sigma_p$; $(\Sigma w)_i$: the covariance of asset i with the portfolio
- **Component VaR** $C_i = w_i \, \partial \mathrm{VaR}_\alpha / \partial w_i$: the part of VaR attributed to asset i
  - Euler's theorem for a function homogeneous of degree 1: $\sum_i w_i \, \partial \rho / \partial w_i = \rho$, so the components add up, $\sum_i C_i = \mathrm{VaR}_\alpha$ Tasche (2008)
- ES, no distributional assumption (continuous losses)
  - the mean loss of asset i on the portfolio's tail days
  - $C_i^{ES} = \mathbb{E}[-w_i R_i \mid L_p \geq \mathrm{VaR}_\alpha(L_p)]$, $\quad \sum_i C_i^{ES} = \mathrm{ES}_\alpha(L_p)$, estimated historically as the fractionally weighted average over the tail days (the weights of historical ES)
  - A negative C_i means the asset **hedges** the portfolio
- Why Euler Tasche (2008)
  - homogeneity of degree 1 makes the contributions add up to the total risk, but other allocation rules do so too
  - Euler is characterised by compatibility with RORAC (return on risk-adjusted capital), for differentiable ρ
  - that is, it correctly shows where a larger weight improves the return per unit of risk

Liquidity-Adjusted VaR

- The question: how much does selling the position in a stressed market add to VaR?
- VaR assumes a sale at the mid price $=$ (bid $+$ ask)$/2$
  - an immediate sale goes at the bid and costs half the **relative spread** $S = (\text{ask} - \text{bid})/\text{mid}$
- The spread widens exactly when markets fall
- Bangia, Diebold, Schuermann & Stroughair (2001): LVaR (Liquidity-adjusted Value-at-Risk), $\mathrm{LVaR}_\alpha = \mathrm{VaR}_\alpha + \frac{1}{2}(\mu_S + k\,\sigma_S)$
  - μ_S, σ_S: the mean and standard deviation of S (here S is the spread, not skewness)
  - k: the number of standard deviations that take the spread to its stressed level (not the EVT k)
- Banca Transilvania, a position of 1 million RON
  - the one-day historical VaR 1% over the last 496 days is 3.16%
- With assumed $\mu_S = 0.5\%$, $\sigma_S = 0.3\%$ and $k = 3$: $\mathrm{LVaR} = 3.86\%$ (+22%)
  - In money: VaR $\approx$ 31 581 RON, spread cost $\frac{1}{2}(0.5 + 3 \times 0.3)\% = 0.70\% \to$ 7 000 RON, LVaR $\approx$ 38 581 RON
- Large positions also move the price (endogenous liquidity)
  - that is why FRTB handles liquidity through liquidity horizons of up to 120 days

VaR 1% vs ES 2.5% on Five Assets (1/2)



MFM_ch7_basel ▶ Interactive

VaR 1% vs ES 2.5% on Five Assets (2/2)

- Historical simulation, daily losses, full samples
  - number above each pair: $\mathrm{ES}_{\mathbf{2.5\%}}/\mathrm{VaR}_{\mathbf{1\%}}$
- Normal benchmark: $2.338/2.326 \approx 1.005$, the same capital for thin tails
  - Student-t: 1.01 at $\nu = 30$, 1.07 at $\nu = 4$
- Data: S&P 500 1.09, BET 1.11, Bitcoin 1.03, EUR/RON 1.10, gold 1.06
  - the switch from VaR 1% to ES 2.5% raises the one-day measure by 3–11%
- ES 2.5% need not exceed VaR 1%
  - the inequality $\mathrm{ES}_\alpha \geq \mathrm{VaR}_\alpha$ is guaranteed only at the same level α
- Stressed calibration and liquidity horizons add more: Seminar 7, C2
- Type of evidence: description

MFM_ch7_basel

From Basel 1996 to FRTB (2019)

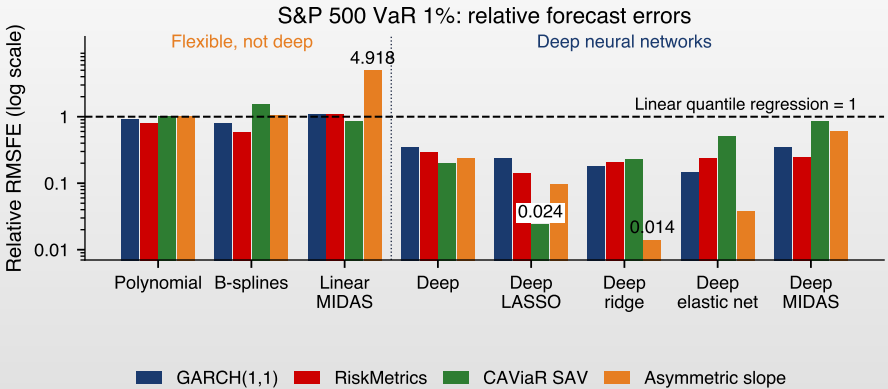
	Basel 1996 Basel Committee (1996)	FRTB 2019 Basel Committee (2019)
Risk measure	VaR 1%, computed daily	ES 2.5%, computed daily for every trading desk
Horizon	10 days; $\sqrt{10}$ scaling allowed	10-day base, scaled by liquidity horizons (large-cap equity 10, small-cap 20, credit spreads up to 120 days)
Data	at least one year of history	calibrated to the most severe 12-month stress period
Capital	max(yesterday's VaR, multiplier $\geq 3 \times$ 60-day average VaR); the multiplier rises by 0–1 after poor backtests (traffic-light zones, Chapter 8)	multiplier $m_c = 1.5$, increased after poor backtests
Backtesting	VaR 1% exceptions (Chapter 8)	one-day VaR 1% and VaR 2.5%: VaR tests, ES sets the capital

- Critique after 2008: VaR ignores the size of tail losses and one-year windows forget crises

Case Study (1/3): Can a Deep Network Improve VaR Forecasts?

- **Question** Chronopoulos, Raftapostolos & Kapetanios (2024) (*Journal of Financial Econometrics*)
  - Does VaR forecasting improve if the conditional quantile may depend on its predictors in an **unknown non-linear** way, learned by a deep neural network?
- Data (paper, Section 3): S&P 500 daily log returns, September 1985 – August 2020
  - $\alpha = 1\%, 5\%, 10\%$
- Predictors x_t: GARCH or RiskMetrics volatility, CAViaR state variables
- Last 2 000 days out of sample, one fixed tuning window (details: Appendix)
- Benchmark: linear quantile regression Koenker & Bassett (1978) on the same predictors
  - i.e. a linear CAViaR-type specification

Case Study (2/3): Relative Forecast Error, $\alpha = 1\%$



MFM_ch7_crk_table6 ▶ Interactive

Case Study (3/3): What the Results Show and Do Not Show

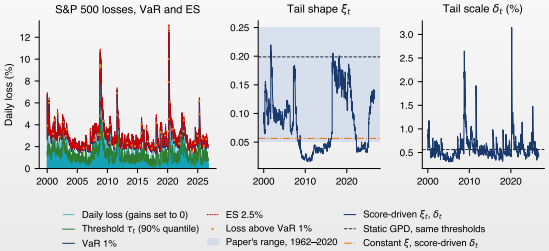
- Chronopoulos, Raftapostolos & Kapetanios (2024), Table 6: RMSFE (Root Mean Squared Forecast Error) relative to linear quantile regression (dashed line = 1), log scale
- Flexible non-network models: 7 of 12 bars at or above 1
  - polynomial, B-splines, linear MIDAS (MIDAS = mixed data sampling regression)
  - all 16 deep-network bars and the 4 deep MIDAS bars below 1 (0.014–0.850)
  - all significant in the DM (Diebold–Mariano, test) test except deep MIDAS for SAV
- Caution: RMSFE is minimised by the conditional **mean**, not the quantile
  - quantile accuracy needs the pinball loss
- Whether the largest cut, 98%, survives rolling-window re-estimation is open
  - the cut comes from one fixed tuning window (Seminar 7, C3)
- Type of evidence: comparison of procedures
  - the results do not isolate the contributions of architecture, penalisation and CV (cross-validation)

MFM_ch7_crk_table6

Second Case Study (1/4): Does the Tail Shape Change over Time?

- D'Innocenzo, Lucas, Schwaab & Zhang (2024), *Journal of Business & Economic Statistics*: the shape and scale of the loss tail change from day to day
  - the threshold τ_t: a dynamic 90% quantile of the loss, as in Patton, Ziegel & Chen (2019)
  - exceedances over the threshold follow a GPD with shape ξ_t and scale δ_t
  - ξ_t and δ_t are updated after each exceedance by the score, the derivative of the log-likelihood: a **score-driven** model
- Why it matters: conditional EVT McNeil & Frey (2000) keeps ξ fixed
  - a conditional GPD tail with $\xi_t = 0.25$ has finite moments of the exceedance only for orders $p < 1/\xi_t = 4$
  - the fourth moment of the exceedance is infinite (in the modelled tail, not necessarily in the unconditional distribution)
  - a larger ξ_t raises ES more than VaR
- The paper's claims, on the S&P 500 (Jul 1962 – Dec 2020, 14 726 days, 1 495 exceedances)
  - ξ_t varies between about 0.05 and 0.25
  - one-off crashes (Black Monday 1987, the 2010 flash crash) raise mainly ξ_t
  - long crises (2008–2009, COVID-19 in 2020) raise mainly δ_t
  - 1.0% of the losses exceed the filtered VaR 1%

Second Case Study (2/4): Tail Shape and Scale of the S&P 500, 2000–2026



- Paths filtered in-sample, with parameters and starting quantile estimated on the whole sample; the recursion for τ_t, ξ_t, δ_t, VaR and ES is causal given them
- Left: VaR 1% peaks at 12.4% on 17 Mar 2020
- Middle: ξ_t between 0.016 (17 Mar 2011) and 0.22 (18 Sep 2001)
  - band: the range reported in the paper
  - orange line: $\xi = 0.056$, when only δ_t is dynamic
- Right: δ_t between 0.30% and 3.15% (17 Mar 2020)
- Dashed lines: static GPD on the same thresholds ($\hat{\xi} = 0.20$)

MFM_ch7_tailshape ▶ Interactive

Second Case Study (3/4): The Model and the Replication on the Course Data

- Threshold: $\tau_{t+1} = (1 - b_\tau)\,\hat{q}_{0.9} + a_\tau\,(1\{L_t > \tau_t\} - 0.1) + b_\tau\,\tau_t$
  - $\hat{q}_{0.9}$: the 90% quantile of the losses; a_τ, b_τ minimise the pinball loss at 90%
- The exceedance $x_t = L_t - \tau_t > 0$ follows a GPD(ξ_t, δ_t), with $f_t = (\ln \xi_t, \ln \delta_t)'$
  - $f_{t+1} = \omega + A\,s_t + B\,f_t$, with $A = \mathrm{diag}(a_\xi, a_\delta)$ and $B = \mathrm{diag}(b_\xi, b_\delta)$
  - s_t: the score scaled by the Fisher information; $s_t = 0$ on days without an exceedance
- VaR and ES at level α: $\mathrm{VaR}_t = \tau_t + \dfrac{\delta_t}{\xi_t}\left[\left(\dfrac{\alpha}{p_t}\right)^{-\xi_t} - 1\right]$, $\mathrm{ES}_t = \dfrac{\mathrm{VaR}_t + \delta_t - \xi_t \tau_t}{1 - \xi_t}$
  - p_t: the share of days with an exceedance up to $t - 1$
- Replication: the same specification, on the course data (2000–2026, 6 717 days, 680 exceedances) instead of CRSP (Center for Research in Security Prices), 1962–2020
  - the states are truncated numerically: $\ln \xi_t \in [-6, 1.5]$ and $\ln \delta_t \in [-6, 4]$
  - comparison by AIC (Akaike information criterion) with a static GPD and with constant ξ and dynamic δ_t, on the same thresholds
- The same estimation on 2000–2015 only (404 exceedances) fails numerically: the states hit their bounds
  - ξ_t sits at the lower bound $e^{-6} = 0.0025$ on 3 939 of 4 024 days
  - ξ_t sits at the upper bound $e^{1.5} = 4.48$ on 5 days, on which ES is infinite
  - with the lower bound moved to e^{-12} the optimum follows it, and the log-likelihood is practically unchanged
  - the exponential limit ($\xi = 0$, dynamic δ_t): AIC 454, against 458 with dynamic ξ_t and 515 for the static GPD

Second Case Study (4/4): The Tail Scale Clearly Varies, the Shape Only Weakly

Quantity	Paper, 1962–2020	Course data, 2000–2026
a_ξ; b_ξ (shape)	0.025; 0.9993	0.130; 0.9975
a_δ; b_δ (scale)	0.112; 0.995	0.160; 0.986
a_τ; b_τ (threshold)	0.241; 0.989	0.365; 0.985
Losses above VaR 1%	1.0%	0.91% (61 days)
Mean loss above VaR 1%; mean ES 1% over all days	3.54%; 3.12%	4.02%; 3.65%

- AIC, lower is better: 774 with dynamic ξ_t and δ_t, 772 with constant ξ, 853 for the static GPD
- The shape dynamics add only 1.04 to the log-likelihood, for two extra parameters
- One-off crashes: ξ_t rises more than δ_t after 11 Sep 2001 (+113% against +91%) and after Brexit (+189% against +114%), but not after the flash crash (−12% against +18%)
- Long crises: Sep 2008 – Mar 2009, ξ_t −58% and δ_t +113%; Feb – Apr 2020, ξ_t +60% and δ_t +136%
- On the 163 days above VaR 2.5%, the forecast ES 2.5% averages 3.18% and the loss 3.18%
- On the 61 days above VaR 1%, the forecast ES 1% averages 4.05% and the loss 4.02%
- Type of evidence: in-sample description (the parameters use the whole sample)
  - limit: no out-of-sample comparison on 2016–2026, because the 2000–2015 estimate fails numerically (static GPD: $\hat{\xi} = 0.12$)
  - closeness to $\xi = 0$ does not by itself prevent forecasting: the exponential limit is an alternative to test

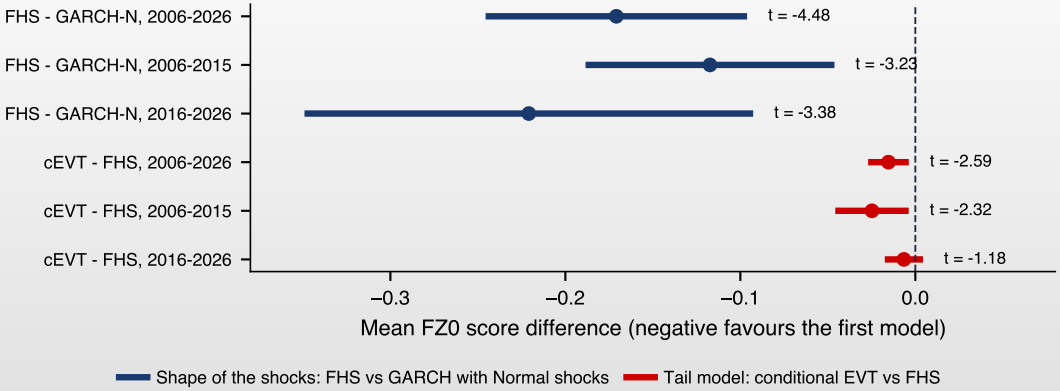
Can Daily Data Rank (VaR, ES) Models? (1/2): Score and Test

- Obstacles for (VaR 1%, ES 1%) Fissler & Ziegel (2016), Nolde & Ziegel (2017), Patton, Ziegel & Chen (2019), Bayer & Dimitriadis (2022), Pele & Mazurencu-Marinescu-Pele (2026)
  - ES alone is not elicitable
  - every day enters the FZ0 score, but extreme observations are rare and weigh heavily on the precision of the comparison
- FZ0 score: a joint penalty for a (VaR, ES) forecast given the realised return, lower is better Patton, Ziegel & Chen (2019)
  - $S(v, e; Y) = -\frac{1}{\alpha e} 1\{Y \le v\}(v - Y) + \frac{v}{e} + \ln(-e) - 1$
  - Y: the realised return; $v_t = -\mathrm{VaR}_{\alpha,t}$ and $e_t = -\mathrm{ES}_{\alpha,t}$: the forecast quantile and tail mean, both negative; the first term penalises returns below v
- Test: $d_t = S_t^A - S_t^B$, the daily score difference of models A and B
  - H_0: $\mathbb{E}[d_t] = 0$ (equal accuracy) vs H_1: $\mathbb{E}[d_t] \neq 0$; Diebold–Mariano statistic $t = \bar{d}/\widehat{\mathrm{SE}}_{\mathrm{HAC}}(\bar{d})$, with $\bar{d}$ the mean of d_t, approximately $N(0,1)$ under H_0; $t < -1.96$: A significantly better at 5%
  - We report the interval for $\mathbb{E}[d_t]$ separately on subperiods fixed before the analysis (2006–2015, 2016–2026)

Can Daily Data Rank (VaR, ES) Models? (2/2): What the Researcher Checks

- An AI assistant can draft the literature map, the code and a referee report
  - the researcher checks
  - Sign and level: VaR 1% $= -q_{1\%}$, positive for a loss (recompute one day of FZ0 step by step)
  - Look-ahead bias: each forecast uses only data before its day
  - Scoring: joint (VaR, ES) scores only, never the MSE (mean squared error) of ES alone Gneiting (2011)
  - Test power: count the tail days the comparison rests on
  - Every DOI (Digital Object Identifier) is resolved (reusable prompts: Appendix)

Mini-Case: Can We Rank ES Models? (1/2)



MFM_ch7_ai_discovery_case ▶ Interactive

Mini-Case: Can We Rank ES Models? (2/2)

- Claim tested: "conditional EVT is the best (VaR 1%, ES 1%) model for the S&P 500"
  - replicate first: FHS exceedance rate 1.54% (80 of 5 210 days, 2006–2026)
- Mean FZ0: conditional EVT 1.267, FHS 1.282, GARCH–Normal 1.453
  - bars: HAC 95% intervals for $\mathbb{E}[d_t]$
  - in the chart, cEVT = conditional EVT, GARCH-N = GARCH with Normal shocks
- All six Diebold–Mariano t-statistics
  - negative: the first model scores better; 5% critical value -1.96
  - FHS vs GARCH–Normal: 2006–2026 $t = -4.48$; 2006–2015 $t = -3.23$; 2016–2026 $t = -3.38$
  - cEVT vs FHS: 2006–2026 $t = -2.59$; 2006–2015 $t = -2.32$; 2016–2026 $t = -1.18$
- Decision rule: a gain counts as demonstrated only if significant at 5% in both halves
  - the two halves are pre-registered (fixed before the analysis)
  - the full-sample t is reported but is not decisive
- The shape of the shocks matters robustly: Normal shocks lose in both halves
- The EVT gain over FHS is ten times smaller
  - significant in 2006–2015 but not in 2016–2026 ($[-0.017, 0.004]$)
  - by the rule above the gain is not demonstrated, and this does not prove the two models are equivalent
- Type of evidence: comparison of procedures

MFM_ch7_ai_discovery_case

Your Turn: A Team-Project Seed

- Question: does a dynamic semiparametric (VaR, ES) model Patton, Ziegel & Chen (2019) beat FHS and conditional EVT?
  - out of sample, at ES 2.5%
  - Markets: the S&P 500, the BET and Bitcoin
  - Model to test: the one-factor GAS (generalised autoregressive score) model of Patton, Ziegel & Chen (2019) (their Section 2.3), estimated by minimising the FZ0 loss
  - Status: an optional extension for deeper study; the seminar team-project menu is Seminar 7, C2 and C3
- Replicate first: the S&P 500 FHS VaR 1% of 1.97% and the FZ0 ranking of the mini-case
- Deliverables
  - A one-page pre-registration: models, levels, scores, subperiods, power (tail days) and what would falsify the claim
  - Code on the daily market data of the course (starting point: Quantlet)
  - A table of all comparisons with adjusted p-values
  - A referee-style self-critique
- Log prompts and accepted outputs in `AI_USE.md`
  - every error caught in `AI_ERRORS.md`: flipped sign, wrong level label, look-ahead bias, invented DOI

Which Method When?

Situation	Method	Fails when
Quick check, long stable history	Historical simulation, 500–1000 days	volatility changes; nothing beyond the worst loss
Daily risk after a volatility shock	FHS or conditional EVT, with a bootstrap interval	the GARCH model is misspecified
VaR 0.5% and deeper, stress tests	Peaks over threshold (GPD), conditional or not	threshold badly chosen; clustered exceedances
Multi-day horizon, options	Filtered Monte Carlo	model risk dominates
Risk budgeting across desks or assets	Euler contributions to ES	few tail days behind each contribution
Regulatory capital	ES 2.5%, 10 days, stressed period, VaR backtests	the stress window is short and noisy

- Every VaR or ES is an estimate
  - report it with its tail count and an interval

Review Questions and the Link to Chapter 8

Retrieval questions

- Why can VaR reward concentration?
- Why can ES not reward it?
- Why does the precision of ES depend on $n\alpha$ rather than on n?
- When is $\sqrt{10}$ scaling too low?
- When is it too high?

Chapter 8: backtesting

- How often did losses exceed VaR?
- Were the exceedances independent?
- Coverage tests and the Basel traffic-light zones
- Comparing (VaR, ES) forecasts with joint scores Fissler & Ziegel (2016)

Further reading: Franke, Härdle & Hafner (2019), Härdle, Chen & Overbeck (2017), Čížek, Härdle & Weron (2011), McNeil, Frey & Embrechts (2015)
Related work, entropy and market risk: Pele, Lazar & Dufour (2017), Pele & Mazurencu-Marinescu-Pele (2019), Pele, Lazar & Mazurencu-Marinescu-Pele (2019)

The Practical Decision

- **What I choose**
  - for tomorrow's VaR 1% and ES 1%: FHS, as a comparative benchmark; S&P 500 after 18 September 2026: VaR 1% $= 1.97\%$ (90% interval $[1.90, 2.09]\%$), ES 1% $= 2.56\%$
  - the FZ0 advantage of FHS over GARCH with Normal shocks is relative and does not validate nominal coverage (a 1% exceedance rate)
  - conditional EVT at 1%: a better FZ0 score than FHS over 2006–2026 ($t = -2.59$), but not in 2016–2026 ($t = -1.18$)
  - VaR 0.5% and VaR 0.1%: a separate evaluation; the 1% results do not show that EVT is better at these levels
  - for ES 2.5% over a historical window: the fractional estimate with a HAC interval; S&P 500: 3.76%, HAC SE 0.26 pp on 6 717 days
- **Under which conditions**
  - the diagnostics show no important residual dependence at the thresholds and lags examined: extremal index 0.99 after filtering, against 0.19 before
  - enough tail days: on 500 days ES 2.5% rests on $n\alpha = 12.5$ losses and its HAC interval is 2.55 pp wide
  - the HAC interval for ES requires a finite second moment of the excess beyond VaR, $\mathbb{E}[(L - v)_+^2] < \infty$ (tail index above 2, estimated in Chapter 1), plus stationarity and weak dependence
  - $\sqrt{10}$ scaling: no autocorrelation is not enough, today's volatility and the tails matter too; for the BET the empirical ratio is 1.22
- **What result would make me drop it**
  - coverage rejected by the Kupiec test of Chapter 8: already here 1.54%, 80 exceedances in 5 210 days against 52.1 expected ($LR_{uc} = 12.97$, $p = 0.0003$), so FHS stays a comparative benchmark only
  - FHS exceedances that cluster in the independence tests
  - a pre-registered half in which FHS no longer beats GARCH with Normal shocks on the FZ0 score (here $t = -3.23$ and $t = -3.38$)

Student-t Fit to the S&P 500, Step by Step

- Maximum likelihood estimates on daily returns in %: $\hat{m} = 0.068$, $\hat{s} = 0.702$, $\hat{\nu} = 2.666$
  - Implied standard deviation $\hat{s}\sqrt{\hat{\nu}/(\hat{\nu}-2)} = 1.40\%$, not the sample $\hat{\sigma} = 1.21\%$: the scale $\hat{s}$, not $\hat{\sigma}$, enters the formulas
- **VaR**: quantile $t_{1\%} = t_{\hat{\nu}}^{-1}(0.01) = -5.028$
  - $\mathrm{VaR}_{1\%} = -(\hat{m} + \hat{s}\, t_{1\%}) = -0.068 + 0.702 \times 5.028 \approx 3.46\%$
- **ES**: density $g_{\hat{\nu}}(t_{1\%}) = 0.00490$, so $g_{\hat{\nu}}(t_{1\%})/\alpha = 0.490$
  - tail factor $(\hat{\nu} + t_{1\%}^2)/(\hat{\nu} - 1) = 16.77$
  - $\mathrm{ES}_{1\%} = -\hat{m} + \hat{s}\,\dfrac{g_{\hat{\nu}}(t_{1\%})}{\alpha}\,\dfrac{\hat{\nu} + t_{1\%}^2}{\hat{\nu} - 1} = -0.068 + 0.702 \times 0.490 \times 16.77 \approx 5.71\%$
- Why ES overshoots: with $\hat{\nu}$ close to 2 the tail factor is large
  - ES is 5.71% against the historical 5.03%, while VaR matches

MFM_ch7_methods_assets

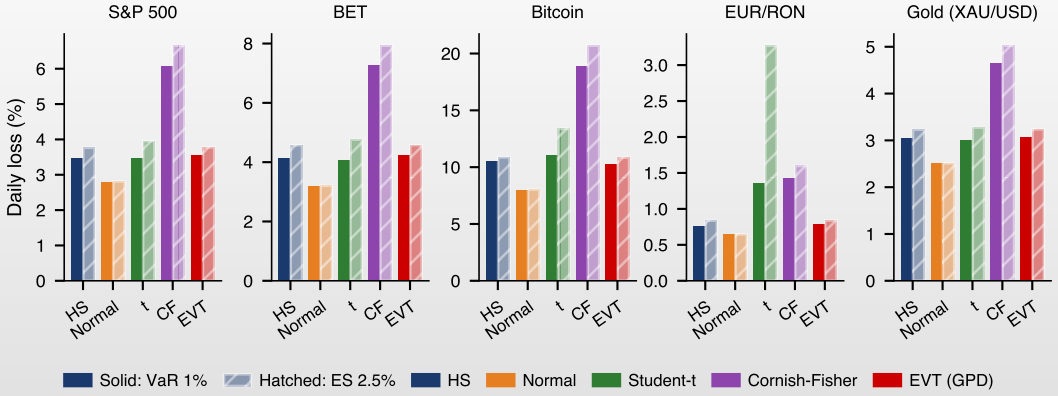
How Do We Compare Tail Shapes Fairly?

- To isolate tail shape, compare the models at the same mean and standard deviation (Seminar 7, A2)
- Example: zero mean, standard deviation 2% and $\nu = 4$
  - the Student-t scale is $s = 2\sqrt{(4-2)/4} = 1.414\%$, the Normal scale is 2%
- With $s = 2\%$ in both models the Student-t standard deviation would be larger
  - the difference would mix a scale effect with a tail-shape effect
- Fitted separately to the same data, the two models have different scales
  - the differences in the Estimation Methods section mix both effects

ES as an Optimisation Problem: Rockafellar–Uryasev

- **Theorem** Rockafellar & Uryasev (2000): for any integrable loss L
  - $\mathrm{ES}_\alpha(L) = \min_{v\in\mathbb{R}} F(v), \quad F(v) = v + \frac{1}{\alpha}\mathbb{E}[(L-v)_+], \quad (x)_+ = \max(x,0)$
  - v: a candidate threshold; $F(v)$: the threshold plus the expected excess of the loss over it, scaled up by $1/\alpha$
  - The minimiser is $\mathrm{VaR}_\alpha(L)$, the minimum value is $\mathrm{ES}_\alpha(L)$; $F(\mathrm{VaR}_\alpha)$ is the discrete ES formula of the two-bond slide
- F is convex in v and, for a portfolio loss $-w^\top R$, jointly convex in (w, v)
  - w: vector of weights; R: vector of asset returns
  - Convexity together with positive homogeneity gives subadditivity: $\mathrm{ES}_\alpha(L_1+L_2) = 2\,\mathrm{ES}_\alpha(\frac{1}{2}L_1+\frac{1}{2}L_2) \le \mathrm{ES}_\alpha(L_1)+\mathrm{ES}_\alpha(L_2)$ (proof in the Appendix)
  - With historical scenarios, minimising ES becomes a linear program (the Portfolio VaR section)
- S&P 500 losses, $\alpha = 2.5\%$: $\arg\min F = 2.51\%$ (the minimising v), $\min F = 3.76\%$
  - $\min F$ equals the historical ES 2.5% exactly
  - $\arg\min F$ equals the historical VaR 2.5% only after rounding: the empirical minimiser is an order statistic

Methods Compared on Five Assets (1/2)



MFM_ch7_methods_assets

Methods Compared on Five Assets (2/2)

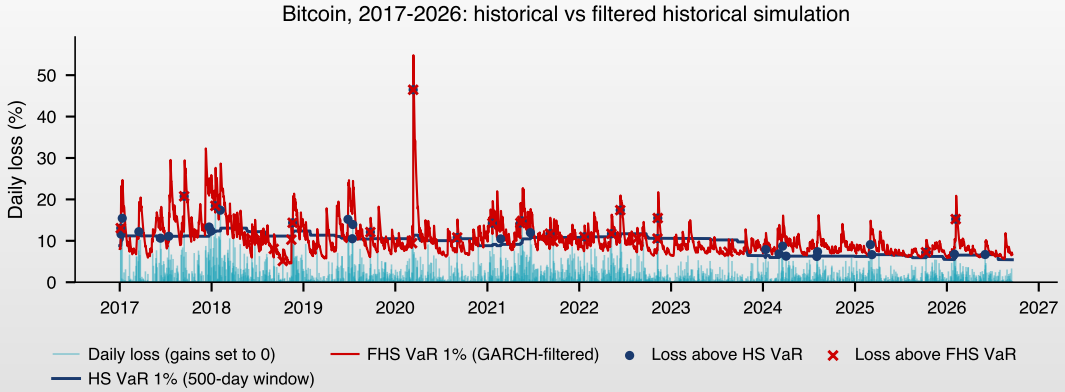
- Solid bars: $\mathrm{VaR}_{1\%}$; hatched bars: $\mathrm{ES}_{2.5\%}$
  - daily losses in %, full samples (all numbers in the Appendix)
  - bar labels: CF = Cornish–Fisher, t = Student-t
  - XAU/USD (XAU: the ISO 4217 code for gold): the gold price in US dollars per troy ounce
- Normal: too low everywhere (Bitcoin $\mathrm{VaR}_{1\%}$ 7.99% vs 10.51% historical)
- Cornish–Fisher: overshoots, since excess kurtosis exceeds 7 in every series
  - BET 7.25% vs 4.13%
- Student-t fails on EUR/RON: $\hat{\nu} = 1.25$ (infinite variance)
  - unrealistic ES 2.5% of 3.27%, against 0.84% historical
  - Many near-zero reference-rate changes drive $\hat{\nu}$ down
- EVT stays close to historical simulation for VaR 1% in all five series

MFM_ch7_methods_assets

Bootstrap and HAC: How the Intervals Are Built

- **Bootstrap** (i.i.d.): draw n losses with replacement from the sample and recompute $\widehat{\mathrm{VaR}}$ or $\widehat{\mathrm{ES}}$
  - Repeat $B = 2\,000$ times; the 95% interval runs from the 2.5% to the 97.5% quantile of the B recomputed values
- **Moving-block bootstrap**: draw blocks of 20 consecutive days instead of single days
  - join $n/20$ of them
  - Calm and turbulent stretches stay intact inside each block, so every resample keeps the volatility clustering
- **HAC long-run variance** of a series y_t (here $1\{L_t > \mathrm{VaR}_\alpha\}$ or $(L_t - \mathrm{VaR}_\alpha)_+$): $\hat\gamma_0 + 2\sum_{j=1}^{m}\left(1 - \frac{j}{m+1}\right)\hat\gamma_j$
  - $\hat\gamma_j$: sample autocovariance of y_t at lag j, $\hat\gamma_0$ its variance; Bartlett weights $1 - j/(m+1)$ shrink distant lags; lag $m = \lfloor 4(n/100)^{2/9} \rfloor$ ($m = 10$ for the full S&P 500 sample)
  - It replaces the i.i.d. variance $\hat\gamma_0$ in the SE formulas. Toy check: $\hat\gamma_0 = 1$, $\hat\gamma_1 = 0.3$, all other $\hat\gamma_j = 0$, $m = 1$: long-run variance $1 + 2 \times 0.5 \times 0.3 = 1.3$, so the SE grows by $\sqrt{1.3} \approx 1.14$
- Worked numbers for the BET: Seminar 7, B1 (i.i.d., HAC and block-bootstrap intervals for VaR 1%)

Historical vs Filtered Historical Simulation: Bitcoin (1/2)



MFM_ch7_hs_fhs_btc

Historical vs Filtered Historical Simulation: Bitcoin (2/2)

- Same design, forecasts from January 2017, 7 days a week
  - days with a loss above VaR (3 548 days): historical 1.16%, filtered 0.76%
  - GARCH with Normal shocks 1.58%
- Filtered VaR peaked at 54.83% in March 2020, after the crash of 2020-03-12
  - the crash: a log loss of 46.47%, i.e. a price fall of 37.17%
- Volatility has fallen since 2022 and the 500-day window follows with a long delay
  - crypto as a distinct asset class: Pele et al. (2023)

MFM_ch7_hs_fhs_btc

Worked Forecast: S&P 500 on the Next Trading Day

- AR(1)-GARCH(1,1) on 2000–2026: $\hat{\phi} = -0.053$, $\hat{\omega} = 0.025$, $\hat{\alpha}_1 = 0.1204$, $\hat{\beta}_1 = 0.8604$
  - persistence $\hat{\alpha}_1 + \hat{\beta}_1 = 0.981$
  - Forecasts for the first trading day after 18 September 2026: $\hat{\mu}_{t+1} = 0.056\%$, $\hat{\sigma}_{t+1} = 0.725\%$ (sample standard deviation 1.21%: a calm market)
- Residual quantile $\hat{\xi}_{1\%} = -q_{1\%}(\hat{z}) = 2.79$, above $-z_{1\%} = 2.326$
  - the filtered shocks are still heavy-tailed
  - $\mathrm{VaR}_{1\%,t+1} = -0.056 + 0.725 \times 2.791 = 1.97\%$
  - ES step: minus the average of the lowest 1% of the residuals is 3.604 (Normal: 2.665), so $\mathrm{ES}_{1\%,t+1} = -0.056 + 0.725 \times 3.604 \approx 2.56\%$
- Unconditional historical VaR 1%: 3.45%
- GARCH with Normal shocks gives a VaR 1% of 1.63% for the same day
  - all five assets: Appendix

How Uncertain Is the Next-Day VaR Forecast?

- Two estimated inputs Francq & Zakoïan (2015)
  - the GARCH parameters $\theta = (c, \phi, \omega, \alpha_1, \beta_1)$, through $\hat{\sigma}_{t+1}$
  - the residual quantile $\hat{\xi}_\alpha$
  - Residual bootstrap Christoffersen & Gonçalves (2005): simulate AR(1)-GARCH(1,1) paths with resampled $\hat{z}$, re-estimate on each path, forecast from the **original** history (algorithm in the Appendix)
- S&P 500, 999 paths: FHS $\mathrm{VaR}_{1\%} = 1.97\%$
  - 90% interval $[1.90, 2.09]\%$
  - at 95% each tail would hold only about 25 values
- The two sources
  - $\hat{\sigma}_{t+1} = 0.725\%$ $[0.708, 0.743]\%$
  - $\hat{\xi}_{1\%} = 2.79$ $[2.68, 2.88]$
  - Only the quantile varies: $[1.89, 2.03]$; only σ_{t+1} varies: $[1.92, 2.02]$: the residual quantile (about 67 tail residuals) is the larger source
- GARCH with Normal shocks (1.63%) lies far below the interval
  - the choice of tail model matters more than the estimation error of a given model

The GPD Tail Estimator on the S&P 500, Step by Step

- Inputs: $n = 6\,717$ daily losses, threshold $u = 1.852\%$, $N_u = 336$ exceedances
  - maximum likelihood $\hat{\xi} = 0.188$, $\hat{\beta} = 0.898$
- Step 1: $\dfrac{n}{N_u}\,\alpha = \dfrac{67.17}{336} = 0.200$
  - the share of the exceedances that lie beyond VaR 1%
  - $0.200^{-\hat{\xi}} = 1.353$
- Step 2: $\widehat{\mathrm{VaR}}_{1\%} = u + \dfrac{\hat{\beta}}{\hat{\xi}}\,(1.353 - 1) = 1.852 + 4.779 \times 0.353 \approx 3.54\%$
- Step 3: $\widehat{\mathrm{ES}}_{1\%} = \dfrac{3.540}{0.812} + \dfrac{0.898 - 0.188 \times 1.852}{0.812} = \dfrac{3.540 + 0.550}{0.812} \approx 5.04\%$
- Historical values for comparison: 3.45% and 5.03%
  - all five series: Appendix

MFM_ch7_gpd_tail

Clustered Extremes: the Extremal Index

- Stationary losses with clustering: if $nP(L > u_n) \to \tau$, then $P(M_n \leq u_n) \to e^{-\theta\tau}$, $\theta \in (0, 1]$ Embrechts, Klüppelberg & Mikosch (1997)
  - u_n: a threshold rising with n; τ: the expected number of exceedances of u_n in n days; M_n: the maximum of the n losses
  - θ = extremal index, $1/\theta$ = mean cluster size; for the maximum, n dependent days act like $n\theta$ independent ones
- S&P 500, GPD threshold ($N = 336$ exceedances): intervals estimator Ferro & Segers (2003) gives $\hat{\theta} = 0.19$
  - i.e. 62 clusters of 5.4 exceedances on average
  - GPD on cluster maxima: $\hat{\xi} = 0.152$ (SE 0.146) vs 0.188 (SE 0.065) on all exceedances: still positive, but its 95% interval now includes 0
- Standardised GARCH residuals: $\hat{\theta} = 0.99$
  - after filtering, threshold exceedances no longer cluster (estimator and declustering: Appendix)
  - $\hat{\theta} \approx 1$ concerns the clustering of extremes, not independence of the whole series
- This is why conditional EVT fits the GPD after the filter

Filtered Monte Carlo over h Days

- Simulate many future paths from the fitted model
  - read VaR and ES from the simulated losses
  - handles autocorrelation, volatility clustering and heavy tails at once, as well as non-linear pay-offs (options)
- For each path $j = 1, \ldots, M$ (M: the number of simulated paths)
  - draw $z_{t+1}^{(j)}, \ldots, z_{t+h}^{(j)}$ with replacement from the standardised residuals $\hat{z}$
  - update $\sigma^{(j)}$ and $r^{(j)}$ day by day with the AR(1)-GARCH(1,1) recursion, starting from today's state
  - aggregate: $L^{(j)} = -\sum_{i=1}^{h} r_{t+i}^{(j)}$, the h-day loss of path j (i: day within the path)
- $\widehat{\mathrm{VaR}}_\alpha$, $\widehat{\mathrm{ES}}_\alpha$ = historical estimates on the M simulated losses
  - simulation error falls as $1/\sqrt{M}$

Minimum-ES Portfolio: a Linear Program

- With N historical scenarios R_s, the Rockafellar–Uryasev objective becomes a **linear program** Rockafellar & Uryasev (2000)
  - $\min_{w,v,u}\ v + \frac{1}{\alpha N}\sum_s u_s$ subject to $u_s \geq -w^\top R_s - v$, $u_s \geq 0$, $\sum_i w_i = 1$, $w \geq 0$
  - $s = 1, \ldots, N$: the scenarios; u_s: an auxiliary variable equal, at the optimum, to the excess $(-w^\top R_s - v)_+$ of scenario s over the threshold v; the optimal v is the portfolio VaR
  - Constraints: fully invested ($\sum_i w_i = 1$), no short sales ($w \geq 0$)
- Same four assets, 3 017 common days, $\alpha = 2.5\%$: weights 31% / 44% / 25% / 0%
  - $\mathrm{ES}_{2.5\%} = 1.67\%$, against 2.16% for the 40/30/20/10 portfolio
- Long-only minimum variance: 35% / 43% / 22% / 0%, $\mathrm{ES}_{2.5\%} = 1.68\%$
  - for these assets the two criteria almost agree
  - they part ways for options and credit
- Several portfolios can attain the same minimum
  - report the solution together with its tail scenarios

FRTB Terms: Desks and Horizons

- The question: how does FRTB turn a daily ES into capital?
  - capital that accounts for crises and for the time needed to exit
- **MAR**: the market-risk chapter of the consolidated Basel Framework
  - MAR12, MAR33 and so on are its chapters (MAR33: capital requirements under the internal models approach)
- **Trading desk**: a defined group of traders and trading accounts with one business strategy
- Under FRTB internal models are approved desk by desk (MAR12)
- **Internal-models approach**: capital from the bank's own ES model
  - the model is supervisor-approved and replaces the standardised formulas
- **Liquidity horizon** (LH): the time needed to exit or hedge a position
  - in stressed markets, without moving its price
- Each risk factor gets 10, 20, 40, 60 or 120 days (MAR33.12)

FRTB ES: Liquidity Horizons and the Stress Period

- ES over liquidity horizons, with $T = 10$ days (MAR33):
  $$\mathrm{ES} = \sqrt{\mathrm{ES}_T(P)^2 + \sum_{j\geq 2} \left(\mathrm{ES}_T(P,j)\sqrt{(\mathrm{LH}_j - \mathrm{LH}_{j-1})/T}\right)^2}$$
  - $\mathrm{ES}_T(P,j)$: the 10-day ES when only the risk factors with $\mathrm{LH} \geq \mathrm{LH}_j$ move, the others held constant; $\mathrm{LH}_1 = 10 < \mathrm{LH}_2 = 20 < \cdots < \mathrm{LH}_5 = 120$ days
- **Stressed period**: the 12-month window that gives the current portfolio its largest ES
  - searched back to at least 2007 (MAR33.6–33.7)
  - Calibration: $\mathrm{ES} = \mathrm{ES}_{R,S}\,\mathrm{ES}_{F,C}/\mathrm{ES}_{R,C}$; R/F: the reduced/full set of factors, S: the stressed period, C: the last 12 months
- **Multiplier**: capital is the larger of the latest figure and $m_c\times$ the 60-day average
  - $m_c = 1.5$ plus 0 to 0.5 after poor backtests (MAR33.41–33.42)
- **Teaching simplification** of Seminar 7, C2: $m_c \times \mathrm{ES}^{(10)}_{2.5\%}\sqrt{\mathrm{LH}/10}$
  - it is the formula above when all factors share one LH

Case Study: the Conditional Quantile as a Neural Network

- Model (Eqs. 5–10): the conditional return quantile is a neural network of the predictors
  - $Q_{r_t}(\alpha \mid x_t) \approx G_\alpha(x_t, w)$
  - $Q_{r_t}(\alpha \mid x_t)$: the α-quantile of r_t given x_t; G_α: a feed-forward network with ReLU (rectified linear unit, $\max(\cdot, 0)$) hidden layers; w: its weights
- Estimation: the fitted weights w^* minimise the pinball loss plus a penalty on w
  - penalty: none, LASSO = L_1, ridge = L_2, or elastic net = both
- The architecture is not fixed in advance: it is chosen by time-series cross-validation
  - depth and width (from $\{1, 5, 10\}$), the learning rate and the penalty (Appendix)
- Target: the daily return quantile 10 days ahead (Eq. 23)
  - $\widehat{\mathrm{VaR}}_{t+10|\mathcal{F}_t} = -G_\alpha(x_t, w^*)$; $\mathcal{F}_t$: the information available at day t
- Evaluation: RMSFE relative to the benchmark
  - RMSFE: root mean squared forecast error
  - below 1 = smaller error than linear quantile regression
  - Significance: Diebold–Mariano (DM) test of equal forecast accuracy

Data Used in This Chapter

Series	Type	Period	Days
S&P 500	US equity index	2000–2026	6 717
BET	BVB equity index	2000–2026	6 670
Bitcoin	crypto-asset, 7 days a week	2014–2026	4 384
EUR/RON	BNR reference rate	2005–2026	5 352
Gold	USD per troy ounce, weekdays	2000–2026	6 956
Portfolio	SPY, TLT, GLD (ETFs) and Bitcoin	2014–2026	3 017
Banca Transilvania	BVB share (liquidity example)	2024–2026	496

- Daily closing values up to 18 September 2026
  - daily log returns in %, $r_t = 100\,(\ln P_t - \ln P_{t-1})$
- Each series on its own calendar
  - the portfolio uses the days on which all four prices exist
- ETFs: prices adjusted for dividends
  - SPY tracks the S&P 500, TLT long US Treasury bonds, GLD gold

Back

Answers to the Quick Checks

- **Which position is riskier?**
  - A: $\mathrm{VaR}_{10\%} = q_{0.9}(L_A) = 9\%$; $\mathrm{ES}_{10\%}$ = mean of L_A on $[9, 10] = 9.5\%$
  - B: $P(L_B > 0) = 5\% \leq 10\%$, so $\mathrm{VaR}_{10\%} = 0$; $\mathrm{ES}_{10\%} = (20 \times 0.05 + 0 \times 0.05)/0.10 = 10\%$
  - VaR calls A riskier (9 vs 0), ES calls B riskier (10 vs 9.5): VaR never sees the 5% chance of losing 20
- **How many days make a precise VaR?**
  - (a) SE $\approx \sqrt{0.05 \times 0.95/1\,000}\,/\,0.1031 = 0.0069/0.1031 \approx 0.067$ pp; 95% half-width $1.96 \times 0.067 \approx 0.13$ pp
  - (b) SE falls like $1/\sqrt{n}$: $4 \times 1\,000 = 4\,000$ days, about 16 years of trading days
  - (c) $n\alpha = 10$ losses: a single extreme day moves the average a lot; precision is governed by $n\alpha$, not n
- **From one day to ten days**
  - (a) $2 \times \sqrt{10} = 6.32\%$; (b) $6.32 \times \sqrt{1.21} = 6.32 \times 1.1 \approx 6.96\%$
  - (c) Probably too low: volatility is expected to rise towards its long-run level during the ten days, so the ten-day variance rises
  - For VaR, the direction is confirmed by filtered Monte Carlo, not by the variance alone

Back

Proof of the Rockafellar–Uryasev Representation

- **Statement**: for continuous losses, $\mathrm{ES}_\alpha(L) = \min_v F(v)$ Rockafellar & Uryasev (2000)
  - $F(v) = v + \mathbb{E}[(L - v)_+]/\alpha$
  - the minimum is attained at $v = \mathrm{VaR}_\alpha$
- **Proof**: F is convex, with right derivative $1 - P(L > v)/\alpha$, which is ≥ 0 exactly when $v \geq \mathrm{VaR}_\alpha$
- Value of the minimum: $F(\mathrm{VaR}_\alpha) = \mathrm{VaR}_\alpha + \mathbb{E}[(L - \mathrm{VaR}_\alpha)_+]/\alpha = \mathbb{E}[L \mid L \geq \mathrm{VaR}_\alpha] = \mathrm{ES}_\alpha$
- **Corollary**: $(w, v) \mapsto v + \mathbb{E}[(-w^\top R - v)_+]/\alpha$ is jointly convex
  - minimising over v keeps convexity
  - ES is convex in w; with positive homogeneity this proves subadditivity

Back

When Is VaR Subadditive After All?

- **Jointly elliptical** losses, $\alpha < 0.5$: VaR is subadditive (slide "Coherent Risk Measures (2/2)") McNeil, Frey & Embrechts (2015)
  - Normal or t marginals alone are not enough: the joint distribution must be elliptical
- **Heavy but not too heavy tails**: VaR is subadditive far enough in the tail Daníelsson et al. (2013)
  - for multivariate regularly varying losses with tail index above 1
  - asymptotically, as $\alpha \to 0$
  - Estimated tail indices of equity and FX (Foreign eXchange) returns are in that range (between 4.8 and 6.5 in the Extreme Value Theory section), but the result is asymptotic, not a guarantee at a given α
- It fails for discrete jumps (defaults), extremely heavy tails and some option pay-offs
- VaR vs ES in practice: ES needs more data for the same precision Yamai & Yoshiba (2005)

Back

VaR Aggregation under Dependence Uncertainty

- Margins $F_1, \dots, F_d$ known, copula unknown: $\underline{\mathrm{VaR}}_\alpha \le \mathrm{VaR}_\alpha(\sum_i L_i) \le \overline{\mathrm{VaR}}_\alpha$ Embrechts, Puccetti & Rüschendorf (2013)
  - the bounds are the minimum and maximum over all copulas (Chapter 6)
  - F_i: distribution function of the loss L_i of asset i, d: number of assets; copula: the dependence structure that joins the margins
  - The comonotone value $\sum_i \mathrm{VaR}_\alpha(L_i)$ (all losses moving perfectly together) is **not** the upper end: VaR is not subadditive

Rearrangement Algorithm and Model Risk

- **Rearrangement algorithm**: discretise the tail beyond each $\mathrm{VaR}_\alpha(L_i)$ in N points (one column per asset)
  - Order each column opposite to the sum of the others until the row sums are flat; worst VaR $\approx$ smallest row sum
- 40/30/20/10 portfolio of the Portfolio VaR section, daily $\mathrm{VaR}_{1\%}$, empirical margins, $N = 2\,000$
  - Historical dependence (empirical quantile of the portfolio loss) 1.95%; comonotone (sum of the empirical stand-alone VaRs) 3.63%; worst case 4.76% (1.31$\times$ comonotone)
  - The Portfolio VaR section uses the Normal formula, hence its 1.77% and undiversified 3.11% for the same portfolio
- Capital based on VaR carries dependence model risk
  - of the same order as the diversification benefit

VaR and ES by Method and Asset

	$\mathrm{VaR}_{1\%}$ (%)					$\mathrm{ES}_{2.5\%}$ (%)				
	Hist.	Normal	t	CF	EVT	Hist.	Normal	t	CF	EVT
S&P 500	3.45	2.80	3.46	6.08	3.54	3.76	2.81	3.95	6.65	3.78
BET	4.13	3.19	4.06	7.25	4.22	4.57	3.21	4.76	7.93	4.58
Bitcoin	10.51	7.99	11.06	18.89	10.23	10.86	8.03	13.37	20.69	10.84
EUR/RON	0.76	0.64	1.36	1.43	0.78	0.84	0.65	3.27	1.60	0.84
Gold	3.04	2.50	3.00	4.64	3.07	3.23	2.51	3.27	5.02	3.23

- Daily losses, full samples; CF = Cornish–Fisher
  - EVT = generalised Pareto tail fitted to the largest 5% of losses (the Extreme Value Theory section)
- Excess kurtosis K: S&P 500 10.7, BET 11.1, Bitcoin 12.0, EUR/RON 15.1, gold 7.7
  - Bitcoin needs about three times the S&P 500 buffer, EUR/RON about a fifth

Back

Asymptotic Theory for Historical VaR and ES

- **VaR**: for i.i.d. losses with density f continuous and positive at VaR_α,
  $\sqrt{n}\,(\widehat{\mathrm{VaR}}_\alpha - \mathrm{VaR}_\alpha) \to N\left(0, \dfrac{\alpha(1-\alpha)}{f(\mathrm{VaR}_\alpha)^2}\right)$
  - $\to$: convergence in distribution as $n \to \infty$; the variance of the limit is the asymptotic variance (avar), and SE $= \sqrt{\mathrm{avar}/n}$
- **ES**: the historical estimator is asymptotically Normal Chen (2008)
  - its influence function IF is the one of the slide "Two Usable Standard Errors (2/2)"
  - No term from estimating VaR: $\partial F/\partial v = 0$ at $v = \mathrm{VaR}_\alpha$; if $\mathbb{E}[(L - \mathrm{VaR}_\alpha)_+^2] < \infty$, the variance of the limit of $\sqrt{n}(\widehat{\mathrm{ES}}_\alpha - \mathrm{ES}_\alpha)$ is avar $= \mathrm{Var}((L - \mathrm{VaR}_\alpha)_+)/\alpha^2$
  - Pareto-type tail with index $a > 2$, $\alpha \to 0$: $\mathrm{Var}((L - \mathrm{VaR}_\alpha)_+) \approx \frac{2\alpha\mathrm{VaR}_\alpha^2}{(a-1)(a-2)}$, so SE $\propto \mathrm{VaR}_\alpha/\sqrt{n\alpha}$ ($\propto$: proportional to) Pele & Mazurencu-Marinescu-Pele (2026)

Back

Precision on the S&P 500, Complete Table

	n	$\widehat{\mathrm{VaR}}_{1\%}$	SE i.i.d.	SE HAC	$\widehat{\mathrm{ES}}_{2.5\%}$	SE i.i.d.	SE HAC
2000–2026	6 717	3.45	0.13	0.19	3.76	0.15	0.26
Last 500 days	500	2.73	0.36	0.44	2.98	0.48	0.65

- Density $f(\mathrm{VaR})$ by a Gaussian kernel, bandwidth $\hat{\sigma}\,(4/(3n))^{1/5}$
  - HAC: Bartlett kernel, lag 10 (full sample) and 5 (last 500 days) Newey & West (1994)
- Full sample, $n = 6\,717$: SE of $\mathrm{VaR}_{1\%}$ with $\sigma = 1.2\%$
  - under the Normal distribution it would be 0.055, under t_4 0.12
  - volatility clustering inflates the SE by 46–73% on the full sample and by 22–35% on the last 500 days

Back

GARCH Filters and the Next-Day Volatility Forecast

	$\hat{\phi}$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_1 + \hat{\beta}_1$	$\hat{\sigma}$ (sample)	$\hat{\sigma}_{t+1}$	$-q_{1\%}(\hat{z})$
S&P 500	−0.053	0.1204	0.8604	0.981	1.21	0.73	2.79
BET	0.104	0.207	0.787	0.994	1.40	1.76	2.74
Bitcoin	−0.009	0.127	0.843	0.970	3.49	3.00	2.88
EUR/RON	0.089	0.148	0.852	1.000	0.28	0.08	2.80
Gold	−0.022	0.046	0.943	0.989	1.09	1.43	2.68

- $\hat{\sigma}_{t+1}$: forecast for the first trading day after 18 September 2026 (19 September for Bitcoin), in %
- EUR/RON: $\hat{\alpha}_1 + \hat{\beta}_1 = 1$, the integrated GARCH boundary
  - within numerical precision
  - so there is no finite unconditional variance and the forecast does not revert to a long-run level
- $\hat{\sigma}$ (sample) is the empirical standard deviation of returns
  - it is not the model-implied long-run volatility
- $-q_{1\%}(\hat{z})$ between 2.68 and 2.88, above $-z_{1\%} = 2.326$: the filtered shocks are still heavy-tailed

Back

Conditional VaR and ES for the Next Trading Day

	$\mathrm{VaR}_{1\%}$ (%)				$\mathrm{ES}_{1\%}$ (%)			
	GARCH	FHS	c-EVT	Hist.	GARCH	FHS	c-EVT	Hist.
S&P 500	1.63	1.97	2.03	3.45	1.88	2.56	2.56	5.03
BET	3.80	4.53	4.63	4.13	4.39	6.17	6.20	6.29
Bitcoin	6.89	8.56	8.68	10.51	7.90	12.23	12.19	14.30
EUR/RON	0.18	0.22	0.21	0.76	0.21	0.28	0.28	1.13
Gold	3.32	3.82	3.86	3.04	3.80	4.95	4.93	4.22

- GARCH: GARCH with Normal shocks; Hist.: unconditional, full sample
  - c-EVT: conditional extreme value theory (the Extreme Value Theory section)
- GARCH with Normal shocks is the lowest of the three conditional methods in every row
  - volatility alone does not create the tail

Back

Estimation Risk in Conditional VaR

- $\mathrm{VaR}_{\alpha,t+\mathbf{1}} = -\mu_{t+\mathbf{1}}(\theta) + \sigma_{t+\mathbf{1}}(\theta)\,\xi_\alpha$, $\xi_\alpha = -q_\alpha(z)$
  - **both** θ (GARCH, by QML) and ξ_α (a residual quantile) are estimated
  - θ: the AR(1)-GARCH(1,1) parameters; QML: quasi-maximum likelihood; $\mu_{t+\mathbf{1}}(\theta)$, $\sigma_{t+\mathbf{1}}(\theta)$: the forecasts they imply
- Francq & Zakoïan (2015): estimate the **risk parameter** (θ, ξ_α) jointly
  - conditions: strict stationarity, $\mathbb{E}[z_t^{\mathbf{4}}] < \infty$ and a positive density of z at $q_\alpha(z)$
  - then the conditional VaR is asymptotically Normal, with a variance from $\hat{\theta}$, from $\hat{\xi}_\alpha$ and from their covariance
- **Residual bootstrap** Christoffersen & Gonçalves (2005), for $b = 1, \ldots, B$
  - B: number of bootstrap samples; a star marks a value computed on a bootstrap sample
  - simulate an AR(1)-GARCH(1,1) path of the sample length with resampled $\hat{z}$ and re-estimate θ^* on it
  - $\sigma^*_{t+\mathbf{1}}$ from the **original** returns with θ^*; ξ^*_α from the path's own residuals; $\mathrm{VaR}^*_{t+\mathbf{1}} = -\mu^*_{t+\mathbf{1}} + \sigma^*_{t+\mathbf{1}}\xi^*_\alpha$; 90% percentile interval: the 5% and 95% quantiles of the B values $\mathrm{VaR}^*_{t+\mathbf{1}}$
- Heavy tails and skewness of z slow the convergence
  - Spierdijk (2016) compares interval methods for ARMA–GARCH VaR in exactly this case
  - ARMA: autoregressive moving average conditional mean

Back

CAViaR Estimation Settings and Joint (VaR, ES) Models

- Estimation settings of Engle & Manganelli (2004): 10^4 random starts in $(0,1)^3$
  - the best 10 refined by simplex and quasi-Newton steps
  - the recursion starts at $\mathrm{VaR}_1 = -\hat{q}_\alpha$, minus the empirical α-quantile of the first 300 returns
  - these days stay in the estimation sample
- The pair (VaR, ES) can be **estimated** by minimising a strictly consistent joint score Fissler & Ziegel (2016)
  - Chapter 8 uses the same scores to **evaluate** forecasts
  - Patton, Ziegel & Chen (2019): dynamic semiparametric (VaR, ES) models, e.g. with GAS (generalised autoregressive score) dynamics, estimated by minimising the FZ0 loss; consistency and asymptotic normality
  - Taylor (2019): CAViaR-type recursions for VaR and ES, estimated by an asymmetric Laplace likelihood
  - Dimitriadis & Bayer (2019): linear regression of the quantile and of ES on covariates (realised and implied volatility), with standard errors
- Common thread: the scoring function that ranks forecasts is also the objective function that estimates the model

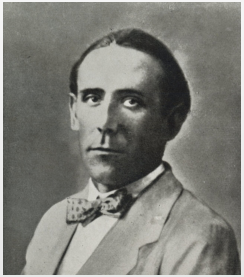
Back

Notation in Extreme Value Theory

- ξ: the GEV/GPD shape
  - $\hat{\xi}_\alpha$ (with a level): the FHS residual quantile $-q_\alpha(\hat{z})$
  - conditional EVT estimates it from a GPD and writes $\hat{z}_\alpha^{GPD}$
- β: GPD scale
  - β_1: GARCH coefficient; $\beta_0, \beta_1, \beta_2$: CAViaR coefficients
- μ, σ: GEV location and scale in the block-maxima section
  - elsewhere the conditional mean and volatility
- u: the GPD threshold
  - in the ES definition, u is the integration variable
  - u_s: the auxiliary variable of the minimum-ES linear program
- α: tail probability
  - α_1: GARCH coefficient; $a = 1/\xi$: tail index

Back

Emil Gumbel and the Statistics of Extremes



Emil Julius Gumbel (1891–1966)

Photo: unknown author (c. 1931); public domain; Wikimedia Commons

- Fisher & Tippett (1928): the maximum of many i.i.d. variables, suitably rescaled, can have only three limit forms
- Gnedenko (1943) proved the result rigorously
- Gumbel (1958) wrote the first textbook and applied it to floods, droughts and life lengths
  - The Gumbel distribution is one of the three limit forms

Back

From Monthly Maxima to a Daily VaR

- Independent days: $P(M_n \leq x) \approx P(L \leq x)^n$, so $\mathrm{VaR}_\alpha \approx H^{-1}((1-\alpha)^n)$
  - $n = 20.9$ trading days per month; H^{-1}: the quantile function of the fitted GEV
  - $\mathrm{VaR}_{1\%} \approx H^{-1}(0.810) = 2.83\%$ vs historical 3.45%; $\mathrm{VaR}_{0.1\%} \approx 6.19\%$
- Volatility clustering breaks the independence assumption, and the approximation is too low here
- With the extremal index: $P(M_n \leq x) \approx F(x)^{n\theta}$, F: the distribution function of the daily loss
  - $\hat{\theta} = 0.35$ (threshold at VaR 1%): $\mathrm{VaR}_{1\%} \approx H^{-1}(0.929) = 4.14\%$ instead of 2.83%; with $\hat{\theta} = 0.19$: 5.09%
  - The correction moves the answer by more than the error it fixes: the block-maxima route to a daily VaR is fragile

Back

GPD Fits for Five Series

	u (%)	N_u	$\hat{\xi}$	SE($\hat{\xi}$)	$\hat{\beta}$	$1/\hat{\xi}$	$\mathrm{VaR}_{1\%}$	$\mathrm{ES}_{1\%}$
S&P 500	1.85	336	0.19	0.065	0.898	5.3	3.54	5.04
BET	1.93	334	0.21	0.066	1.196	4.8	4.22	6.34
Bitcoin	5.33	220	0.16	0.078	2.651	6.1	10.23	14.36
EUR/RON	0.38	268	0.18	0.072	0.217	5.5	0.78	1.14
Gold	1.68	348	0.15	0.062	0.759	6.5	3.07	4.22

- $\hat{\xi}$ between 0.15 and 0.21 for all five series: heavy tails everywhere
  - each more than two i.i.d. standard errors above 0
  - clustering makes these SEs too small
- Tail index $1/\hat{\xi}$ between 4.8 and 6.5
  - the fitted upper tails have finite fourth moments
  - but every 95% interval for ξ reaches above 0.25
  - so finite fourth moments cannot be confirmed (compare Chapter 1)

Back

The Extremal Index: Estimator and Declustering

- **Intervals estimator** Ferro & Segers (2003) from the N exceedance times of u and the gaps T_i between them
  - $\hat{\theta} = \min\left\{1, \dfrac{2\left(\sum_i (T_i - 1)\right)^2}{(N-1)\sum_i (T_i - 1)(T_i - 2)}\right\}$ if $\max T_i > 2$ (else with T_i, T_i^2); it compares the squared mean gap with the mean squared gap: short gaps inside clusters push $\hat{\theta}$ below 1
- **Declustering** Ferro & Segers (2003): target $C = \lfloor \hat{\theta} N \rfloor + 1$ clusters, separated by the $C - 1$ largest gaps
  - Ties: if the $(C-1)$-th and C-th largest gaps are equal, C is reduced until they differ, so the realised number of clusters can be below the target
- Fit the GPD to the cluster maxima
- S&P 500: $\hat{\theta} N = 0.18738 \times 336 = 62.96$, so target $C = \lfloor 62.96 \rfloor + 1 = 63$
  - 62 clusters after the tie rule
  - at the historical VaR 1% ($u = 3.45\%$, $N = 68$): $\hat{\theta} = 0.35$

Back

The Portfolio: Volatilities, Correlations and Contributions

	Weight	σ_i (%)	corr. SPY	corr. TLT	corr. GLD
SPY	40%	1.10	1.00		
TLT	30%	0.93	−0.17	1.00	
GLD	20%	1.03	0.07	0.25	1.00
Bitcoin	10%	4.14	0.23	−0.02	0.09

	σ_i (%)	Marginal VaR	C_i	share	C_i^{ES}	share
SPY	1.10	1.68	0.67	38.1%	0.98	45.4%
TLT	0.93	0.70	0.21	11.8%	0.14	6.6%
GLD	1.03	1.08	0.22	12.3%	0.25	11.5%
Bitcoin	4.14	6.68	0.67	37.8%	0.79	36.6%
Total	$\sigma_p = 0.76$		1.77	100%	2.16	100%

- SPY carries more of the historical ES 2.5% than of the Normal VaR 1%
  - 45.4% vs 38.1%
  - its losses concentrate on the portfolio's worst days more than covariances imply
  - calling this tail dependence needs a separate analysis (Chapter 6)

Back

The Basel Committee and the EU Timetable



The tower of the Bank for International Settlements (BIS) in Basel, seat of the Basel Committee on Banking Supervision

Photo: Michal Pleskowicz (2018); CC BY-SA 4.0; Wikimedia Commons

- The Basel Committee on Banking Supervision sets global capital standards
  - national supervisors turn them into law (in the EU, via the Capital Requirements Regulation)
- **1996**: market risk enters the capital rules Basel Committee (1996)
  - banks may use internal VaR models
- **2019**: FRTB replaces VaR by ES in the internal-models approach Basel Committee (2019)
  - FRTB: the Fundamental Review of the Trading Book standard
  - each jurisdiction sets its own application date
- EU: Delegated Regulation (EU) 2025/1496 postponed the FRTB own-funds requirements to 1 January 2027
  - until then banks apply the earlier VaR-based framework

Back

Case Study, Predictors, Sample and Tuning

- **Data** (Chronopoulos, Raftapostolos & Kapetanios (2024), Section 3.1): S&P 500 daily log returns $r_t = \ln(P_t / P_{t-\mathbf{1}})$
  - September 1985 – August 2020, $T = 9\,053$
  - $\alpha = 1\%$, 5% and 10%
  - Four predictor sets x_t (Eqs. 14–17, following Giacomini & Komunjer (2005)): GARCH(1,1) σ_t; RiskMetrics σ_t with decay factor $\lambda = 0.94$; CAViaR SAV $(\mathrm{VaR}_{t-\mathbf{1}}, |r_{t-\mathbf{1}}|)$; CAViaR asymmetric slope $(\mathrm{VaR}_{t-\mathbf{1}}, r^{+}_{t-\mathbf{1}}, r^{-}_{t-\mathbf{1}})$, $r^{+} = \max(r, 0)$, $r^{-} = \max(-r, 0)$
  - The paper writes τ for our α; $\mathrm{VaR}_{\alpha,t} = -Q_{r_t}(\alpha \mid x_t)$, as in the VaR and Expected Shortfall section

Back

Case Study, Evaluation Design, Rival Models and Tests

- **Design** (paper, Section 3.2): the last 2 000 days are out of sample
  - the first 7 053 tune the model once (fixed forecast window)
  - Time-series CV (cross-validation, Eq. 12): gap $\delta = 5$ days around each left-out point, $\kappa = 20$ random test points; depth and width in $\{1, 5, 10\}$, learning rate and the penalty weight λ (not the RiskMetrics λ) five values in $[0.001, 0.01]$, batch size $\{10, 20\}$, dropout up to 20%, Adam optimiser (stochastic gradient descent with adaptive step sizes), early stopping (paper, Section 1.4)
- **Rivals**: linear quantile regression (the benchmark)
  - quantile polynomial and B-splines, linear MIDAS (Eq. 18)
  - a deep MIDAS network
  - **Evaluation** (paper, Section 3.3): RMSFE relative to the linear benchmark; DM test Diebold & Mariano (1995) on quantile losses with the correction of Harvey, Leybourne & Newbold (1997); CQFE (conditional quantile forecast encompassing) test Giacomini & Komunjer (2005): does one forecast already contain the information of the other; smoothing $\eta = 0.005$
  - Motivation: Patton, Ziegel & Chen (2019) extended CAViaR to (VaR, ES) with parametric dynamics; this paper keeps VaR alone but drops linearity

Back

Case Study, Complete Results (Tables 6 and 7) (1/2)

$\alpha = 1\%$	Polyn.	Splines	MIDAS	Deep	LASSO	Ridge	Elnet	Deep MIDAS
GARCH	0.909	0.791	1.083	0.346	0.236	0.177	0.146	0.346
RiskMetrics	0.780	0.570	1.066	0.289	0.141	0.207	0.239	0.246
SAV	1.000	1.511	0.838	0.198	0.024	0.225	0.498	0.850
Asym. slope	1.000	1.040	4.918	0.235	0.096	0.014	0.037	0.601

- RMSFE relative to linear quantile regression, $\alpha = 1\%$ (Table 6)
  - LASSO, Ridge, Elnet = penalised deep networks
  - all 16 deep entries are below 1 and significant at 1% in the DM test
  - Over all levels, the deep networks cut the error by 50%–98%, the deep MIDAS by 11%–84%; the best penalty changes with the predictor set (elastic net for GARCH, LASSO for RiskMetrics and SAV, ridge for the asymmetric slope)
- At $\alpha = 1\%$ the polynomial and spline regressions do not beat the linear benchmark systematically
  - 4 ratios below 1, 2 equal to 1 and 2 above 1 (largest 1.511; up to 1.716 at other α levels in Table 6)
  - linear MIDAS reaches 4.918
  - the results do not isolate the contributions of architecture, penalisation and CV

Back

Case Study, Complete Results (Tables 6 and 7) (2/2)

- CQFE at $\alpha = 1\%$ (Table 7)
  - the non-linear block encompasses its rivals 757 times (42 losses, 751 inconclusive)
  - the linear block 173 times (15, 170)
  - many inconclusive pairs mean that combining forecasts could still add information

Case Study, Replication and Limitations

- **Replicate** on the daily market data of the course
  - the S&P 500 series starts in January 1990, so 1985–1989 is missing (a data limit)
  - Keep the evaluation window of the paper (the last 2 000 returns up to 31 August 2020, from 19 September 2012): 5 726 tuning days instead of 7 053
  - Build the GARCH and SAV predictors, fit linear and deep quantile regressions at $\alpha = 1\%$ and recompute the GARCH row of Table 6 (Seminar 7, C3: minimum version; the full rows in the replication extension)
- **Limits stated by the authors**
  - One fixed estimation window: recursive or rolling re-tuning over 2 000 days is computationally challenging (paper, Section 3.2)
  - Networks overfit easily, hence penalties, dropout, CV and early stopping (paper, Section 1.4); the output is read only through SHAP (Shapley additive explanations) values and partial derivatives (paper, Section 4)
- **Open extensions** (Seminar 7, C3)
  - Does the gain survive rolling re-estimation on 2020–2026 and the BET?
  - Can a network trained on the FZ0 loss of Patton, Ziegel & Chen (2019) forecast (VaR, ES) jointly, as FRTB needs?

Back

The Discovery Loop, Reusable Prompts (1/2)

- Tools are interchangeable
  - an AI assistant: Claude, ChatGPT, Gemini, GitHub Copilot
  - literature engines: Semantic Scholar, Elicit, Google Scholar
- **Literature**
  - *"List papers since 2015 on comparative backtesting of (VaR, ES) with strictly consistent scoring functions: DOI, scoring function, test, sample size and power findings."*
  - Resolve every DOI on Crossref; check the sign convention of the score in the paper
- **Hypotheses**
  - *"For FHS, conditional EVT and GARCH with Normal shocks, state which features of the tail each captures and what each predicts for the FZ0 differences in calm and in turbulent years."*
- **Data and code**
  - *"Write Python that computes the FZ0 loss for daily VaR 1% and ES 1% forecasts on the S&P 500 from the daily market data of the course, with the sign convention VaR $= -q_\alpha$, and a Diebold–Mariano test with HAC errors."*
  - Replicate first: the FHS exceedance rate 1.54% over 5 210 days (the Conditional VaR and ES section)

Back

The Discovery Loop, Reusable Prompts (2/2)

- **Robustness**
  - *"Write a pre-registered plan: subperiods 2006–2015 and 2016–2026, levels 1% and 2.5%, the FZ0 and a second consistent score, HAC lags 5 and 20, a second index."*
  - Size check: simulate from a fitted GARCH-t and score two forecasts with equal expected FZ0; the test must reject in about 5% of simulations
- **Critique**
  - *"Act as a referee: is my ranking driven by 2008 and 2020, by the score's sign convention or by overlapping estimation windows? Give the test that would expose each."*
- **Report**: every model pair, level and period
  - the number of tail days behind each comparison
- **Multiple testing**: models $\times$ levels $\times$ periods $\times$ assets
  - fix the family in advance and adjust the p-values
  - an unresolvable DOI or a misattributed score: drop the claim and log it in `AI_ERRORS.md`
- Fix the plan before the first run
  - any specification chosen after seeing results is exploratory and is reported as such

Back

Quantinar Courses and the Chapter Code

Advanced courses on quantinar.com

- Measuring Statistical Risk
- XFG Advanced Methods in Quantitative Finance (XFG: Quantlet prefix of *Applied Quantitative Finance*, Härdle, Chen & Overbeck)

Code and data

- Quantlets: github.com/danpele/MFM/tree/main/Quantlets/Ch_07

Back

Acronyms Used in This Material (1/2)

- **AI**: Artificial Intelligence
- **AIC**: Akaike Information Criterion
- **AR**: AutoRegressive (model)
- **ARMA**: AutoRegressive Moving Average (model)
- **BET**: Bucharest Exchange Trading (index)
- **BIS**: Bank for International Settlements
- **BNR**: Banca Națională a României — *the National Bank of Romania*
- **BVB**: Bursa de Valori București — *the Bucharest Stock Exchange*
- **cEVT**: conditional Extreme Value Theory (GARCH filter, EVT tail)
- **CF**: Cornish–Fisher (expansion)
- **COVID-19**: COronaVIrus Disease 2019
- **CQFE**: Conditional Quantile Forecast Encompassing (test)
- **CRSP**: Center for Research in Security Prices
- **CV**: Cross-Validation
- **DM**: Diebold–Mariano (test)
- **ES**: Expected Shortfall
- **ETF**: Exchange-Traded Fund
- **EU**: European Union
- **EVT**: Extreme Value Theory
- **FHS**: Filtered Historical Simulation
- **FRTB**: Fundamental Review of the Trading Book
- **FX**: Foreign eXchange
- **FZ0**: Fissler–Ziegel loss, zero-homogeneous version
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GARCH-N**: GARCH with Normal innovations
- **GAS**: Generalised Autoregressive Score (model)
- **GEV**: Generalised Extreme Value (distribution)
- **GPD**: Generalised Pareto Distribution
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)

Acronyms Used in This Material (2/2)

- **HS**: Historical Simulation
- **IF**: Influence Function
- **LASSO**: Least Absolute Shrinkage and Selection Operator
- **MAR**: Market Risk (chapter of the consolidated Basel Framework)
- **MAR12**: Basel Framework, chapter MAR12 (definition of trading desks)
- **MAR33**: Basel Framework, chapter MAR33 (internal models approach: capital requirements calculation)
- **MC**: Monte Carlo (simulation)
- **MIDAS**: MIxed DAta Sampling (regression)
- **MSE**: Mean Squared Error
- **QML**: Quasi-Maximum Likelihood
- **RMSFE**: Root Mean Squared Forecast Error
- **SAV**: Symmetric Absolute Value (CAViaR specification)
- **SE**: Standard Error
- **SHAP**: SHapley Additive exPlanations
- **US**: United States
- **XAU**: gold, ISO 4217 code (one troy ounce)
- **XFG**: Quantlet prefix of Applied Quantitative Finance (Härdle, Chen & Overbeck)
- **CAViaR**: Conditional Autoregressive Value at Risk
- **CVaR**: Conditional Value at Risk (a synonym of ES)
- **DOI**: Digital Object Identifier
- **i.i.d.**: independent and identically distributed
- **LH**: Liquidity Horizon
- **LVaR**: Liquidity-adjusted VaR
- **POT**: Peaks Over Threshold
- **pp**: percentage points
- **ReLU**: Rectified Linear Unit, the activation $\max(x, 0)$
- **RORAC**: Return On Risk-Adjusted Capital
- **USD**: United States Dollar
- **VaR**: Value at Risk

References (1/4)

- Acerbi, C., Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7), 1487–1503.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Balkema, A.A., de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2(5), 792–804.
- Bangia, A., Diebold, F.X., Schuermann, T., Stroughair, J.D. (2001). Modeling liquidity risk, with implications for traditional market risk measurement and management. In *Risk Management: The State of the Art*, Springer, 3–13.
- Barone-Adesi, G., Giannopoulos, K., Vosper, L. (1999). VaR without correlations for portfolios of derivative securities. *Journal of Futures Markets*, 19(5), 583–602.
- Basel Committee on Banking Supervision (1996). Amendment to the capital accord to incorporate market risks. Bank for International Settlements.
- Basel Committee on Banking Supervision (2019). Minimum capital requirements for market risk. Bank for International Settlements.
- Bayer, S., Dimitriadis, T. (2022). Regression-based expected shortfall backtesting. *Journal of Financial Econometrics*, 20(3), 437–471.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Chen, S.X. (2008). Nonparametric estimation of expected shortfall. *Journal of Financial Econometrics*, 6(1), 87–107.
- Christoffersen, P., Gonçalves, S. (2005). Estimation risk in financial risk management. *Journal of Risk*, 7(3), 1–28.
- Chronopoulos, I., Raftapostolos, A., Kapetanios, G. (2024). Forecasting Value-at-Risk using deep neural network quantile regression. *Journal of Financial Econometrics*, 22(3), 636–669.
- Čížek, P., Härdle, W.K., Weron, R. (Eds.) (2011). Statistical Tools for Finance and Insurance, 2nd ed. Springer.
- Cont, R., Deguest, R., Scandolo, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6), 593–606.
- Cornish, E.A., Fisher, R.A. (1938). Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statistique*, 5(4), 307–320.

References (2/4)

- Daníelsson, J. (2011). Financial Risk Forecasting. Wiley.
- Daníelsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarma, M., de Vries, C.G. (2013). Fat tails, VaR and subadditivity. *Journal of Econometrics*, 172(2), 283–291.
- Daníelsson, J., Zigrand, J.-P. (2006). On time-scaling of risk and the square-root-of-time rule. *Journal of Banking & Finance*, 30(10), 2701–2713.
- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Dimitriadis, T., Bayer, S. (2019). A joint quantile and expected shortfall regression framework. *Electronic Journal of Statistics*, 13(1), 1823–1871.
- D'Innocenzo, E., Lucas, A., Schwaab, B., Zhang, X. (2024). Modeling extreme events: time-varying extreme tail shape. *Journal of Business & Economic Statistics*, 42(3), 903–917.
- Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). Modelling Extremal Events for Insurance and Finance. Springer.
- Embrechts, P., Puccetti, G., Rüschendorf, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, 37(8), 2750–2764.
- Engle, R.F., Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4), 367–381.
- European Commission (2025). Delegated Regulation (EU) 2025/1496. *Official Journal of the European Union*.
- Ferro, C.A.T., Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B*, 65(2), 545–556.
- Fisher, R.A., Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(2), 180–190.
- Fissler, T., Ziegel, J.F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, 44(4), 1680–1707.
- Francq, C., Zakoïan, J.-M. (2015). Risk-parameter estimation in volatility models. *Journal of Econometrics*, 184(1), 158–173.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- Giacomini, R., Komunjer, I. (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics*, 23(4), 416–431.
- Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44(3), 423–453.

References (3/4)

- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), 746–762.
- Gumbel, E.J. (1958). Statistics of Extremes. Columbia University Press.
- Härdle, W.K., Chen, C.Y.-H., Overbeck, L. (Eds.) (2017). Applied Quantitative Finance, 3rd ed. Springer.
- Harvey, D., Leybourne, S., Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13(2), 281–291.
- Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5), 1163–1174.
- Hull, J., White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(1), 5–19.
- J.P. Morgan/Reuters (1996). RiskMetrics – Technical Document, 4th ed. New York.
- Koenker, R., Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.
- Longin, F.M. (2000). From value at risk to stress testing: the extreme value approach. *Journal of Banking & Finance*, 24(7), 1097–1130.
- McNeil, A.J., Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3–4), 271–300.
- McNeil, A.J., Frey, R., Embrechts, P. (2015). Quantitative Risk Management: Concepts, Techniques and Tools, revised ed. Princeton University Press.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Newey, W.K., West, K.D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61(4), 631–653.
- Nolde, N., Ziegel, J.F. (2017). Elicitability and backtesting: perspectives for banking regulation. *Annals of Applied Statistics*, 11(4), 1833–1874.
- Patton, A.J., Ziegel, J.F., Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and Value-at-Risk). *Journal of Econometrics*, 211(2), 388–413.

References (4/4)

- Pele, D.T., Lazar, E., Dufour, A. (2017). Information entropy and measures of market risk. *Entropy*, 19(5), 226.
- Pele, D.T., Lazar, E., Mazurencu-Marinescu-Pele, M. (2019). Modeling expected shortfall using tail entropy. *Entropy*, 21(12), 1204.
- 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., Mazurencu-Marinescu-Pele, M. (2026). Finite-sample precision limits for expected shortfall forecast comparisons. *Mathematics*, 14(13), 2316.
- 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.
- Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119–131.
- Rockafellar, R.T., Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41.
- Rockafellar, R.T., Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7), 1443–1471.
- Spierdijk, L. (2016). Confidence intervals for ARMA–GARCH value-at-risk: the case of heavy tails and skewness. *Computational Statistics & Data Analysis*, 100, 545–559.
- Tasche, D. (2008). Capital allocation to business units and sub-portfolios: the Euler principle. arXiv:0708.2542.
- Taylor, J.W. (2019). Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. *Journal of Business & Economic Statistics*, 37(1), 121–133.
- Yamai, Y., Yoshiba, T. (2005). Value-at-risk versus expected shortfall: a practical perspective. *Journal of Banking & Finance*, 29(4), 997–1015.