








Modelling Financial Markets

Chapter 4: Portfolio Optimisation



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

Can an Estimated Rule Beat Equal Weights?

> *"The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities."*
>
> — *Markowitz (1952)*

- 16 assets and 60 months of data: 152 parameters to estimate, 16 means and 136 variances and covariances
- 2 000 simulated tangency portfolios: the estimated Sharpe ratio has median 1.56
  - the true Sharpe ratio, computed with the population parameters, has median 0.30
- 96% of them have a true Sharpe ratio below that of equal weights, 1/N (0.54)
- Out-of-sample, on 16 ETFs, before costs: Sharpe 0.76 for 1/N, between 0.20 and 0.36 for the GMV rules
  - ETF: Exchange-Traded Fund
  - GMV: global minimum-variance portfolio

▸ Central chart: Estimated versus Realised Frontiers

Why Does Estimating the Optimum Change the Problem?

Easy optimum, hard estimate

- A **portfolio** is a weight vector $\mathrm{w} = (w_1, \ldots, w_N)^\top$ with $\mathrm{w}^\top 1 = 1$
- The optimum is easy to write with known means and covariances, hard to estimate
  - the optimiser loads on the largest estimation errors
  - 16 assets: 16 means + 136 variances and covariances = 152 parameters, estimated from 60 monthly observations
- Can any estimated rule beat equal weights out of sample, after costs?

What the Chapter Contains

Core path

- Mean–variance: global minimum-variance (GMV) and tangency weights, the frontier, constraints
- Estimation risk: why in-sample Sharpe ratios are too optimistic
- Three repairs of the inputs: covariance shrinkage, constraints, Black–Litterman
- Rules that need only Σ: equal risk contribution (ERC) and Hierarchical Risk Parity (HRP)
- Out-of-sample evaluation: turnover and costs, Sharpe-difference tests, Holm and Benjamini–Hochberg
- The application on the Bucharest Stock Exchange (BVB)

- Detailed examples and advanced topics are in Further Topics, after the Conclusions
  - the footer button leads there; the Back button returns to the starting slide

Deeper study

- Case studies: trading costs, volatility timing
- Research assisted by artificial intelligence (AI) and data snooping (searching the same data repeatedly)
- Appendix: derivations and complete tables
- Textbook: Franke, Härdle & Hafner (2019), Part I

What you should be able to do

- derive GMV, tangency and the in-sample Sharpe bias
- compute ERC weights and net Sharpe ratios
- interpret why estimation error favours 1/N
- interpret a Sharpe test after multiple testing
- Further Topics and Appendix: reading for familiarity

Data, Assets and the 1/N Benchmark

Universe	Assets	Returns	Currency
Sectors	9 US sector ETFs	1999/01 – 2026/07	USD
Multi-asset	8 ETFs: equities (3), Treasuries (2), credit (2), gold	2007/05 – 2026/07	USD
Combined	9 sectors + the 7 multi-asset ETFs other than SPY (S&P 500) (16 ETFs)	2007/05 – 2026/07	USD
Romania	8 BVB blue chips; the BET and BET-TR indices	2014/10 – 2026/08	RON

- ETF (exchange-traded fund): a fund traded on an exchange that tracks an index
- Why three universes
  - sectors are similar assets with high correlations
  - multi-asset holds classes with correlations of opposite sign (equities and Treasuries)
  - the combined universe raises N to 16 and amplifies estimation error
- Sample starts: the launch of the sector ETFs (December 1998) and of HYG (April 2007)
- Monthly simple returns from month-end prices, dividends reinvested
  - BET (Bucharest Exchange Trading), the BVB benchmark, is a price index; BET-TR (total return) reinvests dividends
- Excess returns over the one-month US (United States) Treasury bill (Kenneth French Data Library)
- **1/N**: equal weights $w_i = 1/N$, rebalanced every month, with no estimation at all
  - the benchmark of every comparison

▸ Further Topics: the assets in detail

Notation (1/2)

- **Returns**
  - simple return $R_{i,t} = P_{i,t}/P_{i,t-1} - 1$; $P_{i,t}$: the month-end price of asset i
  - excess return $R^e_{i,t} = R_{i,t} - R_{f,t}$ (in the seminar: r_t and r^e_t); $R_{f,t}$: the risk-free rate
  - log return $\ln(P_{i,t}/P_{i,t-1}) = \ln(1 + R_{i,t})$: adds up over time, not across assets
  - portfolios use simple returns, because $R_{p,t} = \sum_i w_i R_{i,t}$ holds exactly
- **Portfolio** of N assets
  - w: vector of weights; $\boldsymbol{\mu}$: vector of expected (excess) returns; 1: vector of ones
  - $\boldsymbol{\Sigma}$: the $N \times N$ covariance matrix of the returns
  - portfolio mean $\mathrm{w}^\top \boldsymbol{\mu}$, portfolio variance $\mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}$

Notation (2/2)

- **Sharpe ratio** (Sharpe (1966)): $SR_p = \mathbb{E}[R_p^e]/\sigma(R_p^e)$
  - the mean excess return per unit of volatility σ
  - monthly figures annualised with $\times 12$ (mean) and $\times\sqrt{12}$ (volatility, Sharpe)
- **Maximum drawdown**: $\min_t(W_t/\max_{s\le t} W_s - 1)$
  - the largest fall of wealth W_t from a previous peak
- **Out of sample** (OOS): the weights held in month t use only data up to month $t-1$
- **In sample**: the same data serve for estimation and for evaluation
- **Rules of this chapter**
  - MV (mean–variance): tangency portfolio from sample moments; GMV: global minimum variance
  - MV-LO (Mean–Variance tangency portfolio, Long-Only) and GMV-LO (Global Minimum-Variance portfolio, Long-Only): their long-only (LO) versions, with no short sales ($w_i \ge 0$); GMV-LW (Global Minimum-Variance portfolio with Ledoit–Wolf covariance): GMV with Ledoit–Wolf (LW) covariance shrinkage
  - ERC (equal risk contribution) and HRP: risk-based rules; 1/N: equal weights

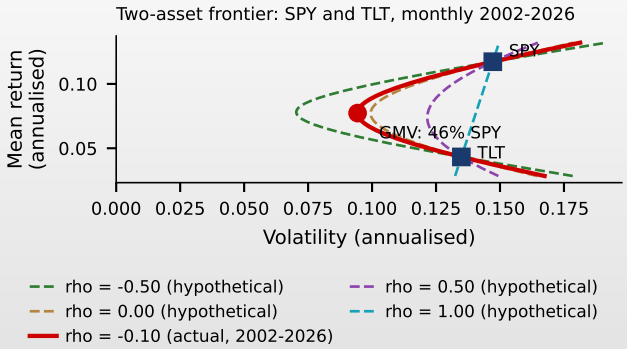
The Markowitz Problem

- The question of this section: which weights do we choose if $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are known?
- Markowitz (1952): investors care about the mean and the variance of portfolio returns
- **Minimum-variance frontier**: minimum variance for every target mean m
  - its upper branch, $m \geq \mu_{\text{GMV}}$ (the mean of the GMV portfolio), is the **efficient frontier**

$$\min_{\mathrm{w}} \ \mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w} \quad \text{s.t.} \quad \mathrm{w}^\top \boldsymbol{\mu} = m, \quad \mathrm{w}^\top 1 = 1$$

  - constraints: the target mean m and full investment (weights sum to 1)
- Utility form, risk aversion $\gamma > 0$: $\max_{\mathrm{w}} \ \mathrm{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}$ s.t. $\mathrm{w}^\top 1 = 1$
  - as γ varies, the solution traces exactly the efficient branch
- Weights may be negative: a **short sale** (selling a borrowed asset)

Two Assets: How Correlation Diversifies



- Two assets, correlation ρ: $\sigma_p^2 = w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho\sigma_1\sigma_2$
  - w: weight of asset 1, σ_1, σ_2: volatilities, σ_p: portfolio volatility
- Minimised at $w^* = \dfrac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$
- SPY (SPDR S&P 500 ETF Trust) and TLT (iShares 20+ Year Treasury Bond ETF), equities and long Treasuries, monthly 2002–2026: $\rho = -0.10$
  - $w^* = 46\%$ SPY, volatility 9.4% (SPY 14.7%, TLT 13.5%)
  - time-varying correlation: Chapter 6
- Chart: means of total returns and annualised volatilities, in decimals ($0.10 = 10\%$)
- The lower ρ, the more the curve bends to the left

MFM_ch4_two_asset ▶ Interactive

The Global Minimum-Variance (GMV) Portfolio

- The problem: $\min_{\mathrm{w}} \mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}$ s.t. $\mathrm{w}^\top 1 = 1$, $\boldsymbol{\Sigma}$ positive definite
- Lagrangian $\mathcal{L} = \mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w} - \lambda(\mathrm{w}^\top 1 - 1)$
  - λ: the Lagrange multiplier of the budget constraint
- First-order condition $2\boldsymbol{\Sigma}\mathrm{w} = \lambda 1$, so $\mathrm{w} = \frac{\lambda}{2}\boldsymbol{\Sigma}^{-1}1$
- The budget gives $\frac{\lambda}{2} = 1/A$, with $A = 1^\top \boldsymbol{\Sigma}^{-1} 1$
- Result: $\mathrm{w}_{\mathrm{GMV}} = \boldsymbol{\Sigma}^{-1}1/A$, variance $\sigma^2_{\mathrm{GMV}} = 1/A$
- Needs only $\boldsymbol{\Sigma}$, not the means

The Tangency Portfolio

- The problem: $\max_{\mathrm{w}} \mathrm{w}^\top \boldsymbol{\mu} / \sqrt{\mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}}$ s.t. $\mathrm{w}^\top 1 = 1$
  - $\boldsymbol{\mu}$: the expected **excess** returns
- First-order condition: $\boldsymbol{\mu} = \dfrac{\mathrm{w}^\top \boldsymbol{\mu}}{\mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}} \boldsymbol{\Sigma} \mathrm{w}$, so $\mathrm{w} \propto \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$
- The budget gives $\mathrm{w}_{\mathrm{MV}} = \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} / B$, with $B = 1^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$
- Maximum Sharpe ratio θ: $\theta^2 = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$
  - needs $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$
- The Sharpe ratio of w_{MV} is $\mathrm{sign}(B)\,\theta$: the maximum only if $B > 0$
- $B < 0$ gives the minimum-Sharpe portfolio
  - for $B = 0$ there is no fully invested solution

Separation and the Frontier in Closed Form



James Tobin, Nobel Prize in Economics 1981

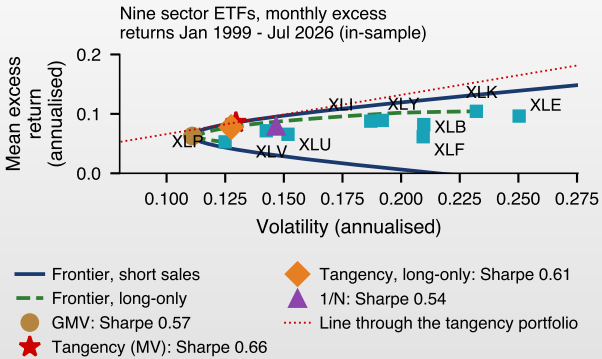
Photo: unknown author (1962); public domain; Wikimedia Commons

- Tobin (1958): add a riskless asset paying R_f
  - every investor holds the **same** risky portfolio, the tangency portfolio, mixed with cash or borrowing
  - risk aversion changes only the mix: **two-fund separation**, the bridge to the Capital Asset Pricing Model (CAPM; Chapter 3)
- Merton (1972): with short sales, the frontier is a hyperbola
  - Σ positive definite, μ not proportional to 1
  - A and B as on the two previous slides; $C = \mu^\top \Sigma^{-1} \mu = \theta^2$, $D = AC - B^2 > 0$

$$\sigma^2(m) = \frac{Am^2 - 2Bm + C}{D}$$

- $\sigma^2(m)$: the smallest variance with mean m
  - a parabola in the (σ^2, m) plane (derivation: ▸ Appendix)
  - the GMV sits at $m = B/A$, $\sigma^2 = 1/A$
  - every frontier portfolio mixes any two distinct frontier portfolios, e.g. GMV and tangency (if $B \neq 0$)

The Frontier of US Sectors



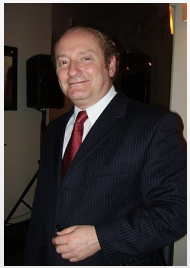
- Monthly excess returns, January 1999 – July 2026 ($T = 331$)
  - axes annualised, in decimals
  - all parameters estimated on the full sample: an **in-sample** picture
- Tangency: Sharpe 0.66; GMV: 0.57 (volatility 11.1%); 1/N: 0.54
- Tangency: short positions (-0.37 financials, -0.21 materials)
  - the largest long position in health care ($+0.37$)
- Long-only tangency: Sharpe 0.61
  - materials, financials and industrials get exactly zero, a **corner solution**
- Dotted line: mixes of the tangency portfolio with cash, from the origin, slope 0.66
- The gain over 1/N is small even with perfect hindsight

MFM_ch4_frontier ▶ Interactive

Constraints

- **Long-only** (LO): $w_i \geq 0$
  - no closed form, solved by quadratic programming
- Position limits: $w_i \leq 0.2$
  - sector or asset-class limits: $\sum_{i \in G} w_i \leq c_G$
  - G a group of assets, c_G its cap
- Leverage limit: $\sum_i |w_i| \leq L$
  - $\sum_i |w_i|$: gross exposure
  - $L = 1$ means no short sales
- Frost & Savarino (1988): constraining weights improved out-of-sample performance
  - the reason: constraints act as shrinkage of the covariance matrix (Jagannathan & Ma (2003); slide Constraints Are Shrinkage Too)

Means Are Hard to Estimate: Merton's Argument

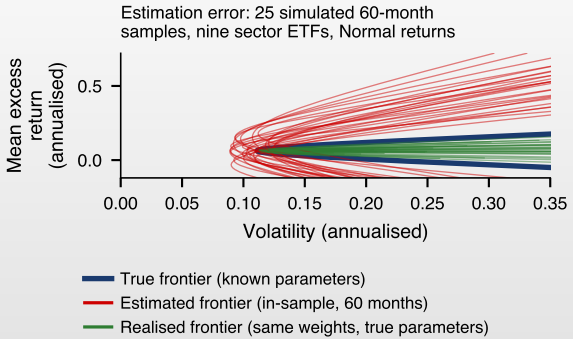


Robert C. Merton, Nobel Prize in Economics 1997

Photo: Digarnick (2006); public domain; Wikimedia Commons

- The question of this section: how much does the optimal portfolio lose when $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are replaced by estimates?
- Merton (1980): Y years sampled h times a year
  - i.i.d. Normal returns (a diffusion) with annual mean μ and variance σ^2
  - $\mathrm{Var}(\hat{\mu}) = \sigma^2/Y$: independent of h; only the **span** of the sample helps
  - $\mathrm{Var}(\hat{\sigma}^2) \approx 2\sigma^4/(hY) \to 0$ as $h \to \infty$ (Normal; in general $(\kappa - 1)\sigma^4/(hY)$, κ the kurtosis)
  - the general formula requires a finite fourth moment, which Chapter 1 questions
- SPY, $Y = 24$, $\sigma = 14.7\%$: standard error 3.0%
  - 95% interval $[5.8\%, 17.6\%]$ for a mean of 11.7%
- A standard error of 1% needs $(0.1472/0.01)^2 \approx 217$ years
- Chopra & Ziemba (1993): errors in means cost about ten times more than errors in variances
  - rules without means (GMV, risk parity) are more robust

Error Maximisation: Estimated versus Realised Frontiers (1/2)



- Michaud (1989): optimisers are **estimation-error maximisers**
  - Best & Grauer (1991): a tiny change in one mean moves many weights
- True parameters = sector moments 1999–2026
  - draw 60 Normal months, optimise with the estimated moments, evaluate with the true parameters
- Chart: 25 simulated 60-month samples
  - red: estimated frontiers, too optimistic
  - green: the same portfolios under the true parameters, all below the true frontier

MFM_ch4_estimation_error ▶ Interactive

Error Maximisation: Estimated versus Realised Frontiers (2/2)

- 2000 simulated tangency portfolios: median estimated maximum Sharpe 1.56
  - median true Sharpe 0.30, true optimum 0.66
- 96% have a true Sharpe below 1/N (0.54), 7.5% below zero
- In 11% of samples $\hat{B} = 1^\top \hat{\Sigma}^{-1} \hat{\mu} < 0$
  - so the tangency formula returns the minimum-Sharpe portfolio

MFM_ch4_estimation_error

The Plug-in Sharpe Ratio Is Biased (1/2)

- i.i.d. $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ returns, T months, N assets (Kan & Zhou (2007))
  - $T\hat{\boldsymbol{\Sigma}}$: Wishart distribution, Appendix
  - $\hat{\theta}^2 = \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}$: the in-sample squared maximum Sharpe ratio, from the maximum-likelihood $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$; θ^2: its true value

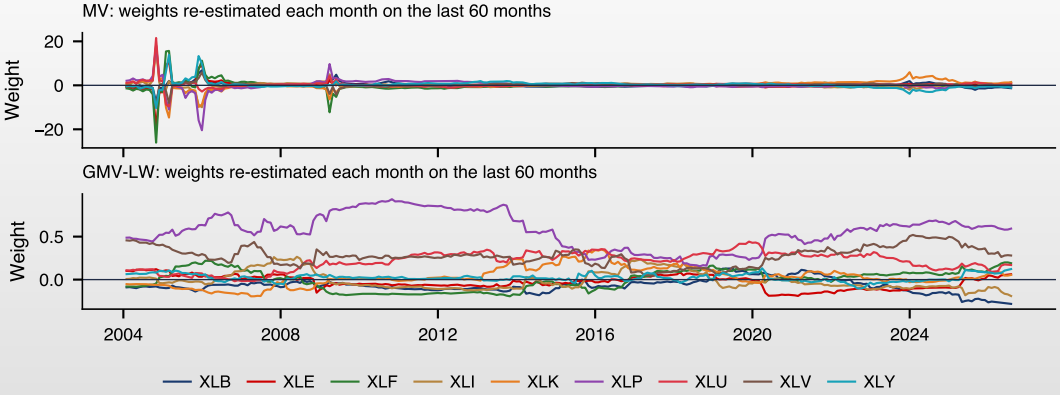
$$\mathbb{E}[\hat{\theta}^2] = \frac{T}{T-N-2}\left(\theta^2 + \frac{N}{T}\right) = \frac{T\theta^2 + N}{T-N-2}, \qquad T > N+2$$

- Sectors: $N = 9$, $T = 60$, true $\theta = 0.66$ a year: $\sqrt{12\,\mathbb{E}[\hat{\theta}^2]} = 1.65$ a year
  - a root mean square, not a median
  - the simulated median 1.56 is lower: $\hat{\theta}$ is right-skewed; derivation: Seminar 4, A8

The Plug-in Sharpe Ratio Is Biased (2/2)

- Unbiased estimator of θ^2: $\hat{\theta}^2_u = \left[(T - N - 2)\,\hat{\theta}^2 - N\right]/T$
  - It can be negative; truncating at zero gives a nonnegative but upward-biased estimator, and $\sqrt{\hat{\theta}^2_u}$ is not unbiased for θ
- Sectors, full sample ($T = 331$), under i.i.d. Normality: the bias correction of $\hat{\theta}^2$ lowers the equivalent point estimate of the Sharpe ratio from 0.66 (plug-in) to about 0.31 a year
  - the correction concerns θ^2: it does not show which part of the observed Sharpe ratio is noise
- Tangency weights: $\mathrm{w} = \mathrm{b}/1^\top \mathrm{b}$ (Britten-Jones (1999))
  - b: the ordinary least squares (OLS) coefficients of a regression of 1 on the excess returns, without a constant
  - derivation: Appendix; Seminar 4, A9 and B6

Unstable Weights (1/2)



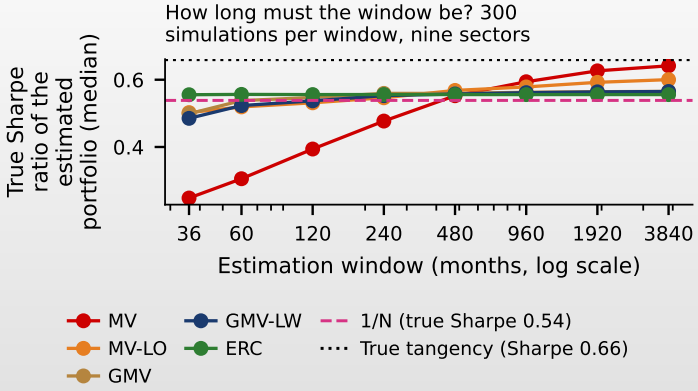
MFM_ch4_oos_backtest ▶ Interactive

Unstable Weights (2/2)

- Nine sector ETFs; a 60-month rolling window
  - one of the two windows of DeMiguel, Garlappi & Uppal (2009) (60 and 120 months)
  - a longer window cuts estimation error, a shorter one tracks regime changes faster
- MV: median gross exposure $\sum_i |w_i| = 4.9$
  - above 3 in 90% of the months
  - largest single weight 26 in absolute value
- Explosions in 2004–2006 and 2009
  - the denominator $1^\top \hat{\Sigma}^{-1} \hat{\mu}$ is close to zero, so the tangency weights have very heavy tails
  - Negative in 12 of 271 windows: MV then holds the minimum-Sharpe portfolio
- GMV-LW (bottom): gross exposure 1.5, largest weight 0.94
- Unstable weights mean high turnover, hence high trading costs

MFM_ch4_oos_backtest

How Long Must the Window Be to Beat 1/N? (1/2)



MFM_ch4_window_length ▶ Interactive

How Long Must the Window Be to Beat 1/N? (2/2)

- DeMiguel, Garlappi & Uppal (2009): 14 models on 7 data sets
  - none beats 1/N consistently in Sharpe ratio, certainty-equivalent return or turnover
  - certainty-equivalent return: mean $-\frac{\gamma}{2}\times$ variance, the riskless return worth the same utility
- In their calibration MV needs a window of about 3000 months for 25 assets
- Chart: 300 simulated samples per window
  - horizontal axis: window, months, log scale
  - population parameters = sector moments
  - vertical axis: the median true Sharpe ratio
- MV beats 1/N only with 480 months (40 years)
  - median true Sharpe 0.55 vs 0.54
  - with 60 months 0.31
- MV-LO, GMV, GMV-LW are close to 1/N from 60 months on
  - ERC is almost insensitive, 0.56
- In this calibration 1/N is nearly efficient (true tangency 0.66)
  - simulation evidence, not an empirical test

MFM_ch4_window_length

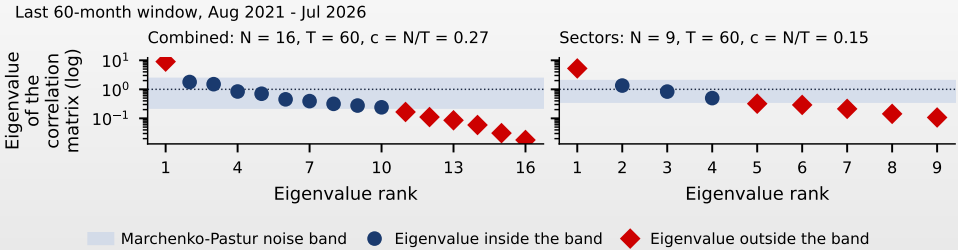
Why the Sample Covariance Matrix Fails

- The question of this section: how do we estimate $\boldsymbol{\Sigma}$ so that $\hat{\boldsymbol{\Sigma}}^{-1}$ does not amplify noise?
- Sample eigenvalues are too dispersed: the largest too large, the smallest too small
- **Condition number**: largest over smallest eigenvalue
  - a large value means $\boldsymbol{\Sigma}^{-1}$ amplifies noise
- Optimisers invert $\boldsymbol{\Sigma}$, so they load on the directions with the smallest estimated variance
- The demeaned sample matrix has rank at most $T - 1$: singular, not invertible, whenever $T \leq N$
- **Marchenko–Pastur** benchmark (Marčenko & Pastur (1967))
  - i.i.d. returns with $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ ($\mathbf{I}$: identity matrix)
  - as $N, T \to \infty$ with $N/T \to c \in (0, 1)$, the sample eigenvalues fill

$$[\lambda_-, \lambda_+] = \left[\sigma^2(1 - \sqrt{c})^2,\ \sigma^2(1 + \sqrt{c})^2\right]$$

  - λ_-, λ_+: the smallest and largest eigenvalue that pure noise produces
  - $c = 16/60 = 0.27$: band $[0.23, 2.30]$ for a correlation matrix ($\sigma^2 = 1$), a ten-fold spread from noise alone
  - Laloux et al. (1999): eigenvalues outside the band are candidates for structure; those inside cannot be told from noise by this comparison alone

Sample Eigenvalues versus the Noise Band (1/2)



- Correlation matrices of monthly excess returns, last 60-month window (August 2021 – July 2026)
  - shaded band: $[\lambda_-, \lambda_+]$ for $\sigma^2 = 1$
- Only the market mode lies above λ_+
  - 9.0 (56% of the trace) for 16 ETFs, 5.2 for nine sectors ($\lambda_+ = 1.92$)
  - with the noise variance rescaled to $\sigma^2 = 1 - \lambda_1/N$ (Laloux et al. (1999)), three eigenvalues exceed the band in each universe

MFM_ch4_inference ▶ Interactive

Sample Eigenvalues versus the Noise Band (2/2)

- Six of 16 eigenvalues lie **below** the band: directions of low standardised in-sample variance
  - such as long TLT and short IEF (iShares 7–10 Year Treasury Bond ETF), scaled by their volatilities
- GMV puts large weights on these combinations
  - if their variance is underestimated, so is the GMV risk

MFM_ch4_inference

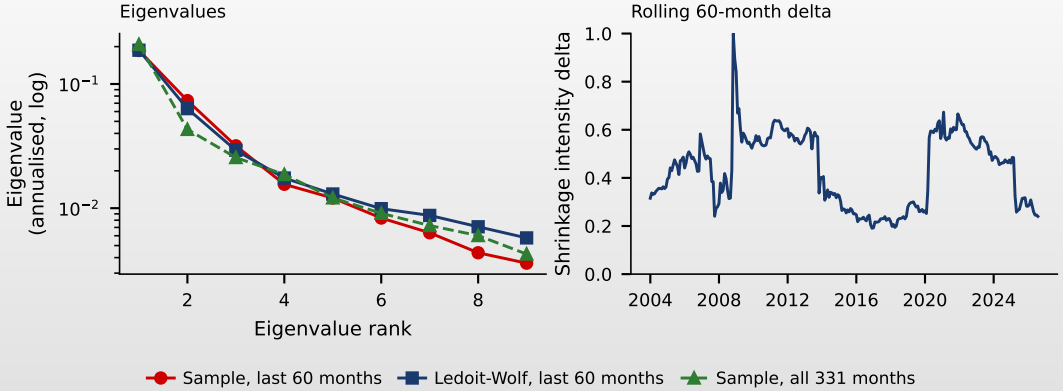
The Shrinkage Estimator

- Shrink the sample matrix S towards a structured target F

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{LW}} = \delta\,\mathsf{F} + (1-\delta)\,\mathsf{S}, \qquad 0 \le \delta \le 1$$

- Possible targets
  - the single-index model (Ledoit & Wolf (2003))
  - the scaled identity (Ledoit & Wolf (2004a))
  - constant correlation (Ledoit & Wolf (2004b))
- The target used here: $f_{ii} = s_{ii}$, $f_{ij} = \bar{r}\sqrt{s_{ii}s_{jj}}$ for $i \neq j$
  - $\bar{r}$: the average sample correlation
- Why this target: the sectors are all US equities with similar correlations
  - the target needs no choice of a market index
- Optimal intensity: $\delta^* = \max\{0, \min\{1, \hat{\kappa}/T\}\}$, $\hat{\kappa} = (\hat{\pi} - \hat{\rho})/\hat{\gamma}$
  - $\hat{\pi} - \hat{\rho}$: the estimated noise of S (Ledoit–Wolf notation; $\hat{\rho}$ is not a correlation)
  - $\hat{\gamma} = \|\mathsf{F} - \mathsf{S}\|^2$: the distance to the target (not the risk aversion γ)
- δ^* falls like $1/T$: in long samples $\delta^* \to 0$ (▸ Appendix)

Shrinkage on US Sectors (1/2)



MFM_ch4_shrinkage ▶ Interactive

Shrinkage on US Sectors (2/2)

- Nine sector ETFs, monthly excess returns
  - left: annualised eigenvalues (log scale)
  - right: δ^* on rolling 60-month windows
- Last window, August 2021 – July 2026: $\delta^* = 0.24$
- The smallest eigenvalues rise
  - the largest is almost unchanged, the second falls from 0.073 to 0.063
- Condition number $51.6 \to 32.5$
  - the matrix on all 331 months (green) has 48.7
- Rolling δ^*: mean 0.43, range 0.19–1.00
- $\delta^* = 1$ only in the window ending in October 2008
  - the crisis month triples the estimated noise $\hat{\pi}$ of S, while the distance $\hat{\gamma}$ to the target barely moves

MFM_ch4_shrinkage

Constraints Are Shrinkage Too

- Jagannathan & Ma (2003): the long-only GMV equals the unconstrained GMV of a modified covariance matrix

$$\tilde{S} = S - (\boldsymbol{\lambda} 1^\top + 1\boldsymbol{\lambda}^\top), \qquad \mathcal{L} = \tfrac{1}{2} w^\top S w - \lambda_0 (w^\top 1 - 1) - \boldsymbol{\lambda}^\top w$$

  - $\boldsymbol{\lambda} \geq 0$: the Lagrange multipliers of the constraints $w_i \geq 0$ in $\mathcal{L}$ (not eigenvalues here); $\lambda_i > 0$ only where the constraint binds
  - The first-order condition $Sw = \lambda_0 1 + \boldsymbol{\lambda}$ and $\boldsymbol{\lambda}^\top w = 0$ give $\tilde{S}w = \lambda_0 1$: the GMV condition for $\tilde{S}$
  - step by step, with the budget $1^\top w = 1$: $\tilde{S}w = Sw - \boldsymbol{\lambda}(1^\top w) - 1(\boldsymbol{\lambda}^\top w) = \lambda_0 1 + \boldsymbol{\lambda} - \boldsymbol{\lambda} - 0$
- $\tilde{s}_{ij} = s_{ij} - \lambda_i - \lambda_j$: a binding constraint on asset i lowers its covariances
  - this shrinks the large estimates
  - why there: the unconstrained GMV shorts assets with high covariances with the others, so $w_i \geq 0$ binds on exactly those assets, whose large covariance estimates are the most likely to be inflated by noise
- Imposing the "wrong" constraint (no short sales) can reduce out-of-sample risk
  - the effect is largest when covariances come from few observations

Constraints as Shrinkage: A Two-Asset Example

- $\sigma_1 = 25\%$, $\sigma_2 = 10\%$, $\rho = 0.5$: $s_{11} = 0.0625$, $s_{22} = 0.0100$, $s_{12} = 0.0125$
- Unconstrained GMV: $w_1^* = (s_{22} - s_{12})/(s_{11} + s_{22} - 2s_{12}) = -0.0025/0.0475 = -0.053$
  - the volatile asset is shorted because it co-moves strongly with the other; volatility 9.93%
- Long-only: $w_1 \geq 0$ binds, w $= (0, 1)$, volatility 10%
  - first-order condition: Sw $= (s_{12}, s_{22}) = (0.0125, 0.0100) = \lambda_0 1 + \boldsymbol{\lambda}$, so $\lambda_0 = 0.0100$, $\boldsymbol{\lambda} = (0.0025, 0)$
- Modified matrix: $\tilde{s}_{12} = 0.0125 - 0.0025 = 0.0100$
  - $\tilde{s}_{11} = 0.0625 - 2(0.0025) = 0.0575$, $\tilde{s}_{22} = 0.0100$
  - correlation lowered from 0.5 to $0.0100/\sqrt{0.0575 \times 0.0100} \approx 0.42$; the variance of asset 1 also falls
  - unconstrained GMV of $\tilde{\mathrm{S}}$: $w_1^* = (0.0100 - 0.0100)/(\cdots) = 0$, the long-only solution

From Covariance Shrinkage to Mean Shrinkage



Fischer Black (1938–1995)

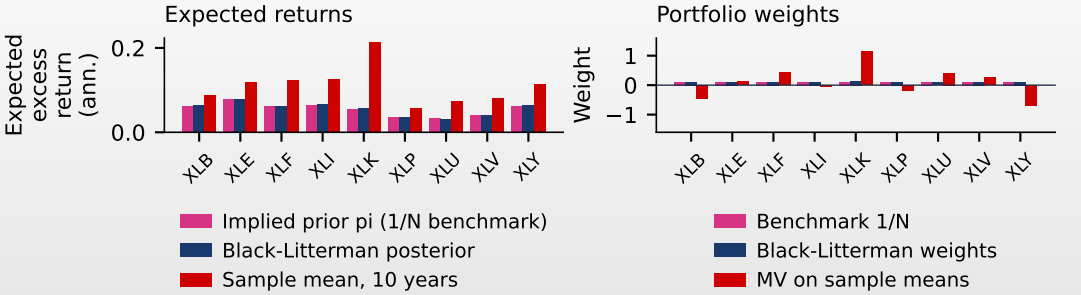
Photo: Dalmatine at English Wikipedia; public domain; Wikimedia Commons

- The question of this section: which means do we use, if the sample means are too imprecise?
- Black & Litterman (1992) shrinks the means towards those implied by a **benchmark** portfolio w_b
- Reverse optimisation: a mean–variance investor with risk aversion $\delta_{\mathbf{RA}}$ and access to cash
  - holds $w^* = (\delta_{\mathbf{RA}} \boldsymbol{\Sigma})^{-1} \boldsymbol{\mu}$ in risky assets
  - the returns that make w_b optimal are

$$\boldsymbol{\pi} = \delta_{\mathbf{RA}} \, \boldsymbol{\Sigma} \, w_b$$

- $\boldsymbol{\pi}$: **implied** (equilibrium) excess returns, not the $\hat{\pi}$ of Ledoit–Wolf
  - with market weights, the CAPM returns
- Here: no market capitalisations for the sectors, so $w_b = 1/N$, $\delta_{\mathbf{RA}} = 2.5$
- Then add the investor's **views**, weighted by their confidence

Black–Litterman on US Sectors (1/2)



MFM_ch4_black_litterman ▶ Interactive

Black–Litterman on US Sectors (2/2)

- Left: annualised means
  - the implied means π
  - the Black–Litterman means after the view
  - the 10-year sample means
- Right: the 1/N, Black–Litterman and sample-mean MV weights
- Σ from monthly excess returns, August 2016 – July 2026
  - 120 months, the other window of DeMiguel, Garlappi & Uppal (2009)
- One relative view: technology (XLK) minus utilities (XLU) = 3% a year (view portfolio: +1 XLK, −1 XLU)
- XLK minus XLU: 2.2% implied, 2.6% after combining with the view (formula on the next slide)
- Weights: XLK from 11.1% to 14.4%, XLU from 11.1% to 7.8%
  - all others unchanged (by hand: Seminar 4, A7)
- With sample means instead (XLK: 21.3% a year)
  - the MV weights become +115% XLK, −68% consumer discretionary (XLY) and −45% materials (XLB)

MFM_ch4_black_litterman

Views and the Posterior

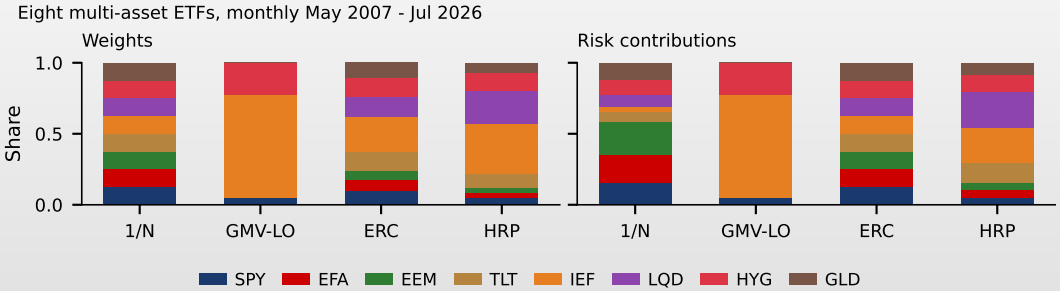
- Prior: $\boldsymbol{\mu} \sim N(\boldsymbol{\pi}, \tau\boldsymbol{\Sigma})$
  - τ scales the uncertainty about $\boldsymbol{\pi}$
- K views: $\mathrm{P}\boldsymbol{\mu} = \mathrm{q} + \boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Omega})$
  - P: $K \times N$ matrix, row k = the portfolio of view k
  - q: the view returns; $\boldsymbol{\Omega}$: their uncertainty
- Posterior mean, a precision-weighted average

$$\boldsymbol{\mu}_{\mathbf{BL}} = \left[(\tau\boldsymbol{\Sigma})^{-\mathbf{1}} + \mathrm{P}^{\top}\boldsymbol{\Omega}^{-\mathbf{1}}\mathrm{P}\right]^{-\mathbf{1}}\left[(\tau\boldsymbol{\Sigma})^{-\mathbf{1}}\boldsymbol{\pi} + \mathrm{P}^{\top}\boldsymbol{\Omega}^{-\mathbf{1}}\mathrm{q}\right]$$

- Black–Litterman weights: $\mathrm{w}_{\mathbf{BL}} = (\delta_{\mathbf{RA}}\boldsymbol{\Sigma})^{-\mathbf{1}}\boldsymbol{\mu}_{\mathbf{BL}}$
- One asset, $\pi = 5\%$, $\tau\sigma^{\mathbf{2}} = 0.00128$, view 8% with variance 0.0004
  - view weight $0.00128/(0.00128 + 0.0004) = 0.76$, so $\mu_{\mathbf{BL}} = 7.29\%$

Weights versus Risk Contributions (1/2)

The question of this section: can we build weights from Σ alone, without means and without inverting Σ?



MFM_ch4_risk_contrib ▶ Interactive

Weights versus Risk Contributions (2/2)

- Eight multi-asset ETFs, full-sample covariance, May 2007 – July 2026 ($T = 231$)
  - volatilities from 6.7% (IEF) to 21.1% (EEM)
- Left: the weights; right: each asset's share of risk
  - share of risk: $w_i(\Sigma \mathrm{w})_i / \sigma_p^2$ (next slide)
  - shares in decimals ($0.5 = 50\%$)
- 1/N: equities and high yield receive half of the weight but 69% of the risk
- ERC (equal risk contribution): 12.5% of the risk from every asset
  - IEF gets the largest weight (25%)
- Long-only GMV: 73% IEF, 22% HYG, 5% SPY, five assets at zero
- HRP (Hierarchical Risk Parity): largest weights IEF (35%) and LQD (23%)
- Volatility: 1/N 9.4%, ERC 7.9%, HRP 7.1%, long-only GMV 5.8%

MFM_ch4_risk_contrib

Risk Contributions

- Volatility is homogeneous of degree one in the weights (Euler)

$$\sigma_p = \sum_{i=1}^{N} w_i \frac{\partial \sigma_p}{\partial w_i} = \sum_{i=1}^{N} \frac{w_i (\Sigma \mathrm{w})_i}{\sigma_p}$$

- **Risk contribution** of asset i: $RC_i = w_i(\Sigma \mathrm{w})_i / \sigma_p$
  - percentage share $RC_i/\sigma_p = w_i(\Sigma \mathrm{w})_i/\sigma_p^2$
- Worked example, 50/50 in SPY and IEF (monthly 2002–2026, annualised)
  - $\sigma_1^2 = 0.02166$, $\sigma_2^2 = 0.00441$, $\sigma_{12} = -0.00098$
  - $(\Sigma \mathrm{w})_1 = 0.5(0.02166) + 0.5(-0.00098) = 0.01034$, $(\Sigma \mathrm{w})_2 = 0.5(-0.00098) + 0.5(0.00441) = 0.00172$
  - $\sigma_p^2 = 0.5(0.01034) + 0.5(0.00172) = 0.00603$, $\sigma_p = 7.76\%$; shares: SPY 85.8%, IEF 14.2%
- Risk budgeting chooses w so that the shares equal preset budgets b_i

Equal Risk Contribution

- ERC (Maillard, Roncalli & Teïletche (2010)): $RC_i = RC_j$ for all i, j
  - computed from a convex problem, then normalised

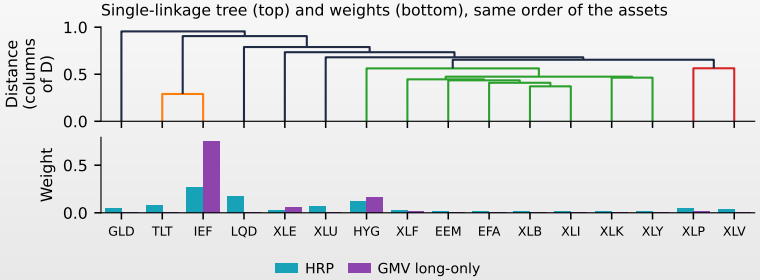
$$\min_{\mathrm{y}>0} \ \tfrac{1}{2}\mathrm{y}^\top \boldsymbol{\Sigma}\mathrm{y} - \tfrac{1}{N}\textstyle\sum_i \ln y_i, \qquad \mathrm{w} = \mathrm{y}/1^\top\mathrm{y}$$

  - y: unnormalised positive weights; the log term keeps every $y_i > 0$ and its first-order condition gives $y_i(\boldsymbol{\Sigma}\mathrm{y})_i = 1/N$ for all i
- For $\boldsymbol{\Sigma}$ positive definite a solution exists and is unique
  - strictly convex objective, infinite at the boundary
- Volatility ordering: $\sigma_{\mathrm{GMV}} \leq \sigma_{\mathrm{ERC}} \leq \sigma_{1/N}$
- Two assets: ERC is **inverse volatility**, $w_1 = \sigma_1^{-1}/(\sigma_1^{-1} + \sigma_2^{-1})$
  - for any $\rho \in (-1, 1)$
  - SPY (14.7%) and IEF (6.6%): ERC holds 31% SPY, volatility 6.1%, each asset 50% of the variance
- ERC is the tangency portfolio if all assets have the same Sharpe ratio and equal correlations

Hierarchical Risk Parity: From a Tree to Weights

- HRP (López de Prado (2016)): three steps, no matrix inversion, all weights positive
  - works even when $\boldsymbol{\Sigma}$ is singular
- **Step 1, tree**
  - correlation distance $d_{ij} = \sqrt{(1-\rho_{ij})/2} \in [0,1]$; ρ_{ij}: correlation of assets i and j
  - Euclidean distance between the **columns** of $\mathrm{D} = (d_{ij})$: $\tilde{d}_{ij} = \sqrt{\sum_k (d_{ki} - d_{kj})^2}$
  - single linkage on $\tilde{d}_{ij}$ (repeatedly merge the two closest clusters) gives a **dendrogram**
- **Step 2, quasi-diagonalisation**
  - order the assets as the leaves of the tree, so that similar ones are adjacent
- **Step 3, recursive bisection**: split the **ordered list** in two halves
  - the halves get α and $1-\alpha$, $\alpha = 1 - V_1/(V_1+V_2)$, $V_j = \tilde{\mathrm{w}}_j^\top \boldsymbol{\Sigma}_j \tilde{\mathrm{w}}_j$ the variance of half j
  - $\tilde{\mathrm{w}}_j$: weights $\tilde{w}_i \propto 1/\sigma_i^2$ inside half j, $\boldsymbol{\Sigma}_j$ its covariance
  - repeat inside each half until single assets remain
- Mantegna (1999): the same distance and the **minimum spanning tree** reveal the hierarchy of a stock market
- Textbook: Härdle & Simar (2019), cluster analysis
- A worked four-asset example: Seminar 4, B4

HRP on 16 ETFs



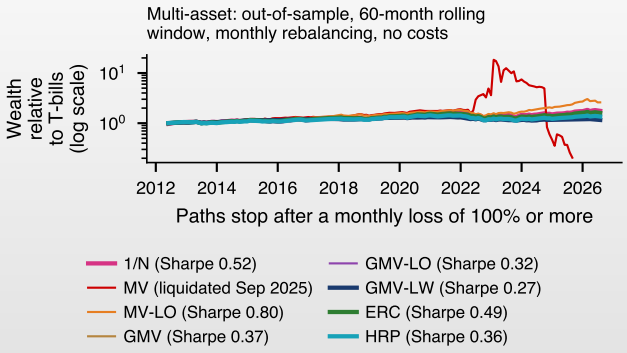
- Nine sectors and seven multi-asset ETFs, monthly, May 2007 – July 2026
  - closest pair TLT and IEF ($\tilde{d} = 0.29$), gold joins last (0.96)
- Equity cluster: cyclical sectors, EFA (iShares MSCI EAFE ETF), EEM and high-yield credit
  - defensive sectors (utilities, staples, health care) and energy join it late
- HRP: all 16 weights positive, largest IEF 26.7%, LQD 17.6%, HYG 12.6%
  - long-only GMV: 75.6% in IEF and only five assets above zero

MFM_ch4_hrp ▶ Interactive

Backtest Design

- The question of this section: does an estimated rule beat 1/N
  - when the weights use only the data available at the decision date?
- The weights of month t use months $t-60, \ldots, t-1$ only (the 60-month rolling window)
- Monthly rebalancing, at the data frequency
  - its cost is measured separately, through turnover
- Eight rules: 1/N, MV, MV-LO, GMV, GMV-LO, GMV-LW, ERC, HRP
- A rule whose wealth reaches zero (a monthly loss of 100% or more) is liquidated
  - its later returns are not an investment outcome
  - it is left out of the Sharpe comparisons, the cost analysis and the tests
- Three universes
  - Sectors: OOS January 2004 – July 2026, 271 months (first window 1999–2003)
  - Multi-asset and combined: OOS May 2012 – July 2026, 171 months
- Measures
  - annualised Sharpe $\sqrt{12}\,\bar{r}^e/\hat{\sigma}(r^e)$ from excess returns
  - maximum drawdown from total returns
  - every difference from 1/N with a test, corrected for multiple testing

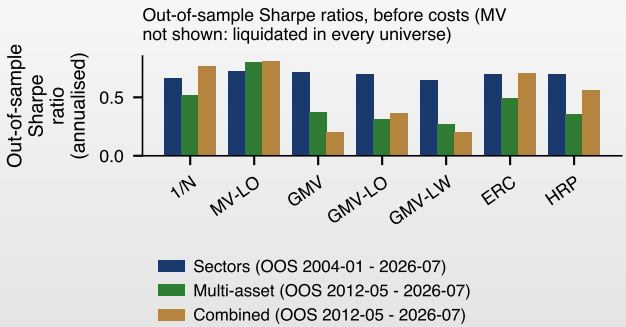
Out-of-Sample Growth: Multi-Asset



- Vertical axis: wealth relative to a Treasury-bill account, log scale
  - eight multi-asset ETFs, before costs
- MV collapses: turnover 44 times wealth per month
  - liquidated in September 2025, after a monthly loss above 100%
- MV-LO has the highest Sharpe (0.80); 1/N 0.52; ERC 0.49
- GMV rules: lowest volatility (3.8%–5.4%), low Sharpe (0.27–0.37)
  - they load on Treasuries, whose excess returns were slightly negative after May 2012 (TLT −0.5%, IEF −0.2% a year)
- MV-LO: better, or lucky? A test decides

MFM_ch4_oos_backtest ▶ Interactive

Sharpe Ratios Across Universes



- Annualised, before costs
  - complete tables for all three universes in the ▸ Appendix
- MV is not shown: its wealth reaches zero in every universe
  - sectors March 2009, multi-asset September 2025, combined January 2023
- Sectors: every other rule lands within 0.07 of 1/N (0.66)
- 1/N has 0.66 here (2004–2026), not 0.54 as in the simulations (1999–2026)
  - the periods differ
- Combined: 1/N 0.76, MV-LO 0.81, ERC 0.71
  - the GMV rules 0.20–0.36
- Turnover per month, sectors: 1/N 0.02, ERC 0.03, GMV-LW 0.14, MV-LO 0.23

MFM_ch4_oos_backtest ▸ Interactive

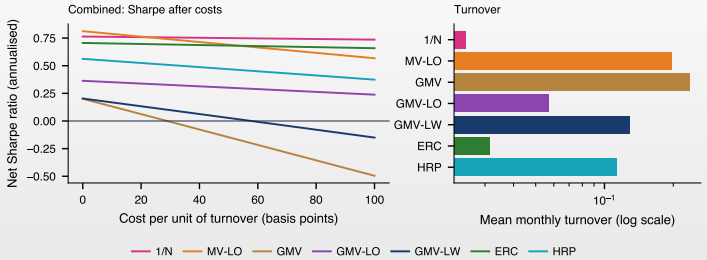
Turnover and Costs: One Worked Month (1/2)

- Between rebalancings weights drift with returns: $w_{i,t}^{+} = w_{i,t}(1 + R_{i,t}) / \sum_j w_{j,t}(1 + R_{j,t})$
  - $R_{i,t}$: **total** return of asset i in month t; the denominator is $1 + R_{p,t}$, one plus the portfolio's total return
- **Turnover**: traded fraction of wealth, $TO_{t+1} = \sum_i |w_{i,t+1} - w_{i,t}^{+}|$
  - summed over all N assets
- Worked month, sectors, 1/N: in January 2004 XLB returned -4.4% and XLK $+4.3\%$
  - so the weights of 11.11% drifted to 10.50% and 11.45%
  - 1/N total return in January 2004: 1.18% (excess 1.11% plus the Treasury bill 0.07%; Seminar 4, B1)
  - XLB: $\frac{1}{9}(1 - 0.044)/1.0118 = 10.50\%$; XLK: $\frac{1}{9}(1 + 0.043)/1.0118 = 11.45\%$

Turnover and Costs: One Worked Month (2/2)

- Rebalancing back to 1/N in February 2004 trades $TO = 0.019$
  - 1/N also trades, selling winners and buying losers
  - XLB and XLK contribute $|0.1111 - 0.1050| + |0.1111 - 0.1145| = 0.0061 + 0.0034 = 0.0095$; the other seven sectors about 0.009 together
  - MV-LO, re-estimated on the new window, trades 0.748 in the same month
- Net return with cost c per unit traded: $R^{\text{net}}_{p,t} = R_{p,t} - c\, TO_t$
  - 1 basis point (bp) $= 0.01\%$, so 20 bp means $c = 0.002$
  - At 20 bp the February trade costs MV-LO $0.002 \times 0.748 = 0.15\%$ of wealth, 1/N $0.002 \times 0.019 = 0.004\%$

Sharpe Ratios after Costs



- Combined universe (16 ETFs)
  - left: net Sharpe (vertical) against the cost c, 0–100 bp per unit of turnover (horizontal)
  - right: mean monthly turnover, log scale, from 0.02 (1/N) to 0.24 (GMV)
  - MV is not shown: liquidated in January 2023
- Net Sharpe turns negative at about 29 bp for GMV and 58 bp for GMV-LW
  - MV-LO overtakes 1/N only for costs below about 22 bp (Seminar 4, B2)
- 1/N and ERC barely move: low turnover is an advantage

MFM_ch4_costs ▶ Interactive

Inference on One Sharpe Ratio: Lo (2002)

- Lo (2002), i.i.d. Normal returns: $\sqrt{T}(\widehat{SR} - SR) \to N(0,\ 1 + SR^2/2)$
  - SR monthly, T the number of months
  - the standard error is $\sqrt{(1 + SR^2/2)/T}$
  - without Normality, the standard error requires a finite fourth moment, which Chapter 1 questions
- 1/N on sectors, $T = 271$: monthly $\widehat{SR} = 0.191$, standard error 0.061
  - Annualised $0.66 \pm 1.96 \times 0.21$: 95% interval $[0.25, 1.08]$
- Annualising by $\sqrt{q}$ ($q = 12$ months per year) is exact only without autocorrelation
  - in general $SR(q) = \eta(q)\, SR$

$$\eta(q) = \frac{q}{\sqrt{q + 2\sum_{k=1}^{q-1}(q-k)\rho_k}}$$

  - ρ_k: autocorrelation of monthly returns at lag k; positive ρ_k lower η, negative ρ_k raise it
- For 1/N on sectors: $\rho_1 = -0.03$, weighted sum $\sum_{k=1}^{11}(12-k)\rho_k = -0.50$
  - $\eta = 12/\sqrt{12 - 2(0.50)} = 12/\sqrt{11.0} = 3.62$ instead of $\sqrt{12} = 3.46$, annual Sharpe $0.191 \times 3.62 = 0.69$
  - $SR(q)$: the Sharpe ratio of the 12-month sum, not of the compounded annual return $\prod_t (1 + R_t) - 1$

Testing a Difference in Sharpe Ratios

- H_0: $SR_1 = SR_2$ for two strategies evaluated on the same months
- Jobson & Korkie (1981), in the form corrected by Memmel (2003) and given in Ledoit & Wolf (2008)
  - for i.i.d. Normal returns: $z = \hat{\Delta} / \sqrt{\hat{V}/T}$
  - $\hat{V} = 2(1-\hat{\rho}) + \frac{1}{2}\left(\widehat{SR}_1^2 + \widehat{SR}_2^2 - 2\widehat{SR}_1\widehat{SR}_2\hat{\rho}^2\right)$
  - $\hat{\rho}$: the correlation of the two strategies
- Ledoit & Wolf (2008): a robust version, by the delta method (linearising f around the estimates)

$$\hat{\Delta} = \widehat{SR}_1 - \widehat{SR}_2 = f(\hat{\mathrm{v}}), \qquad f(\mathrm{v}) = \frac{\mu_1}{\sqrt{\mathbb{E}r_1^2 - \mu_1^2}} - \frac{\mu_2}{\sqrt{\mathbb{E}r_2^2 - \mu_2^2}}$$

  - r_{1t}, r_{2t}: excess returns of the two strategies; μ_1, μ_2: their means; $\mathrm{v} = (\mu_1, \mu_2, \mathbb{E}[r_1^2], \mathbb{E}[r_2^2])$
  - standard error $\sqrt{\nabla f^\top \hat{\boldsymbol{\Psi}} \nabla f / T}$
  - $\hat{\boldsymbol{\Psi}}$: a HAC (heteroskedasticity and autocorrelation consistent) covariance of the four moments
  - HAC handles dependence, not the moment condition: the test requires a finite fourth moment, which Chapter 1 questions
- Strategies on the same assets are highly correlated: small differences can be significant

Further Topics: the HAC (robust standard errors) test step by step

The Block Bootstrap

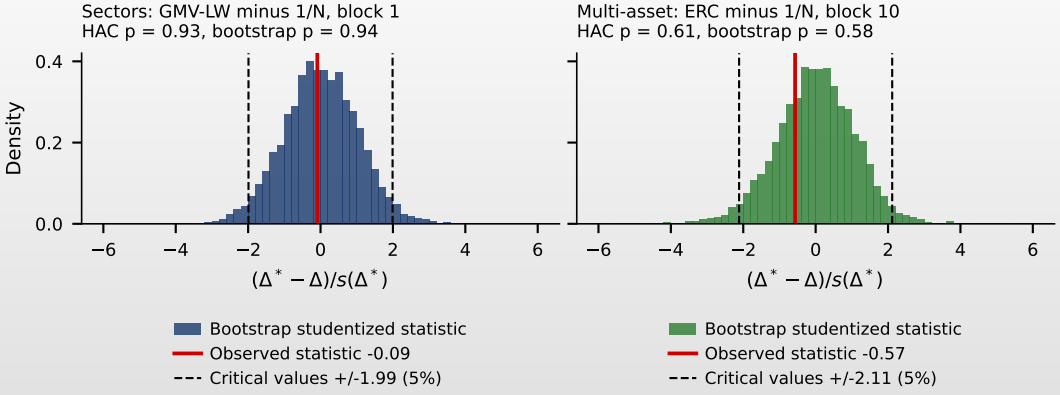


Bradley Efron, inventor of the bootstrap (1979)

Photo: Ryan K. Morris (2007); public domain; Wikimedia Commons

- Returns are dependent (volatility clustering)
  - the bootstrap resamples **blocks** of consecutive months (Künsch (1989))
- Ledoit & Wolf (2008), Section 3.2.2: circular blocks of **pairs** (r_{1t}, r_{2t})
  - wrapping around the end of the sample, so the correlation of the two strategies is kept
- The statistic in draw m: $\tilde{d}^{*m} = |\hat{\Delta}^{*m} - \hat{\Delta}| / s(\hat{\Delta}^{*m})$
  - centred at $\hat{\Delta}$ and divided by the standard error $s(\cdot)$
- Asymptotically pivotal (its limit distribution has no unknown parameters)
  - it can improve on the Normal approximation of the HAC test in finite samples
- $p = \left[\#\{m : \tilde{d}^{*m} \geq |\hat{\Delta}| / s(\hat{\Delta})\} + 1\right] / (M + 1)$, with $M = 4999$ draws
- Block length chosen by the calibration of Ledoit & Wolf (2008), Algorithm 3.1 (▸ Appendix)
  - 1 for all six sector rules
  - between 1 and 10, most often 10, for multi-asset and combined rules

Is Anything Better than 1/N? (1/2)



MFM_ch4_sharpe_test ▶ Interactive

Is Anything Better than 1/N? (2/2)

- Horizontal axis: $(\Delta^* - \hat{\Delta})/s(\Delta^*)$
  - the bootstrap distribution of the studentized difference against 1/N
- Red: the observed statistic $\hat{\Delta}/s(\hat{\Delta})$
  - -0.09 (sectors, GMV-LW) and -0.57 (multi-asset, ERC)
  - dashed: bootstrap critical values ± 1.99 and ± 2.11, in the same order
- Sectors, GMV-LW minus 1/N: -0.02, 95% interval $[-0.36, 0.33]$
  - HAC $p = 0.93$, bootstrap $p = 0.94$ (block 1)
- Multi-asset, ERC minus 1/N: -0.03, $[-0.15, 0.08]$
  - HAC $p = 0.61$, bootstrap $p = 0.58$ (block 10)
  - so correlated with 1/N that the interval is narrow, yet it covers zero
- Sectors: no rule differs from 1/N at 5%
  - the smallest p-value is ERC's ($+0.04$, HAC $p = 0.09$)
- Multi-asset: MV-LO $+0.29$ is **not** significant (HAC $p = 0.12$, bootstrap 0.16)
- The HAC and bootstrap p-values require a finite fourth moment, which Chapter 1 questions

MFM_ch4_sharpe_test

Multiple Testing across Portfolio Rules (1/2)

Universe	Rule minus 1/N	$\hat{\Delta}$	HAC p	Bootstrap p
Sectors	ERC	$+0.04$	0.0915	0.1296
Multi-asset	HRP	-0.16	0.0720	0.0460
Combined	GMV	-0.56	0.0304	0.0496
Combined	GMV-LO	-0.40	0.0477	0.0804
Combined	GMV-LW	-0.56	0.0238	0.0322
Combined	MV-LO	$+0.05$	0.8175	0.8180

- No rule beats 1/N significantly in any universe
- Without a multiple-testing adjustment, at 5%
  - three rules have a significantly lower Sharpe than 1/N by the HAC test (combined GMV, GMV-LO, GMV-LW)
  - three by the bootstrap (multi-asset HRP; combined GMV, GMV-LW)
- The GMV rules have low risk, but even lower return
- The table shows 6 of the 18 two-sided tests, H_0: $\Delta_k = 0$
  - 6 rules $\times$ 3 universes; all 18 in the Appendix

Multiple Testing across Portfolio Rules (2/2)

- Holm over all 18: smallest adjusted p is 0.43 (HAC, 18×0.024) and 0.58 (bootstrap, 18×0.032)
  - nothing survives (next slide; Chapter 3)
- A different question, the best of the 21 configurations (7 rules, without MV, $\times$ 3 universes)
  - the deflated Sharpe ratio (DSR) of Bailey & López de Prado (2014), the probability that the winner's Sharpe ratio is not luck alone, on the common calendar May 2012 – July 2026: $0.94 < 0.95$ for MV-LO on sectors
  - but the verdict hinges on the effective number of trials (▸ Appendix)
  - the DSR uses the winner's kurtosis, so it requires a finite fourth moment, which Chapter 1 questions
- Type of evidence: comparison of procedures (it shows which method works better on the same data, not why)
  - the rules are evaluated on the same months and assets, in these samples

Holm and Benjamini–Hochberg on This Chapter's Tests

- **FWER** (family-wise error rate): the probability of at least one false rejection in the family
- **FDR** (false discovery rate): the expected share of false rejections among the rejections
- Sort the m p-values, $p_{(1)} \leq \cdots \leq p_{(m)}$
  - **Holm** (controls the FWER under any dependence): $\tilde{p}_{(i)} = \max_{j \leq i} \min\{1, (m - j + 1)\, p_{(j)}\}$
  - **Benjamini–Hochberg** (BH, Benjamini & Hochberg (1995)): $\tilde{p}_{(i)} = \min_{j \geq i} \min\{1, \frac{m}{j}\, p_{(j)}\}$
  - BH controls the FDR for valid p-values under independence or under PRDS (positive dependence, Benjamini & Yekutieli (2001)); positive correlations between strategies do not alone establish PRDS
  - under arbitrary dependence: Holm (the main result) or Benjamini–Yekutieli
  - reject where $\tilde{p} < 0.05$ (Holm) or $\tilde{p} <$ the chosen FDR (BH)
- The four smallest two-sided HAC p-values, $m = 18$

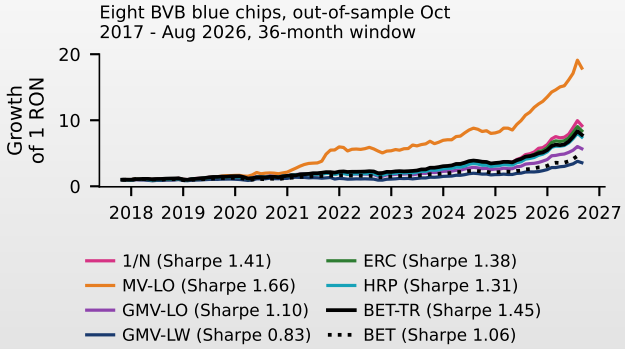
Test	$p_{(j)}$	Holm	$\frac{18}{j}\, p_{(j)}$
Combined, GMV-LW	0.0238	$18 \times 0.0238 = 0.428$	0.428
Combined, GMV	0.0304	$17 \times 0.0304 = 0.516$	0.273
Combined, GMV-LO	0.0477	$16 \times 0.0477 = 0.763$	0.286
Multi-asset, HRP	0.0720	$\min\{1, 15 \times 0.0720\} = 1.00$	0.324

- BH takes the minimum from the bottom up: the first two both become 0.273
  - no rejection at an FDR of 10%

Setup: a Romanian Investor

- The question of this section: does an allocation rule beat the BET-TR index on BVB blue chips
  - for an investor in RON (Romanian leu)?
- Eight blue chips (the most liquid large-company stocks) with complete histories since October 2014, by BVB ticker: Banca Transilvania (TLV), OMV (Österreichische Mineralölverwaltung, Austrian Mineral Oil Administration, OMV group) Petrom (SNP), BRD–Groupe Société Générale (BRD), Transgaz (TGN), Romgaz (SNG), Nuclearelectrica (SNN), Electrica (EL), Transelectrica (TEL)
- A 36-month window, not 60: with 60 months only 83 months would remain out of sample
  - first portfolio October 2017
- Rules: 1/N, MV-LO, GMV-LO, GMV-LW, ERC, HRP
  - no unconstrained MV and GMV, whose weights would be extreme on 36-month windows
- Benchmark: BET-TR (total return, with dividends)
  - BET is shown only to measure the dividend effect
- Sharpe ratios with $R_f = 0$ (no RON money-market series in the data)
  - hence higher than excess-return Sharpe ratios
- Earlier Romanian evidence on diversification: Dragotă, Șerbănescu & Pele (2008)

BVB Portfolios versus BET-TR (1/2)



- Chart: the value of RON 1 invested in October 2017
  - out of sample to August 2026 (107 months), before costs
- BET-TR: mean 24.7% a year, volatility 17.0%, Sharpe 1.45
- BET (no dividends): mean 17.8%, Sharpe 1.06
- Close to BET-TR: 1/N 1.41, ERC 1.38, HRP 1.31

MFM_ch4_bvb ▶ Interactive

BVB Portfolios versus BET-TR (2/2)

- MV-LO: Sharpe 1.66, +0.21 over BET-TR
  - HAC $p = 0.44$, bootstrap 0.53
- GMV-LW: Sharpe 0.83, 0.62 below BET-TR
  - HAC $p = 0.050$, bootstrap 0.10
  - these p-values require a finite fourth moment, which Chapter 1 questions
- Average correlation of the eight stocks 0.44 limits diversification
  - the latest GMV-LO puts 70% in TLV (Seminar 4, C2)
- Type of evidence: comparison of procedures

MFM_ch4_bvb

Which Frontier Can a Large Investor Reach? (1/2)

- The question: which allocation rule stays efficient when trading costs grow with the size of the investor?
- Jensen, Kelly, Malamud & Pedersen (2026), *Review of Financial Studies*
  - how should portfolios be built from machine-learning (ML) return forecasts when every trade moves prices?
- **Implementable efficient frontier** (eq. 12)
  - for each volatility σ, the highest expected return **net of trading costs**

$$k(\sigma) = \max_{\pi} \; E[r_{t+1}^{\text{net}}] \quad \text{s.t.} \quad E[\pi_t^\top \boldsymbol{\Sigma} \pi_t] = \sigma^2$$

  - π_t: vector of risky-asset weights at t (the paper's notation; not the implied returns $\boldsymbol{\pi}$ of Black–Litterman); r_{t+1}^{net}: portfolio return after trading costs
  - $k(\sigma)$: the best net expected return at volatility σ
- Proposition 1: the net Sharpe ratio $k(\sigma)/\sigma$ falls along the frontier
  - a larger investor faces a lower frontier

Which Frontier Can a Large Investor Reach? (2/2)

- Cost as a fraction of wealth W_t (eqs. 4, 7), quadratic in the trade

$$c_t = \tfrac{W_t}{2}(\pi_t - g_t\pi_{t-\mathbf{1}})^\top \mathbf{\Lambda}(\pi_t - g_t\pi_{t-\mathbf{1}}), \qquad \mathbf{\Lambda} = \mathrm{diag}(0.2/V_i)$$

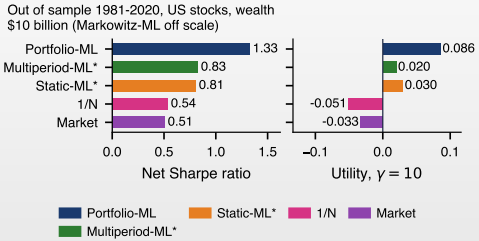
  - g_t: drifts last month's weights; V_i: the average daily dollar volume of asset i

- Optimal rule: trade only part of the way towards an **aim portfolio** A_t (Gârleanu & Pedersen (2013))

$$\pi_t = m\, g_t\pi_{t-\mathbf{1}} + (\mathbf{I} - m)A_t$$

  - A_t: a forward-looking aim, a weighted average of today's and expected future Markowitz portfolios
  - the weighting reflects the persistence of the signals and the frictions; A_t is in general not the current Markowitz portfolio
  - m: an $N \times N$ matrix of trade rates (eq. 14), so $\mathbf{I} - m$ is a matrix too; in the scalar special case of Gârleanu & Pedersen (2013), m is the share of the old position kept each period
  - Portfolio-ML learns A_t from the signals by a ridge regression (least squares with a penalty on the size of the coefficients) whose objective is utility net of costs

The Paper's Results: Net Sharpe and Utility



- Jensen, Kelly, Malamud & Pedersen (2026), Table 2: net of costs, 1981–2020, $\gamma = 10$
  - utility (eq. 11) = annual net return $-\frac{\gamma}{2}\times$ annual variance
- Markowitz-ML (mean–variance on ML forecasts, no costs)
  - gross Sharpe 2.06, turnover 79.36, leverage 74.61: net return extremely negative
- Portfolio-ML: net Sharpe 1.33, utility 0.086 (net return 15%, volatility 11%, cost 0.6% a year)
- Multiperiod-ML* (12 horizons): net Sharpe 0.83, utility 0.020
- Static-ML* (one-period costs; * = extra tuning layer) 0.81, 0.030
  - 1/N 0.54, −0.051; market 0.51, −0.033
  - only cost-aware rules have positive utility
- Type of evidence: comparison of procedures
  - stated limit: on the smallest 20% of stocks, Portfolio-ML trails Multiperiod-ML* and Static-ML*

MFM_ch4_implementable_frontier

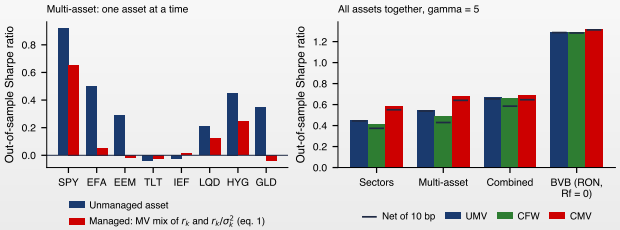
Volatility-Managed Portfolios: The Question

- The question: does the Sharpe ratio rise if the investor cuts exposure when volatility is high?
  - yes, if measured volatility predicts next month's risk and expected returns do not rise in proportion to variance
- Moreira & Muir (2017): the volatility-managed factor k (eq. 1 in DeMiguel, Martín-Utrera & Uppal (2024))

$$r^{\sigma}_{k,t+1} = \frac{c}{\sigma^{2}_{k,t}}\, r_{k,t+1}$$

  - $r_{k,t+1}$: the excess return of factor k in month $t+1$
  - $\sigma^{2}_{k,t}$: the variance of the daily returns of factor k in month t
  - c: the constant that gives r^{σ}_{k} the same volatility as r_k
  - the investor holds the mean–variance combination of r_k and r^{σ}_{k}
- Cederburg, O'Doherty, Wang & Yan (2020): out of sample, managed versions do not systematically beat the unmanaged strategies; estimation error in the combination weights matters
- Barroso & Detzel (2021): after trading costs, the gain is not robust for most of the factors studied
- DeMiguel, Martín-Utrera & Uppal (2024), *Journal of Finance*: the same question for a portfolio of many factors
  - the paper's claim: the conditional mean–variance multifactor portfolio (CMV) beats its unconditional counterpart (UMV) out of sample and net of costs
  - Table II, nine factors, 1977–2020: net Sharpe 1.062 (CMV) against 0.940 (UMV), $p = 0.006$

Volatility Management on the Course Data



- Left: the out-of-sample Sharpe ratio of each multi-asset ETF, May 2012 – July 2026
  - blue: the asset unchanged; red: the estimated mean–variance combination of r_k and r_k/σ_k^2 (eq. 1)
- Right: out-of-sample Sharpe ratio of UMV, CFW (UMV scaled by its volatility) and CMV in the four universes; dash: net of 10 bp
- One asset at a time: volatility management lowers the Sharpe ratio of 22 of the 25 assets (sectors, multi-asset, BVB)
  - EFA: from 0.50 to 0.05; gold (GLD): from 0.35 to -0.04
  - seven falls have $p < 0.05$ in individual, unadjusted HAC tests
  - none after Holm over the family of 25 comparisons: the smallest adjusted p-value is 0.16
- All assets together: CMV has the highest gross Sharpe ratio in every universe
  - sectors 0.59 against 0.45 (UMV); multi-asset 0.69 against 0.55
  - combined 0.69 against 0.67; BVB 1.32 against 1.29
- CFW, which scales the whole UMV portfolio as in Moreira & Muir (2017), stays below UMV in all four universes

MFM_ch4_vol_managed ▶ Interactive

The Paper's Method and the Course Replication

- CMV (eqs. 2–6): the weight of factor k, set at the end of month t and applied to the return $r_{k,t+\mathbf{1}}$, is $\theta_{k,t} = a_k + b_k/\sigma_t$
  - σ_t: the realised volatility of the market in month t, from daily returns
  - $b_k > 0$: less exposure to factor k when the market is volatile
  - $a_k, b_k \geq 0$ maximise $\boldsymbol{\mu}^\top \eta - \widehat{TC}(\eta) - \frac{\gamma}{\mathbf{2}}\, \eta^\top \boldsymbol{\Sigma}\, \eta$, with $\gamma = 5$
  - paper notation: θ here is not the maximum Sharpe ratio, η not Lo's factor; $\eta = (a, b)$; $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$: mean and covariance of the extended returns $(r_{k,t+\mathbf{1}},\ r_{k,t+\mathbf{1}}/\sigma_t)$; $\widehat{TC}$: the mean trading cost
- UMV: the same problem with $b_k = 0$
- CFW: the mean–variance combination of UMV and UMV divided by its variance last month (Moreira & Muir (2017))
- The paper: the cost of each stock from its monthly estimated spread; trades net out across factors (trading diversification)
- The paper: expanding window, the first one 120 months long
- The course replication: the "factors" are the excess returns of the ETFs (positions financed at the T-bill rate)
  - BVB: returns in RON, with $R_f = 0$, as in the rest of the chapter
  - the market: SPY for the US universes, BET-TR for BVB
  - expanding window, first 60 months (BVB: 36): the out-of-sample periods of the chapter's backtest
  - weights optimised without costs, as in Table III, columns 1–3
  - cost: $c = 10$ bp per unit of turnover, with turnover computed per ETF (trades net out)
  - the chapter's liquidation rule: no rule loses 100% in a month, so no rule is liquidated
- Tests: the paper's one-sided stationary bootstrap (mean block 5, 10,000 draws) and the chapter's two-sided HAC test

The Result: A Small, Insignificant Gain

- Gross Sharpe difference CMV − UMV, from unrounded values (one-sided bootstrap p; two-sided HAC p)
  - sectors, 271 months: 0.14 (0.13; 0.26)
  - multi-asset, 171 months: 0.13 (0.25; 0.46)
  - combined, 171 months: 0.02 (0.46; 0.92)
  - BVB, 107 months: 0.03 (0.30; 0.55)
- CMV trades more: monthly turnover 1.08 against 0.21 (multi-asset)
- CMV uses more leverage: mean sum of absolute weights 4.8 against 3.3 (multi-asset)
- The cost per unit of turnover at which the CMV advantage disappears
  - sectors 45 bp, multi-asset 42 bp, BVB 49 bp, combined 7 bp
  - at 10 bp, CMV falls below UMV in the combined universe: net Sharpe 0.646 against 0.655
- Type of evidence: comparison of procedures
  - the paper's direction appears in all four universes: CMV above UMV, the managed combination of a single asset below the unchanged asset
  - the conclusion concerns this implementation (ETFs, expanding window, $\gamma = 5$); it does not show that volatility management fails in general
  - none of the four CMV − UMV differences is significant at 10%

Limits of the Replication

- 8–16 ETFs and 107–271 months, against nine long-short factors over 528 months
- The p-values require a finite fourth moment, which Chapter 1 questions
- BVB: leverage is financed at $R_f = 0$, so the mean is overstated
- US: leverage (mean exposure 3.3–4.8, multi-asset) is financed at the T-bill rate; a borrowing spread would lower the mean more than a few bp of trading costs
- The cost of setting up the initial positions is not included
- In the paper, without netting trades across factors, CMV loses: net Sharpe 0.613 against 0.710 (Table III, Panel C, columns 1–3, optimised without costs)

Open Questions

- Questions
  - does any **estimated** allocation rule beat 1/N out of sample, net of costs?
  - does it still beat 1/N once every rule, window and universe tried is counted?
- Live evidence points both ways
  - Nonlinear shrinkage of Σ improves GMV out of sample in large cross-sections: Ledoit & Wolf (2017)
  - Combinations under estimation risk, no risk-free asset: Kan, Wang & Zhou (2022); with 1/N: Tu & Zhou (2011)
  - Backtests need haircuts for the number of specifications tried: Harvey & Liu (2015)
- Formal statement: H_0: $\Delta_k \leq 0$ for all 18 (rule, universe) pairs
  - against H_1: $\Delta_k > 0$ for at least one k
  - with the FWER or FDR fixed for the whole family (Multiple Testing across Portfolio Rules)
- One pass of the loop with an AI assistant
  - literature with DOIs, competing hypotheses, code on the course's monthly data
  - a pre-registered family of tests, a referee-style critique (prompts in the Appendix)
- Falsified by: a positive $\hat{\Delta}_k$ that survives a family-wise correction, costs and a second sample

What the Researcher Must Check

- **Look-ahead**: weights at t built with month-t returns
  - test: assert that the last estimation date is before t
- **Survivorship and selection**: an ETF list chosen today
  - test: use the universe available at each date, including later delisted ETFs, or disclose the limitation
- **Multiple testing**: 6 rules $\times$ 3 universes already give 18 tests
  - test: Holm (FWER) or Benjamini–Hochberg (FDR) adjusted p-values
- **Power**: for MV-LO minus 1/N on the multi-asset universe, $\hat{\Delta}$ has a HAC standard error of 0.18 on 171 months
  - test: compute the detectable effect before running
- **Return type**: portfolios aggregate **simple** returns, Sharpe ratios use **excess** returns
  - test: recompute one month step by step
- **Invented references**: an unresolvable DOI (Digital Object Identifier) or a title mismatch
  - the claim is dropped and the error goes into `AI_ERRORS.md`

Mini-Case: Data Snooping in This Chapter's Backtest (1/2)

- A typical assistant proposal: "MV-LO beats 1/N on the multi-asset universe"
  - superiority is one-sided: H_0: $\Delta \leq 0$, $p = 1 - \Phi(\hat{\Delta}/\mathrm{SE})$
  - Φ the standard Normal distribution function, SE the standard error
- Replicate: $\hat{\Delta} = +0.29$, HAC standard error 0.18, one-sided $p = 0.06$
  - two-sided HAC 0.12, bootstrap 0.16
  - $z = 0.2863/0.1836 = 1.56$, $p = 1 - \Phi(1.56) = 0.06$
- Power 0.8 for $\Delta = 0.29$, one-sided at 5%, needs about 435 months (36 years)
  - 919 months (77 years) at the Holm level 5%/18
  - months needed: $T_{\mathrm{obs}}\left[(z_{1-\alpha} + z_{0.8})\,\mathrm{SE}/\Delta\right]^2 = 171\,\left[(1.645 + 0.842) \times 0.1836/0.2863\right]^2 \approx 435$
  - at $\alpha = 0.05/18$, $z_{1-\alpha} = 2.77$, which gives ≈ 919

MFM_ch4_ai_discovery

Mini-Case: Data Snooping in This Chapter's Backtest (2/2)

- Test the family, not the winner: 18 tests (6 rules $\times$ 3 universes)
  - one-sided p-values of all 18 in the Appendix
  - Only 7 of 18 differences are positive; the 3 two-sided HAC p-values below 0.05 are all **losses**, which never reject a superiority null
  - One-sided: smallest $p = 0.046$ (sectors, ERC); adjusted Holm 0.82, Benjamini–Hochberg 0.53; 0 discoveries at an FDR of 10%
- Why 0.82 here but 0.43 on Multiple Testing across Portfolio Rules
  - a loss has a large one-sided p (about $1 - p_{\text{two-sided}}/2$, e.g. 0.99 for combined GMV-LW)
  - so the significant losses drop out
  - the smallest one-sided p belongs to a gain, 0.046, and Holm multiplies it by 18
- Interpretation: the data cannot rank these rules
  - the bootstrap interval for MV-LO, $[-0.13, 0.70]$, admits losses and large gains
  - failure to reject is not proof of equality
- Type of evidence: comparison of procedures

MFM_ch4_ai_discovery

Your Turn: A Research Project Idea

- **Questions**
  - does nonlinear shrinkage (Ledoit & Wolf (2017)) beat 1/N on the 16 ETFs of this chapter, after 10–50 bp costs?
  - does a combination of an optimised rule and 1/N (Tu & Zhou (2011)) beat 1/N on the same ETFs, after the same costs?
- **Replicate first**
  - the combined-universe results (1/N Sharpe 0.76, MV-LO 0.81)
  - the adjusted p-values of the mini-case
- **Pre-register**, before any result
  - hypotheses, family of tests, error rate and power
  - a placebo: random long-only weights with the same turnover
- **Report**: code on the course's monthly data, a table of all specifications, a referee-style self-critique
- **Log**: prompts in `AI_USE.md`
  - every error caught and the test that caught it in `AI_ERRORS.md`

The Practical Decision

- **What I choose**
  - 1/N as the default: it estimates nothing, has a monthly turnover of 0.02 and is hard to beat after costs
  - ERC if I want risk balanced across assets: a moderate tilt of 1/N with turnover 0.03 (HRP trades more)
  - GMV or GMV-LW only if the goal is minimum risk: the lowest volatility, but a Sharpe ratio below 1/N on 16 ETFs (0.20–0.36 versus 0.76)
  - not unconstrained MV: its wealth reaches zero in all three universes
- **Under which conditions**
  - 60-month windows and 8–16 assets: estimation error dominates (in simulation, MV beats 1/N only at 480 months)
  - costs above about 22 bp: MV-LO, the best point estimate, no longer beats 1/N on 16 ETFs
  - Black–Litterman only with views whose confidence $\boldsymbol{\Omega}$ is set cautiously: it moves only the assets in the views
  - these verdicts are specific to these samples (sectors 2004–2026, multi-asset and combined 2012–2026)
- **What result would make me drop it**
  - an estimated rule with $\hat{\Delta} > 0$ against 1/N, significant after Holm or Benjamini–Hochberg over the whole family of tests
  - a gain that survives net of costs and on a second sample

Quick Checks with Toy Numbers (1/2)

- **Three rules, two assets**: $\sigma_1 = 30\%$, $\sigma_2 = 15\%$, $\rho = 0.2$
  - compute the weights and volatilities of GMV, ERC and 1/N; check $\sigma_{\text{GMV}} \leq \sigma_{\text{ERC}} \leq \sigma_{1/N}$
  - what share of the 1/N variance comes from asset 1? (Two Assets; Risk Contributions; Equal Risk Contribution)
- **Optimism of the plug-in Sharpe**: $N = 4$ assets, $T = 120$ months, true $\theta = 0.50$ a year
  - what in-sample maximum Sharpe (root mean square, annual) should you expect? (The Plug-in Sharpe Ratio Is Biased)
- **Noise band**: $N = 10$, $T = 40$, correlation matrix
  - compute the Marchenko–Pastur band; is a sample eigenvalue of 2.0 evidence of structure? (Why the Sample Covariance Matrix Fails)

Quick Checks with Toy Numbers (2/2)

- **One view, one asset**: $\pi = 4\%$, $\tau\sigma^2 = 0.0009$, view 10%
  - posterior mean and view weight for $\Omega = 0.0009$ and for $\Omega = 0.0003$ (Views and the Posterior)
- **One Sharpe ratio**: monthly $\widehat{SR} = 0.15$, $T = 120$, i.i.d. returns from the Normal distribution
  - annual Sharpe and its 95% interval: is it significantly positive? (Inference on One Sharpe Ratio)
- **Four tests**: $p = 0.004,\ 0.020,\ 0.030,\ 0.300$
  - Holm and Benjamini–Hochberg adjusted p-values; how many rejections at 5% under each? (Holm and Benjamini–Hochberg)
- Answers: Appendix

Check Yourself, and Next

Retrieval questions

- Why does the in-sample tangency Sharpe overstate the achievable one?
  - plug-in Sharpe; Seminar 4, A8
- By how much does it overstate it for $N = 9$, $T = 60$?
- Why does a long-only constraint act like shrinkage?
  - Jagannathan–Ma and its two-asset example
- Why can a small Sharpe difference between ERC and 1/N be nearly significant while a larger one between GMV and 1/N is not?
  - sectors; HAC test: compare the standard errors 0.023 and 0.159 in the Appendix table

Links to other chapters

- **Chapter 3**: the tangency portfolio and the CAPM
  - factor covariance models
- **Chapters 5–6**: time-varying volatility and correlation for dynamic weights
- **Chapter 7**: portfolio Value-at-Risk and Expected Shortfall
- **Chapter 13**: machine learning for returns and portfolios
  - backtest overfitting

The Assets in Detail

- **Sectors** (9 US sector ETFs)
  - XLB (materials), XLE (energy), XLF (financials), XLI (industrials), XLK (technology)
  - XLP (staples), XLU (utilities), XLV (health care), XLY (discretionary)
- **Multi-asset**
  - equities: SPY (S&P 500), EFA (developed markets outside the US and Canada), EEM (emerging markets)
  - bonds: TLT (US Treasuries, 20+ years), IEF (US Treasuries, 7–10 years), LQD (investment grade), HYG (high yield); gold: GLD
- **BVB** (8 blue chips)
  - TLV Banca Transilvania, SNP OMV Petrom (OMV: Österreichische Mineralölverwaltung, the Austrian oil group), BRD BRD–Groupe Société Générale, TGN Transgaz
  - SNG Romgaz, SNN Nuclearelectrica, EL Electrica, TEL Transelectrica

Nonlinear Shrinkage

- Linear shrinkage mixes two matrices with one δ
  - **nonlinear** shrinkage keeps the sample eigenvectors and corrects each eigenvalue separately
  - $\hat{\Sigma}_{\text{NL}} = \mathrm{U}\,\mathrm{diag}(\tilde{d}_1, \ldots, \tilde{d}_N)\mathrm{U}^\top$, U: the eigenvectors of S, $\tilde{d}_i$: the corrected eigenvalues
  - Ledoit & Wolf (2017): tuned for portfolio selection; Ledoit & Wolf (2020): an analytical formula, no optimisation
- Out of sample, 60-month window, GMV volatility
  - Sectors: sample 11.90%, linear (constant correlation) 11.80%, nonlinear 11.62%
  - 16 ETFs: sample 4.06%, linear 4.38%, nonlinear 4.31%
- A constant correlation fits neither stock–stock (positive) nor stock–bond (negative) pairs
  - the target matters (Seminar 4, B7)
- No estimator dominates
  - nonlinear wins on sectors, the sample matrix on 16 ETFs
  - the choice is made out of sample
- At $N = 16$, $T = 60$ the sample matrix is still competitive
  - it becomes unusable as N approaches T

Black–Litterman: Limiting Cases

- $\boldsymbol{\Omega} \to 0$: the views hold exactly, $\mathrm{P}\boldsymbol{\mu}_{\mathrm{BL}} = \mathrm{q}$
- $\boldsymbol{\Omega} \to \infty$: the views are ignored, $\boldsymbol{\mu}_{\mathrm{BL}} = \boldsymbol{\pi}$
- Without views, the benchmark w_b itself
- With the same $\boldsymbol{\Sigma}$ in the weights, only the assets in the views move
- The same estimator is a mixed regression estimated by generalised least squares (GLS) (Theil & Goldberger (1961))
  - derivation: Appendix; Seminar 4, A6

Black–Litterman: What δ_{RA} and τ Do

- The values of He & Litterman (2002): $\delta_{\mathrm{RA}} = 2.5$, $\tau = 0.05$, $\boldsymbol{\Omega} = \mathrm{diag}(\mathrm{P}\,\tau\boldsymbol{\Sigma}\,\mathrm{P}^{\top})$
- **Risk aversion**: premultiply $\boldsymbol{\pi} = \delta_{\mathrm{RA}}\boldsymbol{\Sigma}\mathrm{w}_b$ by $\mathrm{w}_b^{\top}$

$$\delta_{\mathrm{RA}} = \frac{\mathrm{w}_b^{\top}\boldsymbol{\pi}}{\mathrm{w}_b^{\top}\boldsymbol{\Sigma}\mathrm{w}_b} = \frac{\text{benchmark excess return}}{\text{benchmark variance}}$$

  - e.g. 6% excess return at 15.5% volatility: $0.06/0.155^2 = 2.5$
  - $\boldsymbol{\pi}$ is linear in δ_{RA}: doubling it doubles the implied XLK minus XLU spread of 2.2%
- **Prior uncertainty**: $\tau\boldsymbol{\Sigma}$ is the uncertainty about the mean
  - one reading is $\tau \approx 1/Y$, the variance of a mean estimated over Y years relative to $\boldsymbol{\Sigma}$ (Merton's argument)
  - so $\tau = 0.05$ is about 20 years of data
- With this $\boldsymbol{\Omega}$ and one view, τ cancels
  - the posterior spread is the midpoint $(2.2\% + 3\%)/2 = 2.6\%$
  - the weights do not depend on τ
- The confidence in the view is set by $\boldsymbol{\Omega}$ relative to $\tau\boldsymbol{\Sigma}$, not by τ alone

The HAC Test Step by Step

- Gradient of f with respect to $\mathrm{v} = (\mu_1, \mu_2, \mathbb{E}[r_1^2], \mathbb{E}[r_2^2])$
  - writing $\sigma_i^2 = \mathbb{E}[r_i^2] - \mu_i^2$

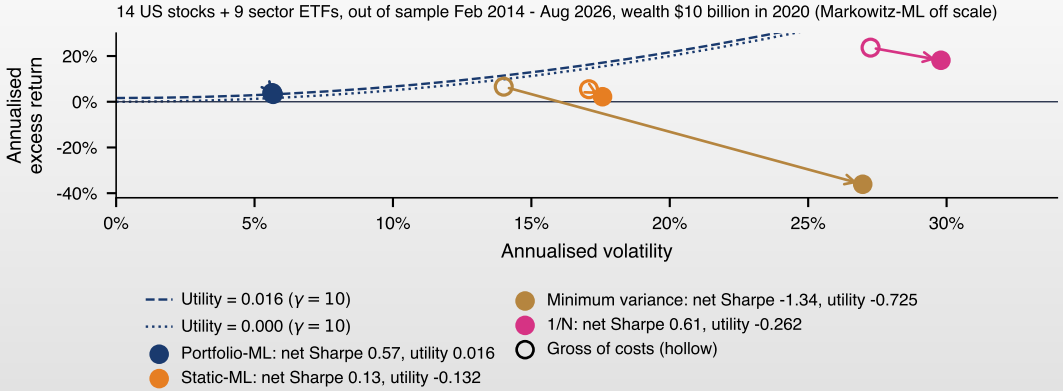
$$\nabla f = \Big(\frac{\mathbb{E}[r_1^2]}{\sigma_1^3},\ -\frac{\mathbb{E}[r_2^2]}{\sigma_2^3},\ -\frac{\mu_1}{2\sigma_1^3},\ \frac{\mu_2}{2\sigma_2^3} \Big)$$

  - these are derivatives in the **uncentred** second moments; a gradient in (μ_i, σ_i^2) does not fit this $\hat{\boldsymbol{\Psi}}$
- $\hat{\boldsymbol{\Psi}}$: long-run covariance of $\mathrm{y}_t = (r_{1t} - \hat{\mu}_1,\ r_{2t} - \hat{\mu}_2,\ r_{1t}^2 - \widehat{\mathbb{E}r_1^2},\ r_{2t}^2 - \widehat{\mathbb{E}r_2^2})$
- HAC p-values in this chapter's tables: Bartlett (Newey–West) kernel
  - $\lfloor 4(T/100)^{2/9} \rfloor = 4$ lags for every T from 107 to 271
  - $z = \hat{\Delta}/\mathrm{SE}$, $p = 2[1 - \Phi(|z|)]$
- The bootstrap studentizes with a different SE (prewhitened quadratic-spectral kernel, Appendix)
  - so the two statistics differ slightly
  - sectors, MV-LO minus 1/N: $\hat{\Delta} = 0.063$; Bartlett SE 0.120, $z = 0.53$, HAC $p = 0.60$; quadratic-spectral (QS) SE 0.127, studentized statistic 0.50

Paper and Course Application: What Is Compared?

	Paper	**Course replication**
Assets	US stocks with market capitalisation above the NYSE (New York Stock Exchange) median (184–805)	14 US stocks + 9 sector ETFs
Signals	115 characteristics (Jensen, Kelly & Pedersen (2023))	5 price and volume signals
Out of sample	1981–2020, monthly	February 2014 – August 2026 (151 months)
Investor	$\gamma = 10$, wealth \$10 billion at the end of 2020, growing with the market	same; wealth scaled by the cumulative SPY return before and after 2020
Utility	annual net return $-\frac{\gamma}{2}\times$ annual variance	same: $12\times$ monthly mean $-\frac{\gamma}{2}\times 12\times$ monthly variance

Gross and Net of Costs on the Course Data (1/2)



MFM_ch4_implementable_frontier

Gross and Net of Costs on the Course Data (2/2)

- Axes: annualised volatility (horizontal) and annualised excess return (vertical)
  - hollow marker: before costs
  - filled: after the cost c_t (eq. 35 of the paper)
- Dashed curve: the utility of Portfolio-ML (0.016)
  - dotted curve: utility 0
- Portfolio-ML: return 3.8% gross and 3.2% net, volatility 5.7%, net Sharpe 0.57
- Utility of Portfolio-ML: $0.0324 - \frac{10}{2} \times 0.0567^2 = 0.016$ (net mean and net volatility, annual)
- 1/N: 18.1% net and Sharpe 0.61, but its 29.8% volatility gives a utility of -0.26
- Minimum variance: $+6.6\%$ gross, -36.1% net
  - its leverage of 2.6 costs 43% a year
- Limits: 23 assets, 5 signals, no factor model for Σ (Seminar 4, C3)

MFM_ch4_implementable_frontier

The Frontier in Closed Form, Step by Step

- Two constraints, two multipliers: $\mathcal{L} = \frac{1}{2}\mathrm{w}^\top \boldsymbol{\Sigma}\mathrm{w} - \lambda_1(\mathrm{w}^\top \boldsymbol{\mu} - m) - \lambda_2(\mathrm{w}^\top 1 - 1)$
- First-order condition: $\boldsymbol{\Sigma}\mathrm{w} = \lambda_1\boldsymbol{\mu} + \lambda_2 1$, so $\mathrm{w} = \lambda_1\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} + \lambda_2\boldsymbol{\Sigma}^{-1}1$
- The constraints, with $A = 1^\top\boldsymbol{\Sigma}^{-1}1$, $B = 1^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$, $C = \boldsymbol{\mu}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$, $D = AC - B^2$

$$\boldsymbol{\mu}^\top \mathrm{w} = \lambda_1 C + \lambda_2 B = m, \qquad 1^\top \mathrm{w} = \lambda_1 B + \lambda_2 A = 1$$

- Solving the 2×2 system: $\lambda_1 = (Am - B)/D$, $\lambda_2 = (C - Bm)/D$
- Variance: $\mathrm{w}^\top\boldsymbol{\Sigma}\mathrm{w} = \mathrm{w}^\top(\lambda_1\boldsymbol{\mu} + \lambda_2 1) = \lambda_1 m + \lambda_2 = \dfrac{Am^2 - 2Bm + C}{D}$
- Check: at $m = B/A$ the variance is $(C - B^2/A)/D = 1/A$, the GMV
- Shapes: σ^2 is quadratic in m (a parabola in the (σ^2, m) plane)
  - $\sigma^2 - A(m - B/A)^2/D = 1/A$ is a hyperbola in the (σ, m) plane

Back

Wishart Moments and the Distribution of Weights

- i.i.d. $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ returns, maximum-likelihood estimators
  - $\hat{\boldsymbol{\mu}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$, $T\hat{\boldsymbol{\Sigma}} \sim W_N(T-1, \boldsymbol{\Sigma})$, independent
  - Wishart: the matrix analogue of χ^2, $T-1$ degrees of freedom
- Inverse Wishart moment, finite for $T > N+2$: $\mathbb{E}[\hat{\boldsymbol{\Sigma}}^{-1}] = \frac{T}{T-N-2}\boldsymbol{\Sigma}^{-1}$
  - an **upward** bias that grows with N/T
- Conditioning on $\hat{\boldsymbol{\mu}}$ and using $\mathbb{E}[\mathrm{x}^\top \mathrm{K} \mathrm{x}] = \mathrm{tr}(\mathrm{K}\,\mathrm{Var}(\mathrm{x})) + \mathbb{E}[\mathrm{x}]^\top \mathrm{K}\,\mathbb{E}[\mathrm{x}]$
  - any random vector x, fixed matrix K, tr: trace
  - gives $\mathbb{E}[\hat{\theta}^2]$ of the lecture (Seminar 4, A8)
- GMV weights with Wishart $\hat{\boldsymbol{\Sigma}}$ and T large enough relative to N
  - $\hat{\mathrm{w}}_{\mathrm{GMV}}$ is unbiased and its exact distribution and covariance are known (Okhrin & Schmid (2006))
- Tangency weights divide by $1^\top \hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\mu}}$, which can be near zero: very heavy tails
- Estimation-risk-optimal combinations: tangency and GMV with a risk-free asset (Kan & Zhou (2007)); without one (Kan, Wang & Zhou (2022))

Back

Tangency Weights by Regression, Britten-Jones (1999) (1/2)

- Britten-Jones (1999): regress the constant 1 on the excess returns, **without** a constant term

$$1 = \mathrm{b}^\top \mathrm{r}_t + u_t, \qquad \hat{\mathrm{b}} = (\mathrm{X}^\top \mathrm{X})^{-1} \mathrm{X}^\top 1_T \propto \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}$$

- r_t: vector of excess returns in month t
  - X: $T \times N$ matrix with rows $\mathrm{r}_t^\top$
  - 1_T: vector of T ones; u_t: regression error
- Sherman–Morrison on $\mathrm{X}^\top \mathrm{X}/T = \hat{\boldsymbol{\Sigma}} + \hat{\boldsymbol{\mu}} \hat{\boldsymbol{\mu}}^\top$ gives the proportionality (Seminar 4, A9)
  - $\hat{\mathrm{b}}/1^\top \hat{\mathrm{b}}$ = sample tangency weights
- $1^\top \hat{\mathrm{b}}$ has the sign of $1^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}$
  - the portfolio is on the efficient branch only if it is positive

Back

Tangency Weights by Regression, Britten-Jones (1999) (2/2)

- OLS t-test of $b_i = 0$: asset i has zero weight in the tangency portfolio
  - F-test of $\mathrm{Rb} = 0$ with $\mathrm{R} = [\mathrm{I}_{N-1}, -1_{N-1}]$: "the tangency portfolio is 1/N"
- Exact finite-sample t and F distributions under i.i.d. Normal returns
  - HAC standard errors otherwise
- Nine sectors, January 1999 – July 2026 ($T = 331$): normalised $\hat{\mathrm{b}}$ reproduces the tangency weights
  - largest $|t| = 1.12$ (financials), HAC 1.19
- H_0: tangency = 1/N, eight restrictions: $F(8, 322) = 0.48$, $p = 0.87$
  - HAC Wald $\chi^2_8 = 4.50$, $p = 0.81$
  - not rejecting is not proof of equal weights

Back

The Optimal Shrinkage Intensity

- Ledoit–Wolf choose the target weight δ in $\delta \mathsf{F} + (1-\delta)\mathsf{S}$
  - to minimise the expected squared (Frobenius) distance to the true $\boldsymbol{\Sigma}$

$$\delta^* = \max\left\{0, \min\left\{1, \frac{\hat{\kappa}}{T}\right\}\right\}, \qquad \hat{\kappa} = \frac{\hat{\pi} - \hat{\rho}}{\hat{\gamma}}$$

- $\hat{\pi}$: sum of the asymptotic variances of the entries of S (noise)
- $\hat{\rho}$: their asymptotic covariances with the target entries
- $\hat{\gamma} = \|\mathsf{F} - \mathsf{S}\|^2$: misspecification of the target (not the risk aversion γ)
- Complete formulas for the constant-correlation target: Ledoit & Wolf (2004b), Appendix
  - code: the shrinkage Quantlet

Back

Shrinking the Means and Other Covariance Estimators

- Bayes–Stein (Jorion (1986)): $\hat{\boldsymbol{\mu}}_{\text{BS}} = (1 - \hat{\phi})\,\hat{\boldsymbol{\mu}} + \hat{\phi}\,\hat{\mu}_{\text{GMV}}\mathbf{1}$
  - Empirical-Bayes shrinkage of every mean towards the mean of the GMV portfolio; $\hat{\phi}$ from a Stein argument, larger when N/T is large
- Tu & Zhou (2011): the convex combination of a sophisticated rule and 1/N
  - it minimises the expected out-of-sample loss
- DeMiguel, Garlappi, Nogales & Uppal (2009): norm constraints $\|\mathrm{w}\|_1 \leq \delta$ or $\|\mathrm{w}\|_2 \leq \delta$
  - they nest the long-only GMV, shrinkage GMV and 1/N
  - the bridge to the lasso portfolios of Chapter 13
- Factor structure plus a sparse residual covariance: POET
  - principal orthogonal complement thresholding, Fan, Liao & Mincheva (2013)
  - the link to Chapter 3
- High-dimensional GMV with shrinkage towards a target portfolio: Bodnar, Parolya & Schmid (2018)

Back

Black–Litterman as Mixed (GLS) Estimation

- Theil & Goldberger (1961): stack the prior and the views as one regression on $\boldsymbol{\mu}$

$$\begin{pmatrix} \boldsymbol{\pi} \\ \mathrm{q} \end{pmatrix} = \begin{pmatrix} \mathrm{I} \\ \mathrm{P} \end{pmatrix} \boldsymbol{\mu} + \begin{pmatrix} \mathrm{e}_1 \\ \mathrm{e}_2 \end{pmatrix}, \qquad \operatorname{Var} \begin{pmatrix} \mathrm{e}_1 \\ \mathrm{e}_2 \end{pmatrix} = \begin{pmatrix} \tau\boldsymbol{\Sigma} & 0 \\ 0 & \boldsymbol{\Omega} \end{pmatrix}$$

- GLS gives exactly $\boldsymbol{\mu}_{\mathrm{BL}}$
  - with covariance $\mathrm{M} = \left[(\tau\boldsymbol{\Sigma})^{-1} + \mathrm{P}^{\top}\boldsymbol{\Omega}^{-1}\mathrm{P}\right]^{-1}$
- Equivalent form: $\boldsymbol{\mu}_{\mathrm{BL}} = \boldsymbol{\pi} + \tau\boldsymbol{\Sigma}\mathrm{P}^{\top}(\mathrm{P}\tau\boldsymbol{\Sigma}\mathrm{P}^{\top} + \boldsymbol{\Omega})^{-1}(\mathrm{q} - \mathrm{P}\boldsymbol{\pi})$
- $\boldsymbol{\Omega} \to 0$: restricted least squares, $\mathrm{P}\boldsymbol{\mu}_{\mathrm{BL}} = \mathrm{q}$
- Predictive covariance $\boldsymbol{\Sigma} + \mathrm{M}$ in the weights would change all weights
  - without views it scales them by $1/(1+\tau)$

Back

Other Risk-Based Rules

- **Inverse volatility**: $w_i \propto 1/\sigma_i$
  - equals ERC when all correlations are equal
- **Maximum diversification** (Choueifaty & Coignard (2008)): maximise $\mathrm{w}^\top \boldsymbol{\sigma} / \sqrt{\mathrm{w}^\top \boldsymbol{\Sigma} \mathrm{w}}$
  - the ratio of the average asset volatility to the portfolio volatility
  - $\boldsymbol{\sigma}$: vector of asset volatilities
- **Minimum variance**: GMV, the extreme risk-based rule
- All ignore expected returns: robust to estimation error
  - but optimal only under special assumptions

Back

Results, Sector Universe

Rule	Mean	Volatility	Sharpe	Turnover	Max. drawdown	Sharpe, 50 bp
1/N	9.4%	14.3%	0.66	0.02	−49%	0.65
MV	liquidated in March 2009 (wealth reaches zero); not ranked					
MV-LO	10.7%	14.7%	0.72	0.23	−42%	0.63
GMV	8.5%	11.9%	0.71	0.23	−33%	0.59
GMV-LO	8.1%	11.5%	0.70	0.09	−34%	0.65
GMV-LW	7.6%	11.8%	0.65	0.14	−31%	0.58
ERC	9.3%	13.2%	0.70	0.03	−47%	0.69
HRP	8.9%	12.8%	0.70	0.08	−45%	0.66

- January 2004 – July 2026
  - turnover: monthly mean
  - drawdown of total-return wealth
- Last column: net of 50 basis points per unit of turnover

Back

Results, Multi-Asset Universe (8 ETFs)

Rule	Mean	Volatility	Sharpe	Turnover	Max. drawdown	Sharpe, 50 bp
1/N	4.3%	8.4%	0.52	0.02	−21%	0.50
MV	liquidated in September 2025 (wealth reaches zero); not ranked					
MV-LO	7.1%	8.9%	0.80	0.12	−24%	0.72
GMV	1.4%	3.8%	0.37	0.14	−11%	0.16
GMV-LO	1.7%	5.4%	0.32	0.05	−17%	0.27
GMV-LW	1.1%	4.2%	0.27	0.09	−13%	0.15
ERC	3.6%	7.3%	0.49	0.02	−21%	0.47
HRP	2.4%	6.6%	0.36	0.04	−20%	0.32

- May 2012 – July 2026 (171 months)

Back

Results, Combined Universe (16 ETFs)

Rule	Mean	Volatility	Sharpe	Turnover	Max. drawdown	Sharpe, 50 bp
1/N	7.9%	10.3%	0.76	0.02	−17%	0.75
MV	liquidated in January 2023 (wealth reaches zero); not ranked					
MV-LO	8.0%	9.8%	0.81	0.20	−26%	0.69
GMV	0.8%	4.1%	0.20	0.24	−12%	−0.15
GMV-LO	2.0%	5.4%	0.36	0.06	−17%	0.30
GMV-LW	0.9%	4.4%	0.20	0.13	−14%	0.03
ERC	5.5%	7.8%	0.71	0.03	−18%	0.68
HRP	4.0%	7.0%	0.56	0.11	−18%	0.47

- May 2012 – July 2026

Back

All 18 Tests against 1/N (1/3), Sectors

Rule	$\hat{\Delta}$	HAC SE	p HAC	Holm	p one-sided	p bootstrap	Block
MV-LO	+0.06	0.120	0.60	1.00	0.299	0.63	1
GMV	+0.05	0.159	0.75	1.00	0.377	0.76	1
GMV-LO	+0.04	0.112	0.73	1.00	0.363	0.73	1
GMV-LW	−0.02	0.166	0.93	1.00	0.537	0.94	1
ERC	+0.04	0.023	0.09	1.00	0.046	0.13	1
HRP	+0.04	0.040	0.32	1.00	0.162	0.36	1

- January 2004 – July 2026, $T = 271$
  - $\hat{\Delta}$: annualised Sharpe of the rule minus that of 1/N, before costs
- HAC: Bartlett, 4 lags, two-sided
  - Holm over all 18 two-sided HAC tests
  - one-sided: $1 - \Phi(\hat{\Delta}/\mathrm{SE})$
  - bootstrap: studentized circular block bootstrap, two-sided, block from Algorithm 3.1

Back

All 18 Tests against 1/N (2/3), Multi-Asset

Rule	$\hat{\Delta}$	HAC SE	p HAC	Holm	p one-sided	p bootstrap	Block
MV-LO	$+0.29$	0.184	0.12	1.00	0.059	0.16	10
GMV	-0.15	0.226	0.51	1.00	0.745	0.54	1
GMV-LO	-0.20	0.137	0.14	1.00	0.930	0.15	8
GMV-LW	-0.25	0.191	0.19	1.00	0.903	0.20	4
ERC	-0.03	0.060	0.61	1.00	0.694	0.58	10
HRP	-0.16	0.089	0.07	1.00	0.964	0.05	6

- May 2012 – July 2026, $T = 171$
  - same columns as for the sectors

Back

All 18 Tests against 1/N (3/3), Combined

Rule	$\hat{\Delta}$	HAC SE	p HAC	Holm	p one-sided	p bootstrap	Block
MV-LO	$+0.05$	0.210	0.82	1.00	0.409	0.82	10
GMV	-0.56	0.261	0.03	0.52	0.985	0.05	1
GMV-LO	-0.40	0.203	0.05	0.76	0.976	0.08	6
GMV-LW	-0.56	0.249	0.02	0.43	0.988	0.03	2
ERC	-0.06	0.114	0.61	1.00	0.695	0.62	10
HRP	-0.20	0.152	0.18	1.00	0.908	0.20	8

- May 2012 – July 2026, $T = 171$
- One-sided family: Holm $18 \times 0.046 = 0.82$
  - Benjamini–Hochberg $\min\{18 \times 0.046,\ \frac{18}{2} \times 0.059\} = 0.53$ for the two smallest

Back

Bootstrap Details and the Block Length

- Studentization: $s(\hat{\Delta}^*)$ from the block sums of each resample
  - $s(\hat{\Delta})$ on the data from a prewhitened quadratic-spectral kernel (Andrews & Monahan (1992)) with automatic bandwidth (Andrews (1991))
- Ledoit & Wolf (2008), Algorithm 3.1 (grid and K of Sections 4–5, applied to each of the 18 pairs)
  - choose the block b whose bootstrap interval has actual coverage closest to 95%
  - Fit a first-order vector autoregression to the pair of returns; resample its residuals by a stationary bootstrap (Politis & Romano (1994)), mean block 5
  - Generate $K = 5000$ pseudo-samples of length T; in each, build the 95% interval for $b \in \{1, 2, 4, 6, 8, 10\}$
  - $\hat{g}(b)$: share of intervals that contain the pseudo-true $\hat{\Delta}$; pick $\tilde{b} = \arg\min_b |\hat{g}(b) - 0.95|$
- A non-studentized bootstrap with a fixed 6-month block is not this test
  - its block is not calibrated for coverage

Back

The Deflated Sharpe Ratio (1/2)

- Bailey & López de Prado (2014): the best of N Sharpe ratios is biased upwards even if every true Sharpe ratio is zero
- Expected maximum under H_0: all true Sharpe ratios zero, N independent trials
  - monthly units as T
  - on this slide N counts trials (configurations), not assets
  - $\gamma_E = 0.5772$: Euler's constant; $\hat{V}$: the variance of the N estimates
  - Φ: standard Normal distribution function; e: Euler's number

$$SR_0 = \sqrt{\hat{V}}\left[(1-\gamma_E)\,\Phi^{-1}\left(1-\tfrac{1}{N}\right)+\gamma_E\,\Phi^{-1}\left(1-\tfrac{1}{Ne}\right)\right]$$

- Deflated Sharpe ratio: the probability that the true Sharpe of the winner exceeds SR_0
  - with skewness $\hat{\gamma}_3$ and kurtosis $\hat{\gamma}_4$ of the winner

$$\mathrm{DSR} = \Phi\left(\frac{(\widehat{SR}-SR_0)\sqrt{T-1}}{\sqrt{1-\hat{\gamma}_3\widehat{SR}+\frac{\hat{\gamma}_4-1}{4}\widehat{SR}^2}}\right)$$

Back

The Deflated Sharpe Ratio (2/2)

- This chapter: 21 configurations (7 rules $\times$ 3 universes), aligned on the common calendar May 2012 – July 2026 ($T = 171$); winner MV-LO on sectors, 0.92 a year
  - As 21 independent trials: $SR_0 = 0.48$ a year, DSR $= 0.94 < 0.95$
  - The configurations are strongly correlated (up to 0.99), so fewer are independent: $N = 8$ gives DSR 0.98, $N = 3$ gives 0.99
  - the effective N handles the dependence between configurations, not the autocorrelation of returns
  - the DSR is an illustrative approximation: a calibrated probability only under the formula's assumptions
- Holm adjusts the whole family of pairwise tests against 1/N
  - the DSR asks whether the single best configuration is better than luck

Back

Beyond Equities, Digital Assets

- Petukhina, Trimborn, Härdle & Elendner (2021): cryptocurrencies in portfolio allocation rules
  - gains depend on the rule and on liquidity
- Pele et al. (2023): are cryptos becoming alternative assets?
- Bitcoin: annualised volatility 172%
  - monthly, 2010–2026, one surviving coin chosen ex post
  - for comparison, IEF 6.6%
- Risk contributions grow with volatility
  - risk-based rules give Bitcoin a small weight automatically (Seminar 4, B8)
- Robust mean–variance for crypto-assets: Pele (2023)

Back

Case Study, Method in Detail (1/2)

- **Cost** (eqs. 4, 7)
  - π_t: risky-asset weights; $1 - 1^\top \pi_t$: the weight of the risk-free asset
  - $g_t = \mathrm{diag}\big((1 + r_t^f + r_t)/(1 + g_t^w)\big)$ (eq. 6); r_t^f: risk-free rate, r_t: asset returns, g_t^w: growth of wealth
  - $\mathbf{\Lambda} = \mathrm{diag}(0.2/V_i)$; V_i: average daily dollar volume over 6 months; trading 1% of V_i moves the price by 0.1% (eq. 35, Frazzini, Israel & Moskowitz (2018))
- **Optimal rule** (Proposition 3): $\pi_t = m\, g_t \pi_{t-\mathbf{1}} + (\mathrm{I} - m) A_t$, generalising Gârleanu & Pedersen (2013)
  - A_t: a weighted sum of today's and expected future Markowitz portfolios; m solves eq. (14)
- **Portfolio-ML** (Proposition 4, eq. 40): $A_t = \mathrm{diag}(1/\sigma_{i,t})\, f(s_t)\beta$
  - $\sigma_{i,t}$: volatility of asset i; f: p random Fourier features of the signals s_t
  - $\hat{\beta} = (E_T[\tilde{\mathbf{\Sigma}}_t] + \lambda \mathrm{I})^{-\mathbf{1}} E_T[\tilde{r}_{t+\mathbf{1}}]$: a ridge regression ($\lambda$: penalty) that maximises utility net of costs in closed form (eqs. 24–26)
  - it learns the weights directly, as Brandt, Santa-Clara & Valkanov (2009) do without costs

Back

Case Study, Method in Detail (2/2)

- **Data** (Section 5.1): US common stocks with market capitalisation above the NYSE (New York Stock Exchange) median (184–805 stocks)
  - 115 characteristics of Jensen, Kelly & Pedersen (2023), ranked into $[0, 1]$
  - December 1952 – December 2020, out of sample 1981–2020, monthly
- **Covariance $\mathbf{\Sigma}$**: 13 characteristic themes and 12 industries as factors
  - exponential weights with half-lives of 378 days (correlations) and 126 days (variances)
  - investor $\gamma = 10$, wealth \$10 billion at the end of 2020, growing with the market (Section 5.1.3)
- **Tuning** (Table 1), chosen each year by the utility of validation backtests since 1971
  - $\lambda \in \{0, e^{-\mathbf{10}}, e^{-\mathbf{9.8}}, \ldots, e^{\mathbf{10}}\}$, $p \in \{2^{\mathbf{6}}, \ldots, 2^{\mathbf{9}}\}$, $\eta \in \{e^{-\mathbf{3}}, e^{-\mathbf{2}}\}$; η: standard deviation of the random weights
- **Rivals** ("*" marks an extra tuning layer, Section 5.3)
  - Markowitz-ML (eq. 32), Static-ML (one-period cost, eq. 34), Multiperiod-ML (forecasts for 12 horizons, eq. 22)
  - factor sorts, 1/N, minimum variance, market

Back

Case Study, Complete Results

- **Table 2**: Portfolio-ML turnover 0.25 a month
- Portfolio-ML needs no second tuning layer, unlike Static-ML* and Multiperiod-ML*
- **Table 3**: posterior probability that Portfolio-ML has the higher utility
  - 99.9% against Static-ML* and Multiperiod-ML*, 100% against Markowitz-ML
- **Figure 1(B)** (Section 6.2), $\gamma = 10$: size is a cost
  - with \$10 billion, net excess return 14.6% at 11% volatility; with \$1 billion, 16.3%
- **Figure 5**: net of costs, quality is the most important theme
  - short-term reversal is almost irrelevant
  - monthly autocorrelation -0.04 for 1-month reversal, 0.94 for book-to-market (Section 6.3)
  - persistent signals are cheap to trade
- **Limits stated by the authors**
  - fees and other frictions; computing power and ML methods not available in real time early in the sample (Section 6.1)
  - costs and risk treated as constant over the horizon of each optimisation (Section 5.1.3)
  - among the smallest 20% of stocks, Portfolio-ML ranks behind Multiperiod-ML* and Static-ML* (Figure 8)

Back

Case Study, the Course Replication (1/2)

- **Course data**: 14 US stocks and 9 sector ETFs
  - five signals: 1-month log return, log momentum from $t-6$ and from $t-12$ to $t-1$, 1-year volatility, 6-month dollar volume (formulas in Seminar 4, C3)
  - cost rule, γ, wealth and grid as in the paper; $\boldsymbol{\Sigma}$ with the same exponential weights; validation from 2011
- **Portfolio-ML**: cost 0.6% a year, leverage 0.32
- Net utility: Static-ML -0.13, minimum variance -0.72
  - Markowitz-ML extremely negative (leverage 18)
- **Probability** of a higher utility for Portfolio-ML: a Normal i.i.d. approximation, conditional on the estimated procedure, not an exact posterior probability
  - 99.5% against Static-ML, 99.9% against minimum variance, 93.4% against 1/N
- Annual utility difference with a 95% HAC interval
  - against 1/N: 0.28 $[-0.18, 0.73]$; against Static-ML: 0.15 $[-0.001, 0.30]$; against minimum variance: 0.74 $[0.25, 1.23]$

Back

Case Study, the Course Replication (2/2)

- Same ranking of cost-aware against cost-agnostic rules as Table 2
  - Portfolio-ML wins with small, slowly traded positions, not with a higher gross Sharpe ratio (0.68 against 0.87 for 1/N)
- **Limits**
  - 23 assets instead of 184–805; five signals instead of 115
  - no factor model for $\boldsymbol{\Sigma}$, whose window reaches 10 years only in August 2014 (common prices from August 2004)
  - return forecasts with $p \geq 2^6$ and three validation years instead of ten; no starred rivals
- **Open extension** (Seminar 4, C3)
  - implementable frontiers for wealth of \$1, 10 and 100 billion and $\gamma \in \{1, 5, 10, 20, 100\}$
  - 1/N's trading costs computed at each wealth level

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 peer-reviewed studies since 2015 comparing shrinkage or risk-based rules with 1/N out of sample; for each give DOI, N, T, cost assumption and test."*
  - Resolve every DOI on Crossref and open the table cited
- **Hypotheses**
  - *"Give three competing explanations for a Sharpe gain over 1/N (better $\boldsymbol{\Sigma}$, a low-volatility premium, luck) and the testable implication of each for turnover, subperiods and factor loadings."*
- **Data and code**
  - *"Write Python that reads the course's daily price files, aligns prices on common dates and runs a 60-month rolling GMV-LW and 1/N; weights at t may use data up to $t-1$ only."*
  - Replicate first: 1/N 0.76, GMV-LW 0.20 on the combined universe

Back

The Discovery Loop, Reusable Prompts (2/2)

- **Robustness**
  - *"Draft a pre-registration: the family of tests (rules × universes × windows × costs), the error rate (FWER or FDR), the subperiods and a placebo, fixed before any result is seen."*
- **Critique**
  - *"Act as a referee: give the three strongest reasons my Sharpe gain is spurious and the test that would expose each."*
- **Report**: every specification tried, adjusted p-values, the power of the design, code and seeds
  - any specification chosen after seeing results is exploratory

Back

Answers to the Quick Checks

- **Three rules**: $s_{11} = 0.09$, $s_{22} = 0.0225$, $s_{12} = 0.2 \times 0.30 \times 0.15 = 0.009$
  - GMV: $w_1 = (0.0225 - 0.009)/(0.09 + 0.0225 - 0.018) = 1/7$, volatility 14.3%; ERC (inverse volatility): $w_1 = 1/3$, volatility 15.5%; 1/N: 18.1%
  - 1/N: $w_1(\Sigma w)_1 = 0.5 \times 0.0495 = 0.02475$ of $\sigma_p^2 = 0.0326$: asset 1 carries 76% of the risk with 50% of the weight
- **Plug-in Sharpe**: monthly $\theta^2 = 0.25/12 = 0.0208$
  - $(120 \times 0.0208 + 4)/(120 - 6) = 0.0570$; $\sqrt{12 \times 0.0570} = 0.83$ a year, against a true 0.50
- **Noise band**: $c = 10/40 = 0.25$, band $[(1 - 0.5)^2, (1 + 0.5)^2] = [0.25, 2.25]$
  - 2.0 lies inside, so this comparison cannot separate it from noise
- **One view**: equal precisions give view weight $0.0009/(0.0009 + 0.0009) = 0.5$, posterior 7%
  - with $\Omega = 0.0003$, weight $0.0009/0.0012 = 0.75$, posterior $4\% + 0.75 \times 6\% = 8.5\%$
- **One Sharpe ratio**: SE $= \sqrt{(1 + 0.15^2/2)/120} = 0.092$
  - annual $0.52 \pm 1.96 \times 0.32$, interval $[-0.10, 1.14]$: not significant
- **Four tests**
  - Holm 0.016, 0.060, 0.060, 0.300: one rejection
  - Benjamini–Hochberg 0.016, 0.040, 0.040, 0.300: three rejections

Back

Courses, Code and Data

Advanced courses on quantinar.com

- Hierarchical Clustering
- Minimum Spanning Tree
- Multivariate Statistical Analysis

Code and data

- All charts and tables: github.com/danpele/MFM/tree/main/Quantlets/Ch_04
- Risk-free rate: Kenneth French Data Library

Back

References (1/6)

- Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3), 817–858.
- Andrews, D.W.K., Monahan, J.C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, 60(4), 953–966.
- Bailey, D.H., López de Prado, M. (2014). The deflated Sharpe ratio: correcting for selection bias, backtest overfitting, and non-normality. *Journal of Portfolio Management*, 40(5), 94–107.
- Barroso, P., Detzel, A. (2021). Do limits to arbitrage explain the benefits of volatility-managed portfolios?. *Journal of Financial Economics*, 140(3), 744–767.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.
- Best, M.J., Grauer, R.R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies*, 4(2), 315–342.
- Black, F., Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5), 28–43.
- Bodnar, T., Parolya, N., Schmid, W. (2018). Estimation of the global minimum variance portfolio in high dimensions. *European Journal of Operational Research*, 266(1), 371–390.
- Brandt, M.W., Santa-Clara, P., Valkanov, R. (2009). Parametric portfolio policies: exploiting characteristics in the cross-section of equity returns. *Review of Financial Studies*, 22(9), 3411–3447.
- Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance*, 54(2), 655–671.
- Chopra, V.K., Ziemba, W.T. (1993). The effect of errors in means, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 19(2), 6–11.

References (2/6)

- Cederburg, S., O'Doherty, M.S., Wang, F., Yan, X.S. (2020). On the performance of volatility-managed portfolios. *Journal of Financial Economics*, 138(1), 95–117.
- Choueifaty, Y., Coignard, Y. (2008). Toward maximum diversification. *Journal of Portfolio Management*, 35(1), 40–51.
- DeMiguel, V., Garlappi, L., Nogales, F.J., Uppal, R. (2009). A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Management Science*, 55(5), 798–812.
- DeMiguel, V., Garlappi, L., Uppal, R. (2009). Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy?. *Review of Financial Studies*, 22(5), 1915–1953.
- DeMiguel, V., Martín-Utrera, A., Uppal, R. (2024). A multifactor perspective on volatility-managed portfolios. *Journal of Finance*, 79(6), 3859–3891.
- Dragotă, M., Şerbănescu, C.I., Pele, D.T. (2008). Portfolio diversification and market share analysis for Romanian insurance companies. In *Risk Management and Value*, World Scientific, 277–298.
- Fan, J., Liao, Y., Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B*, 75(4), 603–680.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- Frazzini, A., Israel, R., Moskowitz, T.J. (2018). Trading costs. SSRN Working Paper 3229719.
- Frost, P.A., Savarino, J.E. (1988). For better performance: constrain portfolio weights. *Journal of Portfolio Management*, 15(1), 29–34.
- Gârleanu, N., Pedersen, L.H. (2013). Dynamic trading with predictable returns and transaction costs. *Journal of Finance*, 68(6), 2309–2340.
- Härdle, W.K., Simar, L. (2019). Applied Multivariate Statistical Analysis, 5th ed. Springer.

References (3/6)

- Harvey, C.R., Liu, Y. (2015). Backtesting. *Journal of Portfolio Management*, 42(1), 13–28.
- He, G., Litterman, R. (2002). The intuition behind Black-Litterman model portfolios. SSRN Working Paper 334304.
- Jagannathan, R., Ma, T. (2003). Risk reduction in large portfolios: why imposing the wrong constraints helps. *Journal of Finance*, 58(4), 1651–1683.
- Jensen, T.I., Kelly, B., Malamud, S., Pedersen, L.H. (2026). Machine learning and the implementable efficient frontier. *Review of Financial Studies*, 39(10), 3035–3078.
- Jensen, T.I., Kelly, B., Pedersen, L.H. (2023). Is there a replication crisis in finance?. *Journal of Finance*, 78(5), 2465–2518.
- Jobson, J.D., Korkie, B.M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4), 889–908.
- Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, 21(3), 279–292.
- Kan, R., Wang, X., Zhou, G. (2022). Optimal portfolio choice with estimation risk: no risk-free asset case. *Management Science*, 68(3), 2047–2068.
- Kan, R., Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, 42(3), 621–656.

References (4/6)

- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Laloux, L., Cizeau, P., Bouchaud, J.-P., Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical Review Letters*, 83(7), 1467–1470.
- Ledoit, O., Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10(5), 603–621.
- Ledoit, O., Wolf, M. (2004a). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365–411.
- Ledoit, O., Wolf, M. (2004b). Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management*, 30(4), 110–119.
- Ledoit, O., Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850–859.
- Ledoit, O., Wolf, M. (2017). Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. *Review of Financial Studies*, 30(12), 4349–4388.
- Ledoit, O., Wolf, M. (2020). Analytical nonlinear shrinkage of large-dimensional covariance matrices. *Annals of Statistics*, 48(5), 3043–3065.
- Lo, A.W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36–52.

References (5/6)

- López de Prado, M. (2016). Building diversified portfolios that outperform out of sample. *Journal of Portfolio Management*, 42(4), 59–69.
- Maillard, S., Roncalli, T., Teïletche, J. (2010). The properties of equally weighted risk contribution portfolios. *Journal of Portfolio Management*, 36(4), 60–70.
- Mantegna, R.N. (1999). Hierarchical structure in financial markets. *European Physical Journal B*, 11(1), 193–197.
- Marčenko, V.A., Pastur, L.A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4), 457–483.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- Merton, R.C. (1972). An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis*, 7(4), 1851–1872.
- Merton, R.C. (1980). On estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics*, 8(4), 323–361.
- Michaud, R.O. (1989). The Markowitz optimization enigma: is 'optimized' optimal?. *Financial Analysts Journal*, 45(1), 31–42.
- Moreira, A., Muir, T. (2017). Volatility-managed portfolios. *Journal of Finance*, 72(4), 1611–1644.
- Okhrin, Y., Schmid, W. (2006). Distributional properties of portfolio weights. *Journal of Econometrics*, 134(1), 235–256.

References (6/6)

- Pele, D.T. (2023). Robust Markowitz for crypto-assets. Quantinar course, Bucharest University of Economic Studies.
- Pele, D.T., Wesselhöfft, N., Härdle, W.K., Kolossiatis, M., Yatracos, Y.G. (2023). Are cryptos becoming alternative assets?. *The European Journal of Finance*, 29(10), 1064–1105.
- Petukhina, A., Trimborn, S., Härdle, W.K., Elendner, H. (2021). Investing with cryptocurrencies – evaluating their potential for portfolio allocation strategies. *Quantitative Finance*, 21(11), 1825–1853.
- Politis, D.N., Romano, J.P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313.
- Sharpe, W.F. (1966). Mutual fund performance. *Journal of Business*, 39(S1), 119–138.
- Theil, H., Goldberger, A.S. (1961). On pure and mixed statistical estimation in economics. *International Economic Review*, 2(1), 65–78.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25(2), 65–86.
- Tu, J., Zhou, G. (2011). Markowitz meets Talmud: a combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, 99(1), 204–215.
- Data: Kenneth R. French Data Library, mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Acronyms used in this material (1/3)

- **AI**: Artificial Intelligence
- **BET**: Bucharest Exchange Trading (index)
- **BET-TR**: BET Total Return (index)
- **BVB**: Bursa de Valori București — *the Bucharest Stock Exchange*
- **CAPM**: Capital Asset Pricing Model
- **DSR**: Deflated Sharpe Ratio
- **ERC**: Equal Risk Contribution portfolio
- **ETF**: Exchange-Traded Fund
- **FDR**: False Discovery Rate
- **FWER**: Family-Wise Error Rate
- **GLS**: Generalised Least Squares
- **GMV**: Global Minimum-Variance portfolio
- **GMV-LO**: Global Minimum-Variance portfolio, Long-Only
- **GMV-LW**: Global Minimum-Variance portfolio with Ledoit–Wolf covariance
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **HRP**: Hierarchical Risk Parity
- **HYG**: iShares iBoxx High Yield Corporate Bond ETF
- **IEF**: iShares 7–10 Year Treasury Bond ETF
- **LO**: Long-Only, no short sales
- **LQD**: iShares iBoxx Investment Grade Corporate Bond ETF
- **ML**: Machine Learning
- **MV**: Mean–Variance tangency portfolio estimated from sample moments
- **MV-LO**: Mean–Variance tangency portfolio, Long-Only
- **NYSE**: New York Stock Exchange
- **OLS**: Ordinary Least Squares
- **OMV**: Österreichische Mineralölverwaltung — *Austrian Mineral Oil Administration (OMV group)*
- **OOS**: Out-Of-Sample — *evaluation on data after the estimation period*
- **POET**: Principal Orthogonal complEment Thresholding
- **US**: United States

Acronyms used in this material (2/3)

- **BH**: Benjamini–Hochberg (procedure controlling the FDR)
- **CFW**: Conditional Fixed-Weight multifactor portfolio
- **CMV**: Conditional Mean–Variance multifactor portfolio
- **DOI**: Digital Object Identifier
- **i.i.d.**: independent and identically distributed
- **LW**: Ledoit–Wolf (shrinkage covariance estimator)
- **QS**: Quadratic Spectral (kernel)
- **RON**: Romanian leu
- **SE**: Standard Error
- **SR**: Sharpe Ratio
- **UMV**: Unconditional Mean–Variance multifactor portfolio
- **USD**: US dollar

Acronyms used in this material (3/3): asset tickers

- **SPY**: SPDR S&P 500 ETF Trust
- **EFA**: iShares MSCI EAFE ETF (developed markets outside the US and Canada)
- **EEM**: iShares MSCI Emerging Markets ETF
- **TLT**: iShares 20+ Year Treasury Bond ETF
- **GLD**: SPDR Gold Shares
- **XLB**, **XLE**, **XLF**, **XLI**, **XLK**: Select Sector SPDR (materials, energy, financials, industrials, technology)
- **XLP**, **XLU**, **XLV**, **XLY**: Select Sector SPDR (consumer staples, utilities, health care, consumer discretionary)
- **TLV**: Banca Transilvania
- **SNP**: OMV Petrom
- **BRD**: BRD–Groupe Société Générale
- **TGN**: Transgaz
- **SNG**: Romgaz
- **SNN**: Nuclearelectrica
- **EL**: Electrica
- **TEL**: Transelectrica