








Modelling Financial Markets

Chapter 3: Factor Models and Asset Pricing



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

Why Do Assets Have Different Average Returns?



Eugene F. Fama. Photo: Bengt Nyman (2013); CC BY 2.0; Wikimedia Commons

"[W]hen the tests allow for variation in β that is unrelated to size, the relation between market β and average return is flat, even when β is the only explanatory variable."
— *Fama & French (1992)*

- The chapter's question: why do some assets earn more on average than others?
  - 25 US (United States) portfolios sorted on size and B/M (book value / market value), Jul 1963 – Jul 2026
  - annualised mean excess returns: 3.9% for small growth firms, 13.3% for small value firms
  - does beta explain this gap of about 9.5 percentage points?
- The classic answer, the CAPM (Capital Asset Pricing Model): the difference is payment for market risk, measured by beta
- The thread of the chapter
  - the CAPM says that once beta is accounted for, nothing is left: alpha is zero
  - we estimate beta and alpha, then test on the data whether alpha is zero
  - when it is not, we add factors and check whether those factors survive

▸ Central chart: the Security Market Line (SML) in the Data

What This Chapter Covers

Route through the chapter

1. What Explains Expected Returns?
2. From Markowitz to the CAPM
3. Estimating Beta
4. Testing the CAPM
5. Multifactor Models
6. Fama–MacBeth Regressions
7. Multiple Testing and the Replication of Factors
8. Factor Investing
9. Conclusions

What you should be able to do

- compute a beta with heteroskedasticity- and autocorrelation-robust errors and shrink it
- compute the joint test of the CAPM alphas and risk premia with robust errors
- interpret a new factor against multiple-testing hurdles and replication evidence
- interpret a factor ETF (Exchange-Traded Fund) as factor exposures and alpha

- Additional subjects: Further Topics; answers to the quick checks: Appendix

What Is a Factor Model?

- A **factor model** explains the return of asset i by a few common variables $f_{1,t}, \ldots, f_{K,t}$, the **factors**

$$R^e_{i,t} = R_{i,t} - R_{f,t} = \alpha_i + \beta_{i,1} f_{1,t} + \cdots + \beta_{i,K} f_{K,t} + \varepsilon_{i,t}$$

- $R^e_{i,t}$: excess return of asset i in period t; $R_{f,t}$: risk-free rate
- K: number of factors
- $\beta_{i,k}$: **factor loading** (exposure) of asset i to factor k
- α_i: the **average** deviation of the excess return from the part explained by the factors
- $\varepsilon_{i,t}$: the asset-specific (idiosyncratic) deviation in **period** t, with zero mean
- If the factors are excess returns of traded portfolios, a model that prices the asset correctly implies $\alpha_i = 0$

What Is a Factor Model Used For?

- The chapter's idea: an average return is **payment for risk**
  - investors demand a premium for common risks: the factors (market, size, value ...)
  - holding the other exposures fixed, a larger exposure β to a factor raises the average return if that factor's premium is positive
  - the pricing restriction is $\mathbb{E}[R_i^e] = \boldsymbol{\beta}_i^\top \boldsymbol{\lambda}$, with $\boldsymbol{\lambda}$ the factor premia; what the exposures leave unexplained on average is α
- Two uses in this chapter
  - Explain an asset: for XLK (Technology Select Sector SPDR Fund), the technology ETF, a one-percentage-point rise in the US market excess return goes with an estimated rise of 1.28 points (Section 3)
  - Evaluate a fund: is its α_i positive once its exposures are priced (Sections 4 and 8)?
- The data and the return definitions are in Further Topics

Which Portfolio Does a Mean–Variance Investor Hold?

- The question of this section: if investors choose only by mean and variance, which portfolio do they hold?
- The second question: what mean return does each asset then earn?
- Step 1: the Markowitz (1952) frontier gives the portfolio with the highest Sharpe ratio
- Step 2: the CAPM (Sharpe (1964); Lintner (1965)) says that this portfolio is the market portfolio itself
- We apply both to the nine long-history US sector ETFs (Select Sector SPDR), available since December 1998
  - they are investable, and their series run without a break from 1999 to 2026
  - the GICS (Global Industry Classification Standard) reclassifications of 2016 (real estate) and 2018 (communication services) moved companies out of these nine funds
  - the nine therefore no longer cover the whole S&P 500 (Standard & Poor's 500), the index of 500 large US companies

The Mean–Variance Frontier

- Markowitz (1952): with weights w, expected excess returns $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, the frontier solves

$$\min_{\mathrm{w}} \ \mathrm{w}^{\top}\boldsymbol{\Sigma}\,\mathrm{w} \quad \text{subject to} \quad \mathrm{w}^{\top}\boldsymbol{\mu} = \mu^{*},\ \mathrm{w}^{\top}1 = 1$$

- Negative weights are **short positions**
- The solution: $\mathrm{w}^{*} = \boldsymbol{\Sigma}^{-1}(\eta_1\boldsymbol{\mu} + \eta_2 1)$
  - η_1, η_2: Lagrange multipliers set by the two constraints
- The weights w^{*} are linear in μ^{*}
  - every frontier portfolio is a combination of two frontier portfolios (two-fund separation)
- In the (σ_p, μ_p) plane the frontier is a hyperbola
  - the branch above the minimum-variance portfolio is the **efficient frontier**

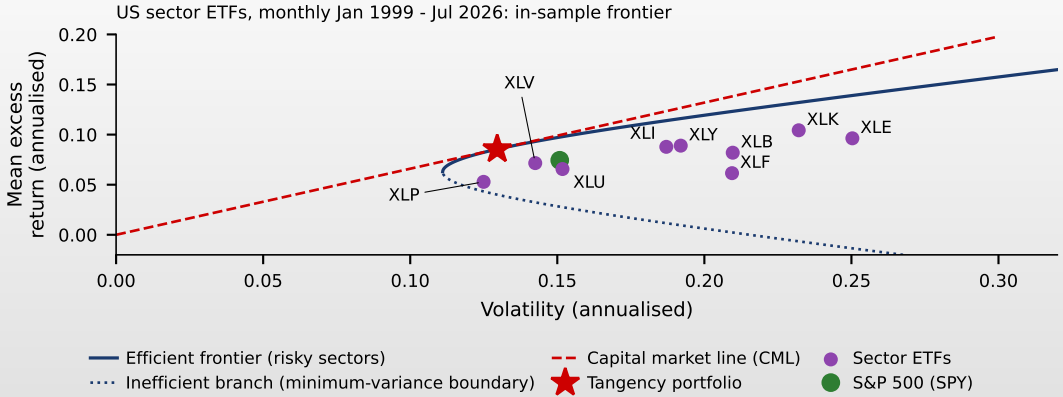
The Sharpe Ratio and the Tangency Portfolio

- **Sharpe ratio** $SR_p = \mathbb{E}[R_p - R_f]/\sigma_p$: the excess return per unit of volatility
- **Tangency portfolio**: the portfolio with the highest Sharpe ratio

$$w_{\text{tan}} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \,/\, 1^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

  - the formula holds when $\boldsymbol{\Sigma}$ is positive definite and the denominator is positive
  - R_p, σ_p: return and volatility of portfolio p
  - the denominator rescales $\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ so that the weights sum to 1
- **CML** (Capital Market Line): all mixes of the risk-free asset and the tangency portfolio, $\mu_p = SR_{\text{tan}}\,\sigma_p$
- If the market is the tangency portfolio, every asset's mean return is linear in its beta
  - Seminar 3, problem A2

The Efficient Frontier of US Sectors (1/2)



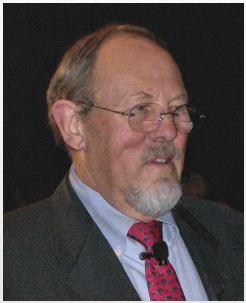
MFM_ch3_frontier ▶ Interactive

The Efficient Frontier of US Sectors (2/2)

- Chart axes: annualised volatility and mean excess return, as fractions ($0.10 = 10\%$)
  - 9 sector ETFs, monthly, January 1999 – July 2026 ($T = 331$ months)
  - the frontier is estimated in-sample
- Tangency portfolio (red star): mean 8.55%, volatility 12.96%, Sharpe ratio 0.66
- SPY (green circle): mean 7.44%, volatility 15.08%, Sharpe ratio 0.49
  - SPY: the ETF that tracks the S&P 500, used as the market proxy
- The CAPM (next slides) requires the market to be the tangency portfolio itself
  - in this sample SPY lies below the frontier
- Sharpe ratio 0.66 against 0.49 in this sample: could the gap come from estimation error?
  - the tangency portfolio is optimised on the same data on which it is evaluated
  - the formal test: Section 4
- The tangency portfolio holds large short positions
  - -0.37 in financials (XLF) and -0.21 in materials (XLB)
  - what each sector holds and the full weights: Further Topics
- The extreme weights arise because the means are estimated with noise (Chapter 4)

MFM_ch3_frontier

The CAPM: Sharpe, Lintner, Black



William F. Sharpe, Nobel Prize in Economics 1990, with Markowitz and Miller

Photo: Larry D. Moore (2007); CC BY 4.0; Wikimedia Commons

- The **CAPM** was formulated by Sharpe (1964) and Lintner (1965)
- All investors hold the same tangency portfolio
  - if they all choose mean–variance efficient portfolios and share the same beliefs
- Since every asset must be held by someone, the tangency portfolio is the **market portfolio**
- Expected excess return is proportional to **beta**

$$\mathbb{E}[R_i] - R_f = \beta_i \left(\mathbb{E}[R_m] - R_f\right), \qquad \beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\mathrm{Var}(R_m)}$$

- R_i, R_m: returns of asset i and of the market portfolio
- $\mathbb{E}[R_m] - R_f$: the **market risk premium**
- β_i: sensitivity of asset i to the market
  - $\beta_i > 1$ amplifies market moves
- Black (1972): a version without unlimited borrowing at the risk-free rate
  - the zero-beta CAPM (Further Topics: a flat SML)

What the CAPM Implies

- Assumptions
  - mean–variance investors with identical beliefs, one period
  - unlimited borrowing and lending at R_f
  - no taxes or trading costs
- Only beta is priced: at the same beta, the model predicts the same expected return, whatever the asset's other characteristics
- Idiosyncratic risk earns nothing, because diversification removes it
- **SML** (Security Market Line): every asset lies on the line (beta, mean)
  - its slope equals the market premium
  - it is the testable statement of the CAPM (Section 4)
- Unlike the SML, the CML contains only efficient portfolios
- Which portfolio is "the market"?
  - the answer decides what a test of the CAPM can establish (Roll's critique, Section 4)

From the CAPM to a Regression: What Alpha Is

- The CAPM is a statement about means; written in excess returns, the relation has no intercept

$$\mathbb{E}[R_i^e] = \beta_i \, \mathbb{E}[R_m^e], \qquad R^e = R - R_f$$

- The data are monthly returns, so we estimate a regression, the **market model**

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \left(R_{m,t} - R_{f,t}\right) + \varepsilon_{i,t}$$

  - **alpha**, α_i: the average excess return that beta does not explain
  - on average, $\alpha_i = \mathbb{E}[R_i^e] - \beta_i \, \mathbb{E}[R_m^e]$: the asset's vertical distance from the SML

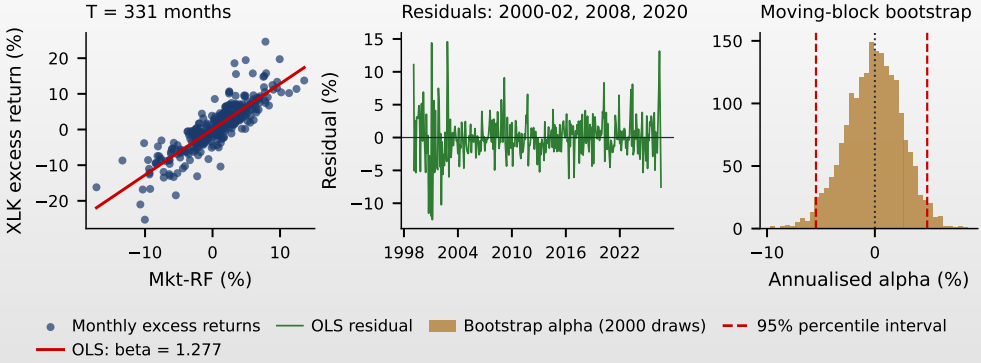
- Everyone wants alpha
  - what performance is left after adjusting for market risk? Of two funds with the same beta, the one with the higher mean return has the higher alpha
  - market exposure can be bought cheaply through an index fund; active managers are paid for alpha

- The CAPM says that in equilibrium $\alpha_i = 0$ for every asset
  - Section 3: we estimate β_i and α_i
  - Section 4: we test whether all alphas are zero

How Precisely Can We Measure Beta?

- The question of this section
  - by how much does an asset move when the market moves by 1%?
  - how sure is that number?
- Beta enters the cost of capital and hedging
- The running example: XLK, the technology ETF, the sector with the highest beta in the sample
- The steps
  - robust standard errors and the choices that change beta
  - rolling betas, shrinkage, the BVB (Bursa de Valori București, the Bucharest Stock Exchange) application

The Beta of the Technology ETF (XLK) in the Data



XLK on Mkt-RF (the US market excess return), monthly excess returns, Jan 1999 – Jul 2026. Left: the OLS (Ordinary Least Squares) line, $\hat{\beta} = 1.277$; middle: the OLS residuals, with spikes in 2000–2002, 2008 and 2020; right: the bootstrap distribution of alpha (details in Further Topics)

MFM_ch3_seminar ▶ Interactive

The Market Model and HAC Errors

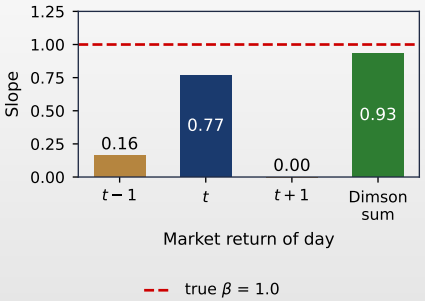
- Market model $R^e_{i,t} = \alpha_i + \beta_i R^e_{m,t} + \varepsilon_{i,t}$
- **OLS** (Ordinary Least Squares): $\hat{\beta}_i = \widehat{\mathrm{Cov}}(R^e_i, R^e_m)/\widehat{\mathrm{Var}}(R^e_m)$
- Variance split $\sigma_i^2 = \beta_i^2 \sigma_m^2 + \sigma^2_{\varepsilon,i}$
  - systematic risk (priced under the CAPM) plus idiosyncratic risk (diversifiable)
  - σ_i^2, σ_m^2: variances of the asset and of the market; $\sigma^2_{\varepsilon,i}$: residual variance
  - $R^2 = \beta_i^2 \sigma_m^2 / \sigma_i^2$ is the systematic share
- Why **HAC** (heteroskedasticity and autocorrelation consistent) errors (Newey & West (1987))
  - volatility clustering (Chapter 1) makes the residuals heteroskedastic
  - the product $R^e_{m,t}\varepsilon_{i,t}$ can be autocorrelated
- HAC changes the estimated covariance matrix of the coefficients, not the OLS estimate
- XLK on Mkt-RF, monthly 1999–2026: $\hat{\beta} = 1.277$ in all three cases
  - standard error: 0.041 (classical), 0.054 (White, heteroskedasticity only), 0.082 (Newey–West, 5 lags)
  - weights, number of lags and assumptions: Further Topics

Choices That Change an Estimated Beta (1/2)

- Frequency and window: usually 3–5 years of monthly data or 1–2 years of daily data
  - a longer window lowers the standard error but mixes different beta regimes (Betas Are Not Constant)
- Market proxy: SPY (S&P 500) or Mkt-RF for US assets, BET (Bucharest Exchange Trading, index) for Romanian stocks
- Data cleaning: an unadjusted capital event creates a fake return
  - Example: 7 September 2023 is the ex-date of a RON (Romanian leu) 1.7225 per share special dividend of Fondul Proprietatea (FP)
  - On that day the close falls by 81% (a false, price-based return), but the total return of the day, with the dividend, is $+6.0\%$

Choices That Change an Estimated Beta (2/2)

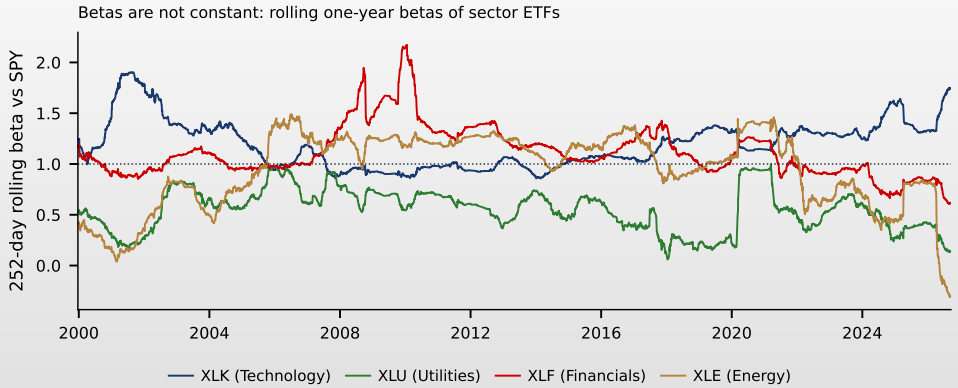
- Asynchronous trading: the closing prices of different assets are not recorded at the same moment
- An illiquid stock reacts to market news with a delay, so the daily OLS beta is biased towards 0
- The Dimson beta (Dimson (1979)): regression on three days of market returns
  - the market return of the previous day, the same day and the next day
  - $R_{i,t} = a + b_{-1} R_{m,t-1} + b_0 R_{m,t} + b_{+1} R_{m,t+1} + \varepsilon_{i,t}$
  - the slopes are added: $\beta^D = b_{-1} + b_0 + b_{+1}$
- The chart shows a simulation
  - a stock with true beta 1 trades on only 75% of days
- The slope on the same-day market return is 0.77, so it understates beta
- The slope on the previous-day market return is 0.16: part of the reaction arrives one day late
- The Dimson beta is 0.93, much closer to 1



▶ Interactive

‣ Further Topics: XLK example

Betas Are Not Constant (1/2)



MFM_ch3_rolling_beta ▶ Interactive

Betas Are Not Constant (2/2)

- The chart: OLS beta against SPY over rolling windows of 252 daily log returns (about one year)
- The beta of financials (XLF) ranged from 0.60 to 2.17
- The XLF beta peaked right after the 2008–2009 banking crisis
- The beta of technology (XLK) ranged from 0.86 to 1.90
- The XLK beta was high in the dot-com bubble (2000–2001)
  - it is 1.75 in the last 252-day window
- The beta of energy (XLE) ranged from -0.31 to 1.49
- The XLE beta is negative in the last window: the covariance of XLE returns with SPY turned negative
- The beta of utilities (XLU), the classic defensive sector, ranged from 0.06 to 1.00
- A single full-sample beta hides these regimes
- Models with time-varying betas are covered in Chapters 5–6

MFM_ch3_rolling_beta

Shrinking Betas: Blume and Vasicek (1/2)

- **Regression to the mean**: a very high or very low estimated beta is partly noise
  - so in the next period it tends to move towards the cross-sectional mean, which is close to 1
- Blume (1971): regression of betas in one period on betas in the previous period
- Hence the practitioner rule: $\beta^{\text{Blume}} = 0.33 + 0.67\,\hat{\beta}$
- Vasicek (1973): a weighted average of the asset's estimate and a prior

$$\beta_i^V = w_i \hat{\beta}_i + (1 - w_i)\bar{\beta}, \qquad w_i = \frac{\sigma_\beta^2}{\sigma_\beta^2 + \text{se}(\hat{\beta}_i)^2}$$

  - the prior: the mean $\bar{\beta}$ of the betas of all assets
  - w_i: the weight of the estimate, larger when $\hat{\beta}_i$ is precise (small SE)
  - $\text{se}(\hat{\beta}_i)$: its standard error
  - σ_β^2: the variance of the **true** betas across assets

Shrinking Betas: Blume and Vasicek (2/2)

- The variance of the estimates across assets adds true variance and noise
  - $\text{Var}_{cs}(\hat{\beta}) = \sigma^2_\beta + \overline{\text{se}^2}$, with $\overline{\text{se}^2}$ the mean squared standard error
- So $\sigma^2_\beta = \text{Var}_{cs}(\hat{\beta}) - \overline{\text{se}^2}$, the empirical-Bayes estimate
  - derivation: Seminar 3, problem A3
- Example: XLK on SPY, monthly 1999–2012: $\hat{\beta} = 1.415$, Newey–West SE 0.139
  - a different benchmark (SPY, not Mkt-RF) and a different period (1999–2012) from the beta of 1.277 on the earlier slides
- Prior from the 9 sectors: $\bar{\beta} = 0.949$ and $\sigma^2_\beta = 0.1159 - 0.0077 = 0.1082$
- The weight on the own estimate: $w = 0.1082/(0.1082 + 0.139^2) = 0.849$
  - it gives the Vasicek beta $\beta^V = 0.849 \times 1.415 + 0.151 \times 0.949 = 1.344$
- Blume beta: $0.33 + 0.67 \times 1.415 = 1.278$
- The beta estimated over the following period, 2013–2026: 1.172
  - both adjustments are closer to it than the unadjusted beta
  - over all 9 sectors, the forecast root mean squared error (RMSE): OLS 0.135, Blume 0.109, Vasicek 0.114

Quick Check: A Beta, Its Error and Its Shrinkage

- A stock, 120 monthly returns: $\hat{\beta} = 0.90$, classical SE 0.05, Newey–West SE 0.09
  1. Give both 95% intervals, with the multiplier 1.96. Does the stock have lower systematic risk than the market (beta below 1) under each?
  2. Which interval do you report, and why does the point estimate not change?
- A cross section of betas has $\mathrm{Var}_{cs}(\hat{\beta}) = 0.10$, $\overline{\mathrm{se}^2} = 0.02$ and mean $\bar{\beta} = 1.0$
  3. Compute σ^2_{β} and the Vasicek beta of an asset with $\hat{\beta} = 1.6$, once with SE 0.2 and once with SE 0.1
  4. Which of the two is shrunk more, and why? Compare both with the Blume beta
- An illiquid BVB stock, daily returns
- Slopes on BET: 0.25 for the previous day, 0.55 for the same day, 0.00 for the next day
  5. Which Dimson beta do you report?
  6. Why is the same-day slope smaller than the Dimson beta?
- Answers: Appendix

Is the CAPM Rejected by the Data?

- The question of this section
  - do mean returns lie on the line $\mathbb{E}[R_i^e] = \beta_i \, \mathbb{E}[R_m^e]$?
  - that is, do all assets have zero alpha?
- Test assets: the 25 size $\times$ B/M portfolios (Fama & French (1993))
- Why these portfolios: they systematically contrast size and B/M, characteristics the literature associates with different mean returns
  - annualised mean excess returns from 3.9% to 13.3%
- The steps
  - the SML in the data
  - the GRS (Gibbons–Ross–Shanken, test) test, which checks all alphas at once
  - the limits of the test

The Test Assets: 25 Size × B/M Portfolios

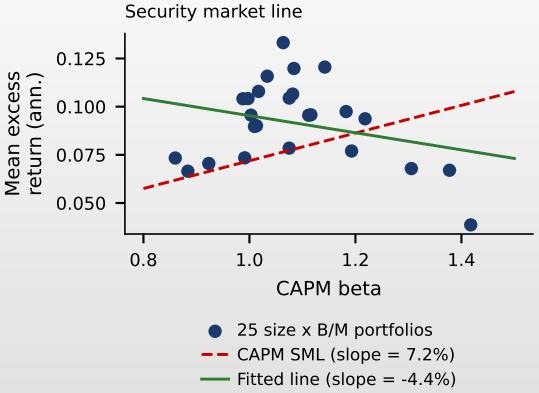
	BM1 (growth)	BM2	BM3	BM4	BM5 (value)
ME1 (small)	3.9	9.4	9.6	11.6	13.3
ME2	6.7	9.7	10.7	10.8	12.1
ME3	6.8	9.6	9.0	10.4	12.0
ME4	7.7	7.8	9.0	10.4	10.4
ME5 (big)	7.3	7.1	7.3	6.7	9.6

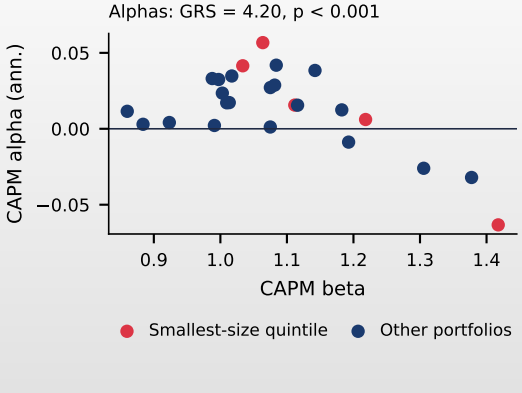
mean annual excess return, %; Jul 1963 – Jul 2026
ME: market equity (capitalisation); BM: B/M quintile

- Every June t, NYSE (New York Stock Exchange), AMEX (American Stock Exchange) and NASDAQ (National Association of Securities Dealers Automated Quotations) stocks are sorted twice, independently (Fama & French (1993))
  - on size: market capitalisation at the end of June t
  - on B/M: book equity for fiscal year $t-1$ divided by market capitalisation in December $t-1$
- The breakpoints are NYSE quintiles (the 20th, 40th, 60th and 80th percentiles)
  - otherwise NASDAQ micro-caps would fill the groups
- Each of the 25 intersections is a value-weighted portfolio, held from July t to June $t+1$
- In every row, the value column (BM5) has a higher mean return than the growth column (BM1); the extremes: small growth (3.9% a year) and small value (13.3%)
- The SMB (Small Minus Big) and HML (High Minus Low) factors (Section 5) use the same characteristics, so they explain these portfolios almost by construction: a high R^2 proves little (Lewellen, Nagel & Shanken (2010))

▶ Interactive

The Security Market Line in the Data (1/2)





MFM_ch3_sml ▶ Interactive

The Security Market Line in the Data (2/2)

- The 25 portfolios, July 1963 – July 2026, $T = 757$
  - re-sorted each June and value-weighted
- The CAPM line has slope 7.2% a year
  - returns and alphas: annual fractions, $0.05 = 5\%$; beta has no unit
- The fitted line has slope -4.4%: higher beta, **lower** average return
  - fitted line: the cross-sectional regression of mean returns on beta
- Betas span only 0.86–1.42
- Alpha is $+5.7\%$ a year for small firms with high B/M
  - and -6.3% a year for small firms with low B/M
- 10 of the 25 alphas have $|t| > 1.96$ (HAC errors)
- In this sample, portfolios with higher betas did not earn higher average returns
- The next question: can all these deviations be consistent with zero?

MFM_ch3_sml

From One Alpha to the GRS Test (1/2)

- **Jensen's alpha** (Jensen (1968)): the market-model intercept α_i
  - The CAPM implies $\alpha_i = 0$ for every asset
- A separate t-test for each asset: XLK's alpha has $t = 0.13$
  - With many assets, some alphas look significant by chance alone (Section 7)
- **The GRS test** (Gibbons, Ross & Shanken (1989)): one test for all assets, $H_0 : \alpha_1 = \cdots = \alpha_N = 0$

$$W = \frac{T - N - 1}{N} \, \frac{\hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}}}{1 + \hat{\mu}_m^2 / \hat{\sigma}_m^2} \;\sim\; F(N,\, T - N - 1)$$

- Notation
  - N: number of test assets; T: number of months; $\hat{\boldsymbol{\alpha}}$: the N estimated alphas
  - $\hat{\boldsymbol{\Sigma}}$: maximum-likelihood residual covariance (divided by T)
  - $\hat{\mu}_m / \hat{\sigma}_m$: Sharpe ratio of the market

From One Alpha to the GRS Test (2/2)

- $\hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}}$: a weighted sum of squared alphas
  - an alpha counts more when the asset's residuals are less volatile
- Decision: reject H_0 when W exceeds the 95% quantile of $F(N, T-N-1)$
- The distribution is exact, conditional on the factors
  - errors i.i.d. multivariate Normal, independent of the factors
  - zero mean and constant nonsingular covariance
- The Sharpe-ratio form (monthly, moments divided by T)

$$\hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} = \widehat{SR}^2(f, \text{assets}) - \widehat{SR}^2(f)$$

  - W is this gain in the squared maximum Sharpe ratio, normalised: a large W means the market is far from the frontier built from the market and the test assets
- For $N = 1$, $W = t^2$: GRS generalises the t-test
  - t: the classical (OLS, not HAC) t-statistic of alpha
  - $F(1, T-2)$ is the distribution of the square of a Student-t variable with $T-2$ degrees of freedom

The GRS Test on the 25 Portfolios

- The 25 size $\times$ B/M portfolios, July 1963 – July 2026 ($T = 757$, $N = 25$)
  - $W = 4.20$, with $p = 8 \times 10^{-11}$
- The CAPM is rejected: the alphas are large relative to their uncertainty
- The rejection concerns the model together with the index used as the market
  - Roll's critique, next slide

Roll's Critique

- Roll (1977): the CAPM is a statement about the **true** market portfolio of all wealth
- The stock indices used as the market omit bonds, real estate, private firms and human capital
- Mathematically, betas computed against **any** efficient portfolio price all mean returns exactly
  - Seminar 3, problem A2
- So a CAPM test really checks whether the index used as the market is efficient
- A rejection can come from the model or from the choice of index

The Choice of Test Assets Changes the Conclusion

- On the 9 sector ETFs of Section 2 the CAPM is not rejected
  - January 1999 – July 2026, $T = 331$, with Mkt-RF as the market
  - GRS $= 0.53$ with $p = 0.85$
- Adding the sectors raises the squared monthly Sharpe ratio by 0.015
  - less than the 0.029 that noise alone would produce (Appendix)
- Over the same 331 months the 25 size $\times$ B/M portfolios reject it
  - GRS $= 2.91$, $p = 9 \times 10^{-6}$
- The sectors are not built to contrast size and B/M
  - forming portfolios by sector imposes no size or B/M contrast comparable with the 5$\times$5 sort
- The power of the test depends on the deviations from the model relative to the residual covariance, on T and on N
- A non-rejection does not confirm the CAPM, nor does it alone prove a lack of power
- Test assets are chosen for economic contrasts set before the test, not for the dispersion of the observed means

Quick Check: Reading a GRS Statistic

- One factor, $N = 4$ test assets, $T = 65$ months
- Monthly Sharpe ratios: factor alone 0.15, tangency portfolio of the factor and the 4 assets 0.25
  1. Compute W from the Sharpe-ratio form. The 95% quantile of $F(4, 60)$ is about 2.53: reject?
  2. Compute the squared Sharpe-ratio gain expected under H_0, $N(1 + \widehat{SR}^2(f))/(T - N - 3)$, and compare it with the observed gain
  3. Same Sharpe ratios, $T = 650$ months: recompute W (95% quantile of $F(4, 645)$: about 2.39). What changed, the gap or the evidence?
- Answers: Appendix

Arbitrage Pricing Theory

- The market beta does not explain mean returns (Section 4)
  - the question of this section: how many factors do we need?
  - and which ones?
- **APT**: Arbitrage Pricing Theory, Ross (1976)
- Assumption: returns follow a K-factor model

$$R_{i,t} = \mathbb{E}[R_i] + \boldsymbol{\beta}_i^\top \mathsf{f}_t + \varepsilon_{i,t}, \qquad \mathbb{E}[\mathsf{f}_t] = 0$$

  - asset-specific risks can be diversified away
  - $\boldsymbol{\beta}_i$: the K loadings of asset i
  - f_t: the unexpected part of the factors
- **No arbitrage** (no sure profit without capital invested) implies $\mathbb{E}[R_i] - R_f = \boldsymbol{\beta}_i^\top \boldsymbol{\lambda}$
- The relation is approximate: it holds for almost all assets, not exactly for each one
- $\boldsymbol{\lambda}$: the **risk premia**, the extra expected return per unit of each exposure
- The theory does not say **which** factors: they must be found in the data

Building Factors: Long–Short Portfolios (1/2)

size: NYSE median; B/M: 30% / 40% / 30%

	Low (growth)	Neutral	High (value)
Small	S_L	S_N	S_H
Big	B_L	B_N	B_H

$$\text{SMB} = \tfrac{1}{3}(S_L + S_N + S_H) - \tfrac{1}{3}(B_L + B_N + B_H)$$

$$\text{HML} = \tfrac{1}{2}(S_H + B_H) - \tfrac{1}{2}(S_L + B_L)$$

- A **characteristic-sorted factor** is the difference between the returns of two portfolios, $f_t = R_{\textbf{long},t} - R_{\textbf{short},t}$
  - the short side finances the long side
- The direction is fixed for each factor
  - HML: high B/M minus low B/M
  - SMB: small minus big capitalisation
- Fama–French: each June, stocks are split into 2×3 groups
  - using the previous year's accounting data
  - each group is a value-weighted portfolio
- The size breakpoint is the median market capitalisation of NYSE (New York Stock Exchange) stocks

Building Factors: Long–Short Portfolios (2/2)

- NYSE breakpoints and value weights stop microcaps from dominating the portfolios
  - **SMB** (Small Minus Big): small minus big stocks
  - **HML** (High Minus Low): high B/M (value) minus low B/M (growth), averaged over small and big firms, hence at the same size
- The same construction on profitability and investment gives RMW and CMA (next slide, which also gives the direction of each)

Factors Added to the Three-Factor Model

- **RMW** (Robust Minus Weak): high minus low operating profitability
- **CMA** (Conservative Minus Aggressive): low minus high asset growth
- **ROE** (Return On Equity): the profitability factor of the q-factor model
- **MOM** (momentum, also UMD, Up Minus Down)
  - stocks with high minus stocks with low returns over months $t-12$ to $t-2$
  - the last month is skipped
- Past winners keep winning for 3–12 months (Jegadeesh & Titman (1993))
- The SMB of the three-factor model (FF3) comes from a single sort (size and B/M)
  - the SMB of the five-factor model (FF5) averages three sorts
  - so FF3 is not a special case of FF5

The Standard Factor Models

Model	Factors	Reference	Motivation
CAPM	Mkt-RF	Sharpe (1964)	equilibrium
FF3	Mkt-RF, SMB, HML	Fama & French (1993)	size and B/M effects (Fama & French (1992))
Carhart	FF3 + MOM	Carhart (1997)	fund evaluation
q-factor	market, size, investment, ROE	Hou, Xue & Zhang (2015)	firms' investment decisions
FF5	Mkt-RF, SMB, HML, RMW, CMA	Fama & French (2015)	dividend-discount model

- FF3 and FF5: the Fama–French three- and five-factor models
  - the FF3 regression is $R^e_{i,t} = \alpha_i + b_i \,\mathrm{MKT}_t + s_i \,\mathrm{SMB}_t + h_i \,\mathrm{HML}_t + \varepsilon_{i,t}$
  - MKT = Mkt-RF; loadings b_i, s_i, h_i
- A good model leaves time-series alphas close to zero
  - GRS tests whether the deviations are large relative to their uncertainty
- The motivation of FF5, the dividend-discount model: Further Topics

Constructing the Momentum Factor

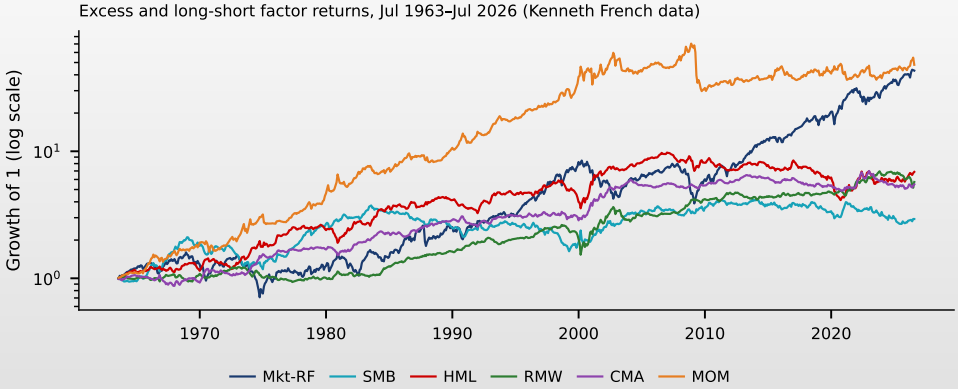
size: NYSE median
return over $t-12 \ldots t-2$: 30% / 40% / 30%

	Losers	Neutral	Winners
Small	S_L	S_N	S_W
Big	B_L	B_N	B_W

$$\text{UMD} = \tfrac{1}{2}(S_W + B_W) - \tfrac{1}{2}(S_L + B_L)$$

- The same 2×3 sort as for SMB and HML, with past return in place of B/M
- Return breakpoints are the 30th and 70th percentiles of NYSE stocks
- The groups are re-formed monthly, not once a year
  - past returns change monthly, accounting data only yearly
- **Formation period**: cumulative return over months $t-12$ to $t-2$, i.e. 11 months
- Month $t-1$ is skipped because of **short-term reversal**: at a one-month horizon, returns reverse (Jegadeesh (1990))
  - causes: bid-ask bounce (the trade price alternates between bid and ask) and liquidity pressure
  - if included, month $t-1$ would mix two effects of opposite sign
- UMD is **cross-sectional** momentum: winners minus losers at the same date
- **Time-series momentum** compares each asset with its own past (Moskowitz, Ooi & Pedersen (2012))
  - long if the 12-month return is positive, short otherwise

Cumulative Returns of the Six Factors, 1963–2026 (1/2)



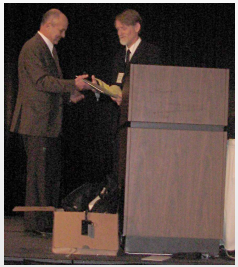
MFM_ch3_factor_cum ▶ Interactive

Cumulative Returns of the Six Factors, 1963–2026 (2/2)

- Monthly factor returns, compounded from 1, on a log scale
- The factors are excess or long–short returns
  - the curves include no interest on capital or on the collateral of short positions
  - so they are not the wealth of a real investor
- Final value of one unit invested, July 1963 – July 2026
  - Mkt-RF 43.2, MOM 48.0, HML 6.9
  - RMW 5.7, CMA 5.5, SMB 2.9
- Momentum peaked before 2009 and lost heavily when the market rebounded sharply in 2009
  - a **momentum crash**
- The value factor was flat or losing from about 2007 to 2020
- Average annual returns
  - Mkt-RF 7.2%, MOM 7.3%, HML 3.6%
  - RMW 3.1%, CMA 3.0%, SMB 2.2%

MFM_ch3_factor_cum

The Two Passes



Eugene Fama, Morgan Stanley–AFA (American Finance Association) Award, 2008

Photo: TonyTheTiger (2008); CC BY-SA 3.0; Wikimedia Commons

- The question: are factor exposures rewarded?
  - that is, do betas explain the differences in mean returns, $\mathbb{E}[R_i^e] = \lambda_{\mathbf{0}} + \boldsymbol{\beta}_i^\top \boldsymbol{\lambda}$?
- Why Fama & MacBeth (1973): for a traded factor, a correct model imposes $\lambda = \mathbb{E}[f]$; the free estimate can differ
  - non-traded factor (e.g. inflation): only the cross-sectional regression gives λ
- **Pass 1** (time series, one regression per asset)
  - $R_{i,t}^e = a_i + \boldsymbol{\beta}_i^\top \mathrm{f}_t + \varepsilon_{i,t}$ gives the matrix $\hat{\mathrm{B}}$ $(N \times K)$
- **Pass 2** (cross section, one regression each month on $\mathrm{X} = [1, \hat{\mathrm{B}}]$)

$$R_{i,t}^e = \lambda_{\mathbf{0},t} + \hat{\boldsymbol{\beta}}_i^\top \boldsymbol{\lambda}_t + u_{i,t}, \qquad \hat{\boldsymbol{\lambda}} = \tfrac{1}{T} \textstyle\sum_t \hat{\boldsymbol{\lambda}}_t$$

- $\lambda_{\mathbf{0}}$: zero if the model is right; in the Black (1972) version, the zero-beta portfolio return minus R_f
- Intuition of pass 2: in a given month, did the portfolios more exposed to HML earn more?
  - that month's slope is the realised premium; the mean of the T slopes, the estimated one
  - the example uses full-sample betas: a description, not a forecast
  - 25 portfolios: HML premium 4.0% a year, factor mean 3.6%; how precise?

Standard Errors for Risk Premia (1/2)

- Four sources of uncertainty, four corrections
  - Fama–MacBeth: the monthly variation of the slopes
  - Shanken: the betas are estimated
  - HAC: dependence over time
  - Kan–Robotti–Shanken (KRS): possible pricing errors (a wrong model)
- **Fama–MacBeth**: the monthly $\hat{\lambda}_{k,t}$ form a sample of size T

$$\text{se}_{FM}(\hat{\lambda}_k) = s(\hat{\lambda}_{k,t})/\sqrt{T}$$

  - $s(\cdot)$: the sample standard deviation over the T months
  - robust to correlation of the residuals across assets in the same month
  - if $\hat{\lambda}_{k,t}$ is autocorrelated, use Newey–West on this series
- **Errors in variables**: pass 2 uses estimated betas, not the true ones, so se_{FM} understates the uncertainty

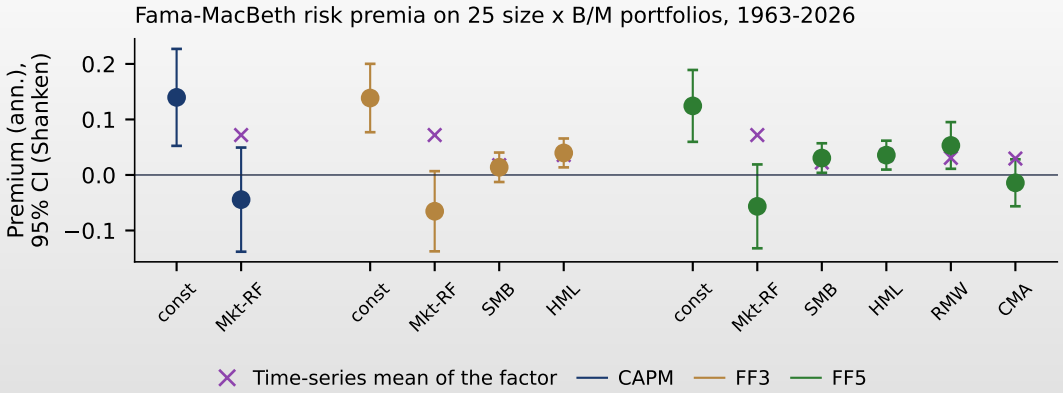
Standard Errors for Risk Premia (2/2)

- The Shanken (1992) correction scales by $(1 + c)$ the part of the variance that does not come from factor variability

$$\mathrm{Var}_{\mathrm{Sh}}(\hat{\boldsymbol{\lambda}}) = (1 + c)\left[\mathrm{Var}_{FM}(\hat{\boldsymbol{\lambda}}) - \tfrac{1}{T}\boldsymbol{\Sigma}_f\right] + \tfrac{1}{T}\boldsymbol{\Sigma}_f, \qquad c = \hat{\boldsymbol{\lambda}}^\top \boldsymbol{\Sigma}_f^{-1} \hat{\boldsymbol{\lambda}}$$

  - Var_{FM}: the Fama–MacBeth variance, $s^2(\hat{\boldsymbol{\lambda}}_t)/T$
  - $\boldsymbol{\Sigma}_f$: factor covariance
  - c: squared Sharpe ratio of the premia, per month (0.03 for FF3)
- Small c (premia small relative to factor volatility, as for monthly FF3): the correction is small
- Large c (for example, a macro factor with low volatility relative to its estimated premium): Fama–MacBeth t-statistics are overstated
- The Fama–MacBeth, Newey–West and Shanken errors assume that the model is correct; the KRS correction for a wrong model: Further Topics

Estimated Risk Premia (1/2)



MFM_ch3_fama_macbeth ▶ Interactive

Estimated Risk Premia (2/2)

- Bars: 95% confidence intervals with Shanken errors, in annual fractions ($0.1 = 10\%$)
- Crosses: the time-series mean of each traded factor, equal to λ if the model is right
  - the comparison with the cross is visual: the mean is also estimated, from the same months
- Estimated premia (25 portfolios, 1963–2026), SE in brackets
  - CAPM: intercept 14.0% (4.4%), market −4.4% (4.8%)
  - FF3: SMB 1.4% (1.4%), HML 4.0% (1.3%)
  - FF5: RMW 5.3% (2.1%), CMA −1.4% (2.2%)
- HML: estimated premium 4.0%, factor mean 3.6%: close
- Market (FF3): estimated premium −6.5%, factor mean 7.2%: far apart
- Market betas differ little across the portfolios, so the slope on them is imprecise
  - 0.86–1.42 in the CAPM and 0.89–1.14 in FF3
- The market premium is therefore **weakly identified**
  - the intercept absorbs the average level of returns (Lewellen, Nagel & Shanken (2010))
- The result shows an association between exposures and mean returns, not the economic source of the risk

MFM_ch3_fama_macbeth

Three Checks of a Factor Model *(deeper)*

- Three checks (Lewellen, Nagel & Shanken (2010))
  - Broaden the test assets (Further Topics)
  - Also report the **GLS** (Generalised Least Squares) R^2: like GRS, it accounts for covariances, but it uses the inverse covariance of the asset returns, not of the residuals
  - Force the premium of each traded factor to equal its mean and drop the intercept: FF3 on the 55 assets (25 portfolios and 30 industries) gives $R^2 = -0.27$ (neither the GLS R^2 nor the free OLS R^2)
- Unweighted cross-sectional R^2: $R^2 = 1 - \sum_i (\bar{R}_i^e - \hat{\mu}_i)^2 / \sum_i (\bar{R}_i^e - \bar{R}^e)^2$
  - $\hat{\mu}_i$: the predicted mean
  - $\bar{R}^e$: the average of the N assets
- $R^2 < 0$ when the model does worse than the plain average $\bar{R}^e$
  - OLS with an intercept cannot do worse than that
- Individual stocks instead of portfolios (Ang, Liu & Schwarz (2020))
  - betas are more dispersed, but estimated with more noise

Quick Check: Premia, Shanken and a Negative R^2

- A non-traded factor (for example, industrial-production growth), volatility $\sigma_f = 1\%$ a month
- $T = 400$ months; pass 2 gives $\hat{\lambda} = 0.5\%$ a month, with $s(\hat{\lambda}_t) = 3\%$
  1. Compute se_{FM} and the Fama–MacBeth t-statistic
  2. Compute c, the Shanken variance and the Shanken t-statistic. Does the factor clear the $t > 3$ hurdle (the exercise's assumption; rationale: Section 7)?
  3. Why would c be tiny for a traded factor with $\sigma_f = 4.5\%$ and the same $\hat{\lambda}$?
- Three portfolios have mean excess returns 2%, 4%, 6% a year
- A model with imposed premia and no intercept fits 6%, 4%, 2%
  4. Compute SST (Total Sum of Squares), SSE (Sum of Squared Errors) and R^2
  5. Why can the R^2 of an OLS regression with an intercept never be negative?
- Answers: ▸ Appendix

The Factor Zoo

- The question: how many of the published factors are real?
  - each was tested as if it were the only one
- Cochrane (2011) called the growing list of factors a *factor zoo*
  - Harvey, Liu & Zhu (2016) catalogue 316 published factors
- A controlled experiment: we test 300 true null hypotheses at 5%
  - on average we expect 15 false rejections
- The inventory of publications (316) is not the tested family, whose size is unknown
- **Data mining** (or p-hacking): searching across many variants until one looks significant
- The Bonferroni correction divides 5% by the number of tests
  - for 316 tests hypothetically fixed in advance: $p \leq 0.00016$, that is $|t| > 3.78$
- Harvey, Liu & Zhu (2016) propose $|t| > 3.0$ for a new factor
  - the hurdle that follows from the multiple-testing corrections (next slide) applied to all factors tested
- The hurdle rises with every factor published: the tested family grows
  - a hurdle that is too low lets false factors through, which then fail to replicate (the replication case study)

Controlling False Discoveries (1/2)

- **FDR** (False Discovery Rate): expected share of false rejections among all rejections
- **FWER** (Family-Wise Error Rate): probability of at least one false rejection (Chapter 2)
- Notation of the table
  - M: number of tests; α: the significance level of the tests (here α is not Jensen's alpha)
  - $p_{(\mathbf{1})} \leq \cdots \leq p_{(M)}$: sorted p-values; $H_{(j)}$: the hypothesis with the j-th smallest p-value
  - **BH**: the Benjamini–Hochberg procedure; **BY**: the Benjamini–Yekutieli procedure

Procedure	**Controls**	**Reject $H_{(j)}$ if**	**Valid under**
Bonferroni	FWER	$p_{(j)} \leq \alpha/M$	any dependence
Holm (Holm (1979))	FWER	$p_{(k)} \leq \alpha/(M-k+1)$ for all $k \leq j$	any dependence
BH (Benjamini & Hochberg (1995))	FDR	$j \leq$ largest k with $p_{(k)} \leq k\alpha/M$	positive dependence
BY (Benjamini & Yekutieli (2001))	FDR	as BH, thresholds divided by $\sum_{k=\mathbf{1}}^{M} 1/k$	any dependence

Controlling False Discoveries (2/2)

- Positive dependence: the test statistics tend to move up together
  - independence is a special case
  - the technical condition is called PRDS, positive regression dependence on a subset (Benjamini & Yekutieli (2001))
- FWER: suitable when no false discovery is acceptable
- FDR: suitable when searching over many candidate factors
  - a small share of false discoveries is accepted

Quick Check: Counting Discoveries

- Five candidate factors, sorted p-values 0.004, 0.012, 0.020, 0.030, 0.200; significance level $\alpha = 5\%$
  1. How many discoveries under the naive rule, Bonferroni, Holm and BH? Write down each threshold you compare with
  2. Under BY the BH thresholds are divided by $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$: how many discoveries remain?
  3. If all five nulls were true and the tests independent, what is the probability that at least one p-value falls below 0.05?
  4. Which procedure would you use in a search over many candidate factors?
  5. How do you limit to 5% the probability of rejecting at least one true null hypothesis?
- Answers: ▸ Appendix

Do Published Factors Replicate?

- **Anomaly**: an average return that the CAPM does not explain
- **Two critiques** of factor research (Jensen, Kelly & Pedersen (2023))
  - Results cannot be reproduced: Hou, Xue & Zhang (2020) replicate only 35% of the anomalies they study
  - Results are the product of testing many factors without correction: Harvey, Liu & Zhu (2016) conclude that most claimed findings are likely false
- **Questions**
  - do published factors have significant CAPM alphas in the original data?
  - are these alphas significant out of sample and in other countries?
  - do they remain significant once multiple testing is taken into account?

Why Replication Matters for This Chapter

- If most factors were false discoveries, multifactor models and factor ETFs (Sections 5 and 8) would rest on noise
- Feng, Giglio & Xiu (2020) show that most new factors add nothing to the existing ones (▸ Appendix)
- Jensen, Kelly & Pedersen (2023) instead see the zoo as many correlated measurements of the same few economic themes
- Their method, empirical Bayes: each factor's alpha is shrunk towards zero and pulled towards the mean of the factors in the same theme
- The quantity tested in the case study is Jensen's alpha (Section 4)
- The classical (non-Bayesian) benchmark is the Benjamini–Yekutieli correction (Section 7)

Most Factors Replicate, but Alphas Fall after the Original Sample (1/3)

Replication rates (Figure 1)

Hou, Xue & Zhang (2020), raw returns	35.0%
Jensen, Kelly & Pedersen (2023), raw returns	55.6%
Raw returns, 119 originally significant	61.3%
CAPM alpha, OLS	82.4%
CAPM alpha, Benjamini–Yekutieli	75.6%
CAPM alpha, empirical Bayes	82.4%
World, empirical Bayes	82.4%

- From 35% (Hou, Xue & Zhang (2020)) to 82.4%
  - With a longer sample, 15 more factors and a different construction, the rate rises to 55.6%
  - Keeping only the factors significant in the original studies, it becomes 61.3%
  - Testing CAPM alphas instead of raw returns, it becomes 82.4%
  - this is the largest step

Most Factors Replicate, but Alphas Fall after the Original Sample (2/3)

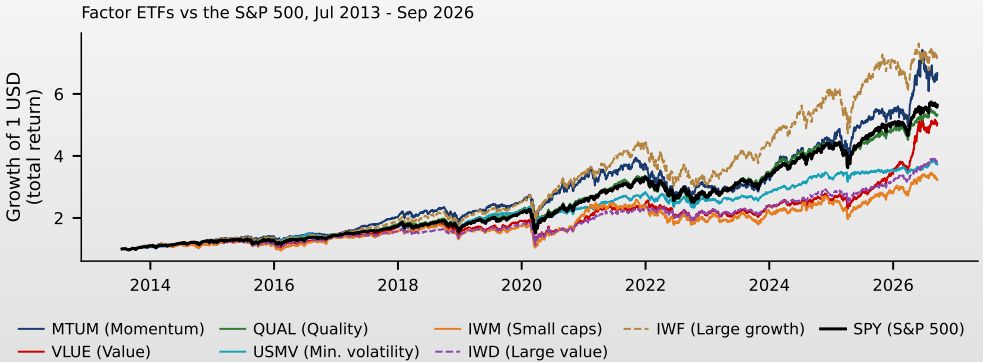
- Empirical Bayes also gives 82.4%
  - pulling towards zero loses some discoveries
  - information borrowed from the theme adds others, and the effects offset
  - complete results: Appendix
- The Bayesian FDR among factors with $z \geq 1.96$ is only 0.1%
  - Bayesian FDR: the estimated share of false discoveries
  - although the hurdle stays at 1.96
- The replication rate refers to the 119 factors significant in the original studies, not to all factors ever proposed

Most Factors Replicate, but Alphas Fall after the Original Sample (3/3)

- **Decay** after the period of the original study
  - On average a factor's alpha falls from 0.49% a month in-sample (the original study's sample)
  - to 0.26% a month out-of-sample, a 47% drop
  - The slope of out-of-sample alpha on in-sample alpha is 0.26 (Figure 8), a coefficient distinct from the mean of 0.26% a month
  - an initial alpha higher by 1 percentage point a month stays higher later by only 0.26 percentage points
  - Statistical shrinkage alone would explain a slope of about 0.9, so the real decay is larger than noise
- The strongest factors decay most
  - a sign of arbitrage or of data mining (Further Topics)
- Two distinct results (Jensen, Kelly & Pedersen (2023))
  - reproduction: for 98 of the 119 originally significant factors, the CAPM alpha (OLS) reproduces the result (82.4%)
  - effect size: the average alpha falls from 0.49% to 0.26% a month (−47%); the cause of the decay stays open

Factor ETFs versus the Market (1/2)

The question: does an investor who buys the factors through ETFs earn more than with the market?



MFM_ch3_factor_etfs ▶ Interactive

Factor ETFs versus the Market (2/2)

- The chart: the value of 1 USD (US dollar) invested on 18 July 2013
  - dividends reinvested (total return)
  - SPY is the black line
- **Smart beta**: long-only ETFs that buy stocks by fixed rules
  - so they keep market exposure and do not replicate the academic long–short factor
- IWD (iShares Russell 1000 Value ETF), IWF (iShares Russell 1000 Growth ETF): the value and growth sub-indices of the Russell 1000 (the 1000 largest US stocks); IWM (iShares Russell 2000 ETF): the Russell 2000 (the next 2000)
- What explains a higher CAGR? The MTUM (iShares MSCI USA Momentum Factor ETF) example
  - daily regression on FF5 + MOM, to July 2026: MOM exposure 0.36
  - annualised alpha -0.2%, $t = -0.11$
  - these data are consistent with zero alpha
- **CAGR** (Compound Annual Growth Rate): $(V_T/V_0)^{1/n} - 1$ over n years
- Above SPY (14.0%)
  - large growth (IWF) 16.3%
  - momentum (MTUM) 15.5%
- Below SPY
  - QUAL (iShares MSCI USA Quality Factor ETF) 13.5%, VLUE (iShares MSCI USA Value Factor ETF) 13.1%, IWD 10.7%
  - USMV (iShares MSCI USA Min Vol Factor ETF) 10.5%, IWM 9.3%
- Thirteen years are a single historical path
  - too short to reject factor premia estimated over six decades

Conclusions: The Steps of a Factor Analysis

1. **Estimate exposures** with HAC errors
   - report sample and proxy
   - shrink extreme betas (Blume (1971), Vasicek (1973))
2. **Test all the alphas at once**
   - GRS measures, normalised, how much the squared maximum Sharpe ratio rises when the test assets are added
   - the CAPM fails on the 25 size $\times$ B/M portfolios (flat SML, GRS rejects)
3. **Estimate premia** in two passes
   - with errors that account for estimated betas and for a possibly wrong model
   - a high R^2 on the 25 portfolios proves little
4. **Guard against the zoo**: use hurdles corrected for multiple testing
5. **Check replication**: 82.4% of the 119 originally significant factors replicate
   - but the average alpha falls by 47% after the original sample
6. **Decompose factor ETFs**: they mostly sell exposure to known factors, not alpha (MTUM: Section 8; the rest: Further Topics)

Check Your Understanding

Retrieval questions

- Why does the HAC standard error of a beta differ from the classical one, while the beta does not? (Further Topics)
- Why can GRS fail to reject the CAPM on sector ETFs but reject it on the 25 portfolios over the same months? (Section 4)
- Why does a useless factor look priced in pass 2? (Further Topics)
- Why can the empirical-Bayes replication rate exceed the Benjamini–Yekutieli rate? (Section 7)

Links to Other Chapters

- Chapter 2, already covered: a non-zero alpha is either an anomaly or the sign of a missing risk factor
  - the joint hypothesis problem
- Chapter 4: factor covariance models and shrinkage for Markowitz
- Chapters 5–6: time-varying volatility, correlation and betas
- Chapter 13: backtest overfitting and machine learning for the cross section of returns

The Practical Decision

- **What I choose**
  - a beta with HAC errors and shrinkage (XLK on SPY, 1999–2012: $\hat{\beta} = 1.415$, Blume 1.278, Vasicek 1.344; realised 2013–2026: 1.172)
  - the GRS test on assets chosen for economic contrasts set in advance: the 25 size $\times$ B/M portfolios, not only the sectors (GRS 2.91 against 0.53, over the same 331 months)
  - Fama–MacBeth premia with KRS errors (with Newey–West lags); FM and Shanken errors only as comparisons; a new factor only above the multiple-testing hurdle
  - robustness to misspecification and autocorrelation does not solve weak identification
- **Under which conditions**
  - the HAC errors require a finite fourth moment, which Chapter 1 questions
  - GRS assumes multivariate Normal errors independent of the factors; a rejection also concerns the index used as the market (Roll)
  - the market premium stays weakly identified when betas differ little (0.89–1.14 in FF3)
- **Which results would change the decision**
  - shrinkage: a larger forecast error than OLS across several assets and periods (now, over 9 sectors, RMSE: OLS 0.135, Blume 0.109, Vasicek 0.114)
  - a model: a negative cross-sectional R^2 with the premia imposed and no intercept (FF3 on 55 assets: -0.27)
  - a published factor: an out-of-sample alpha that is small in economic terms, read together with its interval (on average, alpha falls by 47%)

Further Topics: Additional Subjects

- The chapter's data and the returns used
- Factors as a stochastic discount factor
- Latent factors: PCA (Principal Component Analysis) and IPCA (Instrumented Principal Component Analysis)
- Case study: factor momentum
- AI (artificial intelligence) for scientific discovery
- Then: worked examples and details for each section, in chapter order

The Data: Which Series We Use

Series	What it contains	Frequency	Period
Fama–French factors	returns on five factors: market, size, value, profitability and investment	monthly	1963–2026
Momentum factor	stocks with the highest returns over months $t-12$ to $t-2$ minus those with the lowest	monthly	1963–2026
25 portfolios	US stocks sorted independently into 5 size groups and 5 B/M (book-to-market) groups; the 25 intersections form the portfolios	monthly	1963–2026
Sector ETFs	one ETF for each sector of the S&P 500 (index of 500 large US companies): 9 sectors from 1999, 11 from 2018	daily	1999–2026
Factor ETFs	momentum, value, quality, minimum volatility (from 2013); Russell size, value and growth (from 2000)	daily	2000–2026
Bucharest Stock Exchange (BVB)	11 large stocks (blue chips) and the exchange index, BET	daily	2015–2026

- Sector ETF tickers
  - XLB materials, XLE energy, XLF financials, XLI (Industrial Select Sector SPDR Fund) industrials, XLK technology
  - XLP consumer staples, XLU utilities, XLV (Health Care Select Sector SPDR Fund) health care, XLY consumer discretionary
  - from 2018 also XLRE real estate and XLC communication services

Which Returns This Chapter Uses

- Notation: simple and log returns
  - $R_{i,t} = P_{i,t}/P_{i,t-1} - 1$ (simple return) and $r_{i,t} = \ln(1 + R_{i,t})$ (log return)
  - $P_{i,t}$: the price adjusted for dividends and splits
- Regressions on the French factors use monthly simple returns
  - The factors are portfolio returns, and a portfolio return is computed from simple returns
  - The mean log return would be lower by about $\sigma^2/2$: at an annual volatility of 20%, about 2 percentage points a year
- Rolling betas, BVB betas and principal component analysis (Further Topics) use daily log returns
- At the BVB the daily risk-free rate is treated as a constant, so beta is the same with raw or excess returns

What the Sectors Hold and Their Tangency Weights

ETF	Sector	What it holds (example companies)	Weight
XLB	materials	chemicals, construction materials, packaging, metals and mining (Linde, Sherwin-Williams, Freeport-McMoRan, Newmont)	−0.21
XLE	energy	oil and gas, oilfield services (Exxon Mobil, Chevron, ConocoPhillips)	+0.19
XLF	financials	banks, insurers, asset managers, payment processors (Berkshire Hathaway, JPMorgan Chase, Visa)	−0.37
XLI	industrials	aerospace and defence, machinery, transport (GE Aerospace, Caterpillar, Union Pacific)	+0.18
XLK	technology	software, hardware, semiconductors (Apple, Microsoft, Nvidia)	+0.13
XLP	consumer staples	food, beverages, household products (Procter & Gamble, Coca-Cola, Walmart)	+0.16
XLU	utilities	electricity, natural gas, water (NextEra Energy, Southern, Duke Energy)	+0.27
XLV	health care	pharmaceuticals, biotechnology, medical devices (Eli Lilly, UnitedHealth, Johnson & Johnson)	+0.37
XLY	consumer discretionary	retail, autos, hotels, restaurants (Amazon, Tesla, Home Depot)	+0.29

- Sectors follow the GICS (Global Industry Classification Standard) classification of S&P 500 companies
  - the composition changed: real estate left XLF in 2016; in 2018 Alphabet and Facebook (now Meta), from XLK, and Netflix and Comcast, from XLY, moved to communication services
- The sign of each weight depends on the whole covariance matrix, not on a single pair of sectors
  - example: XLB and XLI have correlation 0.85, and XLI has a higher mean (8.78% vs 8.19%) and lower volatility (18.71% vs 20.96%)
- XLF has the lowest Sharpe ratio of the 9 sectors (0.29) and a correlation of 0.80 with XLI

MFM_ch3_frontier

Factors as a Stochastic Discount Factor *(deeper)*



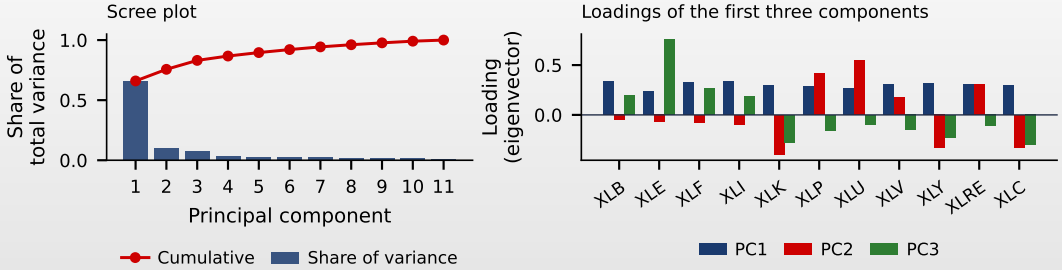
Lars Peter Hansen, Nobel Prize in Economics 2013

Photo: Bengt Nyman (2013); CC BY 2.0; Wikimedia Commons

- The question: is a factor needed to price the assets, once we have the other factors?
  - survey: Cochrane (2011)
- **SDF** (Stochastic Discount Factor) m_t: prices all excess returns, $\mathbb{E}[m_t R^e_{i,t}] = 0$
- m_t is high in bad times
  - an asset that pays off exactly then is insurance, so it earns a low mean return
- A linear factor model is an SDF, $m_t = 1 - \mathrm{b}^\top(\mathrm{f}_t - \mathbb{E}\mathrm{f})$
  - b: the SDF coefficients
- Substitute: $\mathbb{E}[R^e_i] = \mathrm{Cov}(R^e_i, \mathrm{f}_t)^\top \mathrm{b} = \boldsymbol{\beta}_i^\top \boldsymbol{\lambda}$
  - $\boldsymbol{\lambda} = \boldsymbol{\Sigma}_f \mathrm{b}$ (Section 6)
- Two different tests
  - $b_k = 0$: factor k does not help pricing, given the other factors
  - $\lambda_k = 0$: exposure to factor k is not rewarded
  - with correlated factors, $\lambda_k \neq 0$ and $b_k = 0$ can coexist
- b is estimated by GMM (Hansen (1982); Further Topics)
  - requiring the sample mean of $m_t(\mathrm{b})\, \mathrm{R}^e_t$ to be zero for every asset

PCA of the US Sector ETFs (1/2)

The question: which factors emerge from returns if we do not choose them in advance?



Left (scree plot): each component's share of total variance (bars) and the cumulative share (line); right: the loadings (eigenvectors) of the first three components on each ETF

MFM_ch3_pca ▶ Interactive

PCA of the US Sector ETFs (2/2)

- **PCA** (Principal Component Analysis)
  - the linear combinations of returns that explain, in turn, the largest part of the remaining variance
  - formula: next slide
- 11 sector ETFs, daily log returns
  - 20 June 2018 – 18 September 2026 ($T = 2073$ common days)
- PC1 (first Principal Component) explains 66.0% of the variance
  - all loadings positive (0.24–0.34): a near-equal average of the sectors
- PC1 is the market factor: its correlation with SPY is 0.954
- PC2 (9.7%): **defensive versus growth sectors**
  - utilities $+0.55$, staples $+0.42$, real estate $+0.31$
  - vs technology -0.40, consumer discretionary -0.33
- PC3 (7.4%) is dominated by energy
  - the first three components explain 83.1%
- The result describes the sample; it does not identify a cause

MFM_ch3_pca

Principal Components

- **Latent factors**: unobserved factors, estimated from returns alone
  - without saying in advance what they represent
- PCA of the correlation matrix C of standardised returns

$$\mathsf{C} = \mathbf{\Gamma}\mathsf{D}\mathbf{\Gamma}^\top, \qquad \mathsf{y}_t = \mathbf{\Gamma}^\top \mathsf{z}_t$$

  - z_t: standardised returns of day t
  - y_t: the components
  - $\mathbf{\Gamma}$: eigenvectors in columns, the **loadings**
  - D: diagonal matrix of the eigenvalues ℓ_k
- Eigenvalues, sorted: $\ell_1 \geq \cdots \geq \ell_N$ (written ℓ so as not to confuse them with the risk premia λ)
- Component k explains ℓ_k/N of the total variance
- Example, two assets with correlation ρ: $\ell = 1 \pm \rho$
  - the vector $(1,1)/\sqrt{2}$ is the "market"
  - $(1,-1)/\sqrt{2}$ is the difference between the two
- In a crisis ρ rises and the first component dominates

Limits of PCA for Asset Pricing

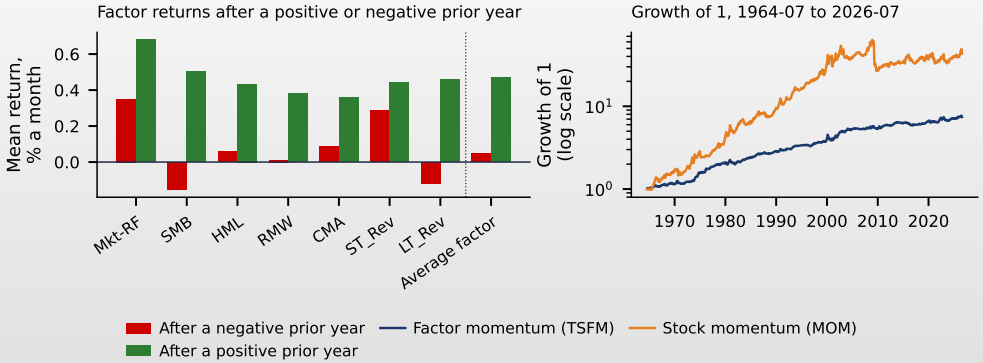
- Loadings are **constant**: PCA cannot follow a stock whose size or value changes
- Components explain **variance**, not necessarily **average returns**
- Individual stocks enter and leave the sample (listings, delistings)
  - so many observations are missing, while classic PCA needs a complete panel
- The sign of each component is arbitrary: flipping an eigenvector changes nothing
- Latent factors are determined only up to a rotation
  - any rotated combination of them fits the data equally well
  - so interpreting them is a choice
- Two extensions of PCA also target average returns
  - risk-premium PCA and the three-pass estimator (Appendix)

Factor Momentum: Does Stock Momentum Only Time Other Factors?

- **Question**: are factor returns autocorrelated, and is the momentum factor MOM (Section 5) only a way of timing the other factors?
- **Why it matters**
  - the Carhart model (FF3 + MOM) and FF5 + MOM add MOM as a separate source of risk
  - if MOM only times other factors, a model needs factor timing (moving capital between factors on signals), not a separate momentum factor
- **Main claims** of Ehsani & Linnainmaa (2022)
  - the average factor earns 0.06% a month after a year of losses
  - and 0.51% a month after a positive year
  - factor momentum concentrates in the factors that explain more of the cross section
  - momentum in high-eigenvalue principal-component factors subsumes most forms of stock momentum
  - conclusion of the authors: momentum is not a distinct risk factor, it times the other factors

Factor Momentum in the French Factors, 1964–2026

Question: does a factor that rose over the past year keep rising in the next month?



MFM_ch3_factor_momentum ▶ Interactive

Factor Momentum: the Course Replication

- **Data**: seven monthly factors of the Kenneth French Data Library, July 1963 – July 2026
  - Mkt-RF, SMB, HML, RMW, CMA (FF5)
  - ST_Rev (short-term reversal): long the losers of month $t-1$, short its winners
  - LT_Rev (long-term reversal): long the losers of months $t-60$ to $t-13$, short their winners
  - MOM stays out of the set: it is the series to explain
- **Signal** $s_{i,t}$: the compounded return of factor i over months $t-12$ to $t-1$, as in time-series momentum (Moskowitz, Ooi & Pedersen (2012))
- **TSFM** (time-series factor momentum): $\mathrm{TSFM}_t = \frac{1}{7}\sum_{i=1}^{7} \mathrm{sign}(s_{i,t})\, r_{i,t}$
  - long the factors with a positive prior year, short the others; first month July 1964 ($T = 745$)
- **Tests**
  - pooled difference of the means after positive and negative years, standard errors clustered by month
  - spanning regressions of MOM on FF5 and TSFM, Newey–West errors
- **Difference from the paper**: the paper uses many more factors, including principal-component factors; seven factors give a weaker spanning test

Factor Momentum: Result and Limits

- Average factor, mean return in the next month
  - after a negative prior year: 0.05% a month
  - after a positive prior year: 0.47% a month
  - difference 0.42 percentage points, $t = 3.62$
  - the paper reports 0.06% and 0.51%
- The mean is higher after a positive year for all seven factors
- TSFM: mean 0.28% a month, $t = 5.48$
  - annual Sharpe ratio 0.67; MOM: 0.49
  - maximum drawdown -15%; MOM: -58% (September 2009)
- FF5 alpha of MOM: 0.74% a month ($t = 4.18$)
  - with TSFM added: 0.32% ($t = 2.27$), 57% lower
  - slope on TSFM 1.76 ($t = 13.8$)
- Alpha of TSFM on FF5 and MOM: 0.08% ($t = 1.81$)
- The result describes the sample
  - factor returns are autocorrelated over twelve months
  - seven factors explain about half of the MOM alpha, not all of it
- Limits
  - one sample, no trading costs
  - the two reversal factors are themselves built from past returns

MFM_ch3_factor_momentum

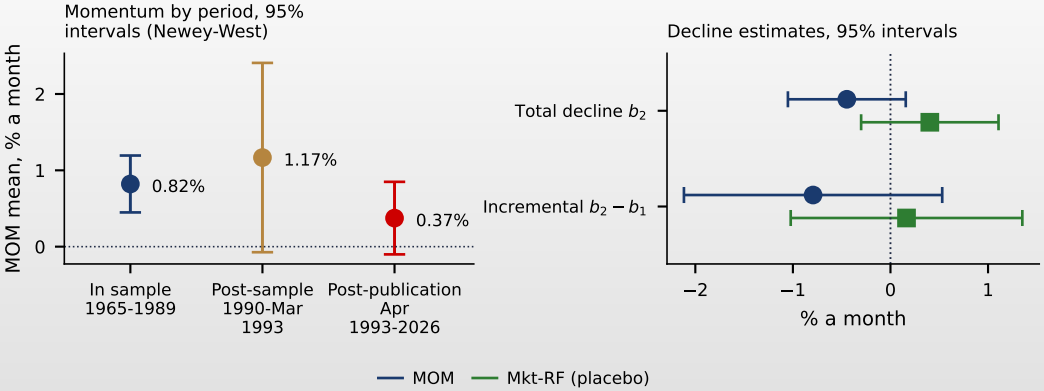
Open Questions and the Discovery Loop

- **Open questions**
  - why do factor premia decay after publication: arbitrage, data mining or a change in risk?
  - does the decay differ across factors and markets?
- Evidence so far
  - McLean & Pontiff (2016): across 97 predictors, returns are 26% lower out of sample and 58% lower after publication
  - the authors attribute the gap to investors trading on the published papers
  - Chen & Zimmermann (2022): open data and code for the published predictors
- **The discovery loop**
  - The AI assistant drafts: a literature table with DOIs (Digital Object Identifiers), competing hypotheses and what each predicts, the regression code, a robustness plan (prompts: Appendix)
  - The researcher fixes before running: the event dates, the list of factors, the multiple-testing correction and the placebo test (a factor where the effect should not appear)
  - Before interpreting: every DOI (Digital Object Identifier) resolves in Crossref, the split-sample t-statistics of Further Topics are reproduced, and the version of the French data is recorded

Mini-Case: Was Momentum Arbitraged Away? (1/3)

- Design of McLean & Pontiff (2016): $\text{MOM}_t = a + b_1 \text{PostSample}_t + b_2 \text{PostPub}_t + e_t$
  - monthly, January 1965 – August 2026 ($T = 740$), Newey–West errors
- Dummy variables
  - $\text{PostSample}_t = 1$ in January 1990 – March 1993 (post-sample)
  - $\text{PostPub}_t = 1$ from April 1993 (post-publication)
- a: the in-sample mean (1965–1989)
- b_2: total decline relative to the in-sample mean
- $b_2 - b_1$: incremental effect of publication

Mini-Case: Was Momentum Arbitraged Away? (2/3)



MFM_ch3_ai_discovery ▶ Interactive

Mini-Case: Was Momentum Arbitraged Away? (3/3)

- In-sample momentum mean: $a = 0.82\%$ a month ($t = 4.32$)
- Decline after publication: $b_2 = -0.45\%$ a month, a mean 54% lower
  - $t = -1.45$, interval $[-1.05\%, 0.16\%]$
- Incremental effect of publication: $b_2 - b_1 = -0.79\%$ ($t = -1.17$)
- Placebo on Mkt-RF: $b_2 = +0.40\%$ ($t = 1.12$)
- The decline is close to the 58% of McLean & Pontiff (2016)
  - but with one factor we cannot reject $H_0 : b_2 = 0$ at 5%
- The result describes the sample
  - a before/after-publication comparison; arbitrage is one possible interpretation, not a demonstrated cause

MFM_ch3_ai_discovery

Project: Do Factor Premia Fall after Publication?

- **Questions**
  - do the Fama–French factors and momentum decay after publication?
  - does the decay differ across factors?
- **Design**
  - one pooled regression of SMB, HML, RMW, CMA and MOM, with each factor's own dummies for the post-sample and post-publication periods
  - Mkt-RF as placebo
  - errors robust to correlation across factors and to autocorrelation
  - whether the decay is larger where arbitrage is easier can only be tested with stock-level data (size, liquidity, short-sale costs)
- **Replicate first**: the split-sample t-statistics and the mini-case estimates

The Data: Sources and Notation

Sources

- Factors and the 25 portfolios: Kenneth French Data Library, monthly returns in percent
- Last month available: July 2026 (for momentum: August 2026)
- ETFs and stocks: daily adjusted closing prices (dividends reinvested)
- Last day available: 18 September 2026

Notation

- **BET**: the BVB benchmark index, a price index (no dividends)
- **Excess return**: $R^e_{i,t} = R_{i,t} - R_{f,t}$
- $R_{f,t}$: one-month US Treasury bill rate (RF in the French data)
- **Mkt-RF**: excess return of the US stock market, weighted by market capitalisation

- Why these data
  - the very different mean returns of the 25 portfolios make the test demanding
  - the ETFs are investable; the BVB applies the methods to the local market

The Cost of Capital: Which Uncertainty Matters?

- CAPM cost of capital: $R_f + \hat{\beta}_i \left(\widehat{\mathbb{E}[R_m]} - R_f\right)$
- The standard error (SE) of the market premium: $15.4\%/\sqrt{63} = 1.9\%$ a year
  - 15.4%: annual market volatility
  - 63: years of data
- Multiplied by the XLK beta, this gives an uncertainty of 2.5% a year in the cost of capital
  - XLK beta: $\hat{\beta} = 1.28$ (Further Topics)
- The uncertainty of beta adds only $0.082 \times 7.2\% = 0.6\%$ a year
  - the SE of beta times the premium
- Conclusion: the uncertainty of the market premium weighs far more than that of beta

HAC Errors: Weights, Lags and Assumptions

- The Newey–West estimator (Newey & West (1987)) adds up the autocovariances of the moments $R^e_{m,t}\varepsilon_{i,t}$ up to lag L
  - with Bartlett weights $1 - \ell/(L+1)$, which make the estimated matrix positive semidefinite
- The lag rule: $L = \lfloor 4(T/100)^{2/9} \rfloor$ (Newey & West (1994))
  - $T = 331$ gives $L = 5$; $T = 757$ gives $L = 6$
- The point estimates remain those of OLS
- HAC errors are not necessarily larger than the classical ones in every sample
- HAC errors handle dependence but require a finite fourth moment, which Chapter 1 questions

Worked Example: Beta of the Technology ETF (1/2)

		Standard error		
	Estimate	**Classical**	**White**	**Newey–West**
α (annual)	0.33%	2.27%	2.32%	2.45%
β	1.277	0.041	0.054	0.082

- XLK on Mkt-RF, monthly January 1999 – July 2026, $T = 331$, $R^2 = 0.742$
- Columns of the table
  - **Classical**: errors assumed homoskedastic and serially uncorrelated
  - **White**: robust to heteroskedasticity (White (1980))
  - **Newey–West**: also robust to autocorrelation, $L = 5$ lags
- The robust SE of β is twice the classical one
- Heteroskedasticity explains part of the increase (White column)
  - autocorrelation of the product $R^e_{m,t}\varepsilon_t$ explains the rest
- The White and Newey–West errors require a finite fourth moment, which Chapter 1 questions

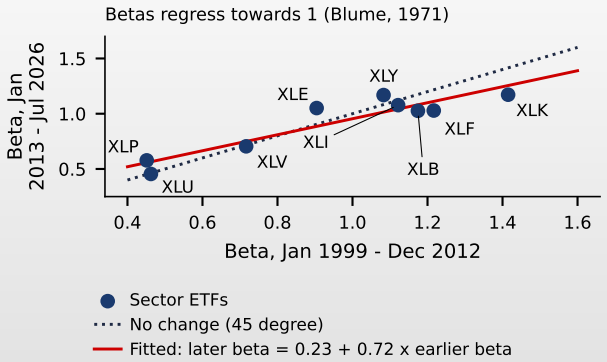
MFM_ch3_seminar

Worked Example: Beta of the Technology ETF (2/2)

- t-statistic of alpha: $t(\alpha) = 0.33/2.45 = 0.13$
- 95% bootstrap interval for alpha: $[-5.5\%, 4.9\%]$
  - bootstrap SE 2.61%, slightly above 2.45%
  - moving-block bootstrap of 12-month blocks of the pairs $(R^e_{\text{XLK},t}, R^e_{m,t})$, which keeps the dependence over time
  - 2 000 resamples (Seminar 3, problem B1)

MFM_ch3_seminar

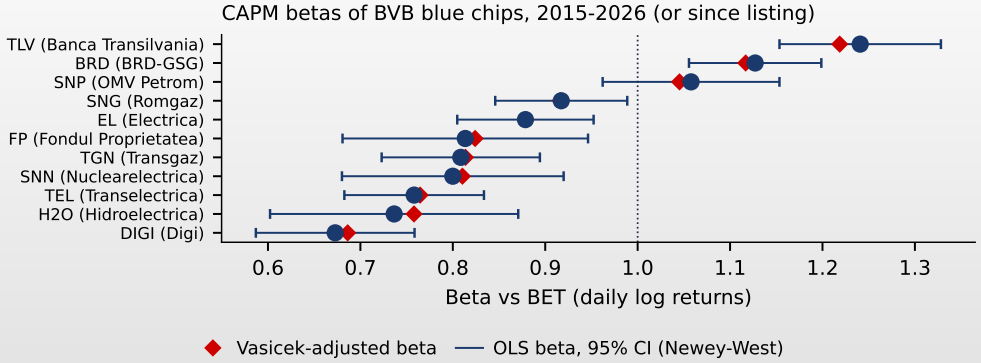
Do Betas Regress Towards One?



- Monthly betas of 9 sector ETFs on SPY: 1999–2012 (horizontal) vs 2013–2026 (vertical)
- Fitted line $\hat{\beta}_{\text{later}} = 0.23 + 0.72\,\hat{\beta}_{\text{earlier}}$
  - very close to Blume's $0.33 + 0.67\,\hat{\beta}$
- Forecast error: the **RMSE** (Root Mean Squared Error)
  - the unadjusted beta has RMSE 0.135
  - the Blume beta has RMSE 0.109
  - the Vasicek beta has RMSE 0.114
- Shrinkage cuts the forecast error by 15–20%
- With only 9 sectors the lesson is qualitative

MFM_ch3_beta_shrink

CAPM Betas on the Bucharest Stock Exchange (1/3)



Legend: OLS beta with its 95% confidence interval (CI), Newey–West errors; Vasicek-adjusted beta

MFM_ch3_bvb_betas

CAPM Betas on the Bucharest Stock Exchange (2/3)

- Daily log returns on BET, from 6 January 2015 or the listing date
  - listing dates: DIGI May 2017, H2O (Hidroelectrica, BVB ticker) July 2023
  - FP prices are adjusted for the special dividend
- Banks are the most market-sensitive
  - Banca Transilvania (TLV) 1.24, with $R^2 = 0.62$
  - BRD (BRD–Groupe Société Générale, BVB ticker) 1.13
- The least sensitive
  - Digi 0.67, with $R^2 = 0.21$
  - Hidroelectrica 0.74 and Transelectrica 0.76
- R^2 is high for TLV partly because TLV has a large weight in BET
  - the stock is part of its own benchmark

MFM_ch3_bvb_betas

CAPM Betas on the Bucharest Stock Exchange (3/3)

- Vasicek shrinkage changes the estimates little
  - daily betas are precise (SE between 0.04 and 0.07)
- Stale prices bias daily betas towards zero
  - the Dimson beta rises from 0.80 to 0.94 for Nuclearelectrica
  - and from 0.81 to 0.905 for Transgaz (Seminar 3, problem B4)
- BET is a price index while the stock prices include dividends
  - alpha is pushed up by roughly beta times the dividend yield of the index
  - beta changes little with BET-TR (BET Total Return, index)

MFM_ch3_bvb_betas

A Flat SML: Black and Betting Against Beta (1/2)



Fischer Black (1938–1995)

Photo: Dalmatine at English Wikipedia (2009 upload); public domain; Wikimedia Commons

- Black (1972): without unlimited borrowing at R_f, expected return stays linear in beta
  - but the line starts at the return of a **zero-beta** portfolio, above R_f
- Prediction: the SML is **flatter** than the CAPM line
- Frazzini & Pedersen (2014): investors who want more risk but cannot borrow buy high-beta assets
  - these become expensive and earn lower future returns

A Flat SML: Black and Betting Against Beta (2/2)

- Their strategy, **BAB** (Betting Against Beta):

$$r_t^{\text{BAB}} = \frac{R_{L,t} - R_{f,t}}{\beta_L} - \frac{R_{H,t} - R_{f,t}}{\beta_H}$$

  - $R_{L,t}$, $R_{H,t}$: returns of the low-beta (L) and high-beta (H) portfolios
  - β_L, β_H: their betas
- The strategy is long $1/\beta_L > 1$ units of L and short $1/\beta_H < 1$ units of H
  - the long leg is financed at R_f; the proceeds of the short leg are lent at R_f
  - it has beta zero and a positive average return in their data
- A plain low-minus-high-beta spread, without scaling, has a negative beta
  - its mean mixes the BAB premium with a negative market exposure (Seminar 3, problem C2)

GRS as a Sharpe-Ratio Test

- An algebraic identity (Gibbons, Ross & Shanken (1989)), with all moments estimated by maximum likelihood

$$\hat{\boldsymbol{\alpha}}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}} = \widehat{SR}^2(f,\mathrm{R}) - \widehat{SR}^2(f)$$

  - $\widehat{SR}^2(f,\mathrm{R})$: squared Sharpe ratio of the tangency portfolio of the factor and the N test assets
  - $\widehat{SR}^2(f)$: that of the factor alone
- So GRS measures how much the Sharpe ratio rises when the test assets are added

$$W = \frac{T-N-1}{N}\cdot\frac{\widehat{SR}^2(f,\mathrm{R}) - \widehat{SR}^2(f)}{1+\widehat{SR}^2(f)}$$

- The monthly Sharpe ratio of the market (Mkt-RF) alone is 0.134 (0.47 a year)
- The monthly Sharpe ratio of the tangency portfolio of the market and the 25 portfolios is 0.405 (1.40 a year)
- $W = 29.24 \times (0.405^2 - 0.134^2)/(1+0.134^2) = 4.20$, with $29.24 = (757-25-1)/25$
- Under the CAPM, noise alone would raise the monthly $\widehat{SR}^2$
  - on average by $N(1+\widehat{SR}^2(f))/(T-N-3) = 0.035$
  - observed: 0.146 (Appendix)

The Generalized Method of Moments (1/2)

- An economic model implies $\mathbb{E}[g_t(\boldsymbol{\theta})] = 0$ for a vector of moment conditions g_t
- **GMM** (Generalized Method of Moments, Hansen (1982)) chooses $\hat{\boldsymbol{\theta}}$ so that the sample mean of the moments is zero

$$\bar{g}(\hat{\boldsymbol{\theta}}) = \frac{1}{T} \sum_{t=1}^{T} g_t(\hat{\boldsymbol{\theta}}) = 0$$

  - exactly identified case: as many moments as parameters
  - with more moments than parameters, not all can be zero, and $\bar{g}$ is made as small as possible with a weighting matrix
- Sandwich variance in the exactly identified case

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) \approx \frac{1}{T} D^{-1} S D^{-1\top}, \qquad D = \mathbb{E}\left[\partial g_t / \partial \boldsymbol{\theta}^\top\right], \qquad S = \sum_{j=-\infty}^{\infty} \mathrm{Cov}(g_t, g_{t-j})$$

  - D: Jacobian of the moments (their derivatives with respect to the parameters)
  - S: long-run variance of the moments, autocovariances included, estimated by Newey–West (Section 3)

The Generalized Method of Moments (2/2)

- The simplest case, a mean: $g_t = R_t - \mu$ gives $D = -1$ and $\mathrm{Var}(\hat{\mu}) \approx S/T$
  - this is the HAC t-statistic of a mean factor return (Further Topics)
- OLS: $g_t = x_t(y_t - x_t^\top \boldsymbol{\theta})$ gives $D = -\mathbb{E}[x_t x_t^\top]$
  - White errors when S has no lags, Newey–West errors otherwise (Further Topics)
- **GRS as GMM**: parameters (α_i, β_i) for $i = 1, \ldots, N$
  - $2N$ moments $\varepsilon_t \otimes (1, f_t)$: each residual times 1 and times the factor ($\otimes$: Kronecker product)
  - the system is exactly identified, so the estimates are the OLS ones, asset by asset; GMM changes only their variance
  - ordering the parameters asset by asset, $(\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots)$, gives $D = -I_N \otimes (Z^\top Z/T)$, with Z the $T \times 2$ matrix of rows $(1, f_t)$
- $\hat{V}_\alpha$: the block of the sandwich matrix that belongs to the alphas, computed with a HAC $\hat{S}$
- Wald statistic $W_{\mathrm{GMM}} = \hat{\boldsymbol{\alpha}}^\top \hat{V}_\alpha^{-1} \hat{\boldsymbol{\alpha}} \to \chi^2_N$ under H_0
- The same machinery gives the misspecification-robust premia errors (Appendix)
  - it also estimates the discount-factor model of Further Topics

When the GRS Assumptions Fail

- **Heavy tails and volatility clustering** (Chapter 1): the F distribution is no longer exact
  - Residual bootstrap under H_0: resample whole rows of residuals, rebuild returns with $\alpha = 0$ (Seminar 3, problem B7)
  - The bootstrap is robust to non-Normality only if the errors are i.i.d. and independent of the factors
  - GMM Wald test (Further Topics), with a HAC $\hat{V}_\alpha$: robust to heteroskedasticity and autocorrelation
- **Too many assets for the number of months**: $\hat{\Sigma}$ has rank at most $T - 2$
  - when $N \geq T - 1$, $\hat{\Sigma}$ cannot be inverted and GRS cannot be computed
  - the power of the test falls well before that limit
- With many assets there are tests that do not invert $\hat{\Sigma}$ (Pesaran & Yamagata (2024); Appendix)

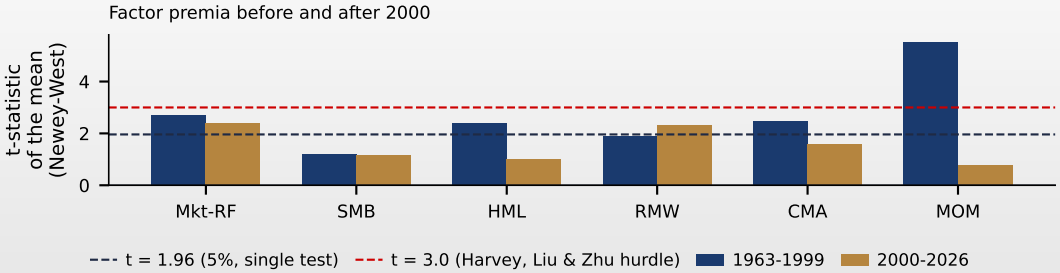
The Motivation of FF5: the Dividend-Discount Model

- Fama & French (2015) start from dividend-discount valuation

$$\frac{M_t}{B_t} = \frac{\sum_{\tau \geq 1} \mathbb{E}[Y_{t+\tau} - dB_{t+\tau}]/(1+r)^{\tau}}{B_t}$$

  - M: market value, B: book value, Y: earnings, dB: investment
  - r: the expected return, the discount rate
- At a given M/B, high expected profitability (RMW) or low expected investment (CMA) implies a high expected return r
- The relation motivates the factors, but does not show that they are priced risks

Stability of the Premia over Time (1/2)



MFM_ch3_factor_tstats

Stability of the Premia over Time (2/2)

- Bars: the Newey–West t-statistic of the mean monthly factor return
  - 1963–1999 vs 2000–2026
  - an arbitrary cut, not the publication date
  - the decay after publication: Further Topics
- Lines: $t = 1.96$ (5%, a single test) and $t = 3$
  - $t = 3$: the multiple-testing hurdle of Harvey, Liu & Zhu (2016) (Section 7)
- The second period has 319 months against 438
  - so t falls by a factor $\sqrt{319/438} \approx 0.85$ even with an unchanged mean
  - compare the means too
- Momentum: t falls from 5.53 to 0.75
  - the annual mean from 10.7% to 2.5%
- Value (HML): t falls from 2.39 to 1.02
  - the annual mean from 4.2% to 2.8%
- Profitability (RMW) is the exception: t rises from 1.90 to 2.31
- The market premium stays significant in both periods ($t = 2.70$ and 2.39)

MFM_ch3_factor_tstats

Standard Errors When the Model Is Misspecified

- The Fama–MacBeth, Newey–West and Shanken errors assume that the model is correct
  - that is, the pricing errors are zero: $\mathrm{e} = \boldsymbol{\mu} - \mathrm{X}\boldsymbol{\gamma} = 0$
  - $\boldsymbol{\mu}$: mean excess returns, $\boldsymbol{\gamma} = (\lambda_0, \boldsymbol{\lambda}^\top)^\top$
- If the model is wrong ($\mathrm{e} \neq 0$), the noise in the estimated betas adds variance to the premia
- The standard errors robust to this are those of Kan–Robotti–Shanken (Kan, Robotti & Shanken (2013); formula: ▸ Appendix)
- Misspecification-robust is not HAC: serial dependence also needs the autocovariances of the moments (Newey–West)

Worked Example: Four Standard Errors for the Value Premium

FF3, HML	$\hat{\lambda}$	SE	t	Significant at 5%?
Fama–MacBeth	3.97%	1.33%	2.99	yes
Newey–West (6 lags)	3.97%	1.62%	2.45	yes
Shanken	3.97%	1.33%	2.99	yes
Kan–Robotti–Shanken (no lags)	3.97%	1.32%	3.00	yes
Kan–Robotti–Shanken (6 lags)	3.97%	1.61%	2.47	yes

- 25 size $\times$ B/M portfolios, 1963–2026
- The Shanken correction leaves the error unchanged, because $c = 0.03$ per month
- What matters most is the autocorrelation of the monthly slopes $\hat{\lambda}_t$
  - Newey–West with 6 lags raises the SE from 1.33% to 1.62%
- For RMW in FF5, Kan–Robotti–Shanken raises the error from 2.1% to 2.6%, and with 6 lags to 3.0%
  - t falls from 2.5 to 2.1, then to 1.8, below 1.96
- The $t > 3$ hurdle (Section 7) is not clearly passed with any standard error

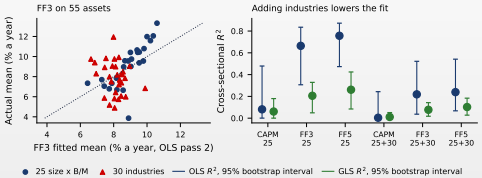
Useless and Weak Factors (1/2)

- A **useless factor** g_t is independent of all returns
  - its true betas are zero
  - the estimated ones are pure noise, of order $1/\sqrt{T}$
- Pass 2 regresses returns on this noise: $\hat{\lambda}_g$ does not go to zero
  - the t-test too often declares the factor priced when the model is wrong (Kan & Zhang (1999))
- The experiment: on the 25 portfolios, we estimate the CAPM plus an invented factor
  - on these portfolios GRS rejects the CAPM (Section 4)
  - the invented factor is drawn independently from the Normal distribution
  - We repeat the experiment 2 000 times and test at the 5% level whether the invented factor's premium is zero
  - A correct test would reject in about 5% of the repetitions
  - With Fama–MacBeth errors, the false premium looks significant in 60% of the repetitions
  - With Shanken errors, it looks significant in 27% of the repetitions

Useless and Weak Factors (2/2)

- The remedy: first test in pass 1 $H_0 : \beta_g = 0$ across all N assets
  - $W = \hat{\beta}_g^\top \widehat{\text{Var}}(\hat{\beta}_g)^{-1} \hat{\beta}_g \sim \chi^2_N$
  - if you do not reject, the premium of the factor cannot be estimated
  - The Wald test rejects in only 6.7% of the repetitions, close to 5%
- **Weak identification**: betas close to zero (a weak factor) or nearly equal across assets
  - nearly equal betas: the market, in our example
  - There are tests that are robust to weak identification (Kleibergen (2009))
  - With many assets, consistency of pass 2 depends on factor strength (Anatolyev & Mikusheva (2022))

A Tougher Test: Adding 30 Industry Portfolios (1/2)



- The 25 portfolios have a strong factor structure (Lewellen, Nagel & Shanken (2010))
  - three factors explain on average 91% of their variation over time (between 79% and 95%)
  - so several models that capture the common size and value variation can reach a high cross-sectional R^2
- Left: actual mean (vertical) against the mean predicted by FF3 (horizontal), % a year
  - an exact model would put all points on the diagonal
  - the industries (red triangles) spread vertically

A Tougher Test: Adding 30 Industry Portfolios (2/2)

- Right panel of the chart: cross-sectional R^2 with 95% bootstrap intervals
  - with 30 industries added, the OLS R^2 of FF3 falls from 0.66 to 0.22
  - and that of FF5 from 0.76 to 0.24
- The FF3 R^2 intervals: $[0.31, 0.84]$ on 25 portfolios, $[0.04, 0.52]$ on 55 assets
  - the marginal intervals do not establish whether the difference is significant: a test needs the same resampled months for both sets of assets and a direct interval for ΔR^2
- The GLS (Generalised Least Squares) R^2 is never above 0.26
  - it weights the errors by the inverse covariance of returns

Comparing Models: Maximum Sharpe Ratios (1/2)

Max. Sharpe ratio (annual)	CAPM	FF3	Carhart	FF5	FF5 + MOM
1963–2026 ($T = 757$)	0.47	0.66	0.97	0.97	1.16
2000–2026 ($T = 319$)	0.48	0.54	0.67	0.97	1.02

- Barillas & Shanken (2018): the better model is the one whose factors allow the highest Sharpe ratio
  - $SR^2(\mathrm{f}) = \boldsymbol{\mu}_f^\top \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\mu}_f$ ($\boldsymbol{\mu}_f$, $\boldsymbol{\Sigma}_f$: means and covariance of the factors)
- The reason: comparing $SR^2(\mathrm{f})$ directly, the test assets no longer matter
  - GRS shows that a model's alphas are small exactly when its factors approach the maximum Sharpe ratio
  - unlike the Hansen–Jagannathan distance, computed on the test assets: Appendix
- When one model contains the other, in-sample SR^2 can only rise

Comparing Models: Maximum Sharpe Ratios (2/2)

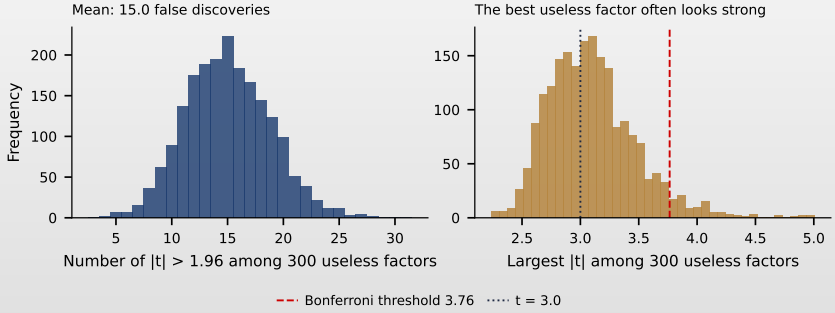
- Then run a **spanning regression**: regress the added factor on the others
  - a non-zero alpha shows that it brings something new
  - HML on the other five factors: $\alpha = 0.9\%$ a year ($t = 0.75$)
  - since 2000: $\alpha = -1.7\%$ ($t = -0.81$)
  - so no significant incremental contribution of HML, as in Fama & French (2015)
- FF5 vs Carhart (neither contains the other): the difference in annual SR^2 is 0.00
  - 95% interval $[-0.57, 0.66]$
- Since 2000: the difference is 0.50, so again not significant
  - 95% interval $[-0.24, 1.48]$
  - stationary bootstrap, Politis & Romano (1994); Seminar 3, problem B8

Worked Example: Ten Candidate Factors

- Ten candidate factors, level $\alpha = 5\%$
  - sorted p-values: 0.0004, 0.002, 0.006, 0.011, 0.018, 0.027, 0.041, 0.070, 0.15, 0.62
- The naive rule rejects every null with $p \leq 0.05$ and finds **7** significant factors
- Bonferroni divides 5% by the 10 tests: only **2** factors pass
  - the threshold is 0.005
- Holm compares the j-th p-value with $0.05/(11 - j)$ and stops at the first one that fails
- The first three p-values pass
  - $0.0004 \leq 0.005$, $0.002 \leq 0.0056$, $0.006 \leq 0.00625$
- The fourth, 0.011, exceeds its threshold 0.0071, so Holm stops at **3** factors
- BH compares the j-th p-value with $0.005\,j$
  - the largest j that passes is 6 ($0.027 \leq 0.030$)
  - so BH declares the first **6** factors significant

Simulation: 300 Useless Factors (1/2)

- The experiment: 300 factors with no true premium
  - each factor's t-statistic is an independent draw from the standard Normal distribution
- We repeat the experiment 2 000 times and count, each time, how many factors pass the hurdle



MFM_ch3_false_discoveries

Simulation: 300 Useless Factors (2/2)

- At the classical hurdle $|t| > 1.96$, on average 15.0 useless factors pass, that is 5% of 300
- At the hurdle $|t| > 3$, on average 0.83 useless factors pass
- In 57% of the repetitions at least one useless factor passes $|t| > 3$
- In half of the repetitions the best useless factor has $|t|$ above 3.06
- The Bonferroni threshold for 300 tests is $|t| > 3.76$
- The right hurdle depends on how many tests were actually tried, unpublished ones included

MFM_ch3_false_discoveries

How Many Factors? (1/2)

- Approximate factor model $x_{it} = \mathbf{\Lambda}_i^\top \mathsf{F}_t + e_{it}$
  - both the number of assets N and of periods T are large
  - x_{it}: standardised return; F_t: factors; $\mathbf{\Lambda}_i$: loadings; e_{it}: idiosyncratic part
  - $V(k)$: mean squared residual after the first k components
- Bai & Ng (2002): choose the k that minimises $\ln V(k) + k \, \frac{N+T}{NT} \ln \min(N, T)$
  - the information criterion IC_{p2}: fit plus a penalty per factor
- Ahn & Horenstein (2013): choose the k where the ratio ℓ_k / ℓ_{k+1} is largest
  - that is, where the eigenvalues drop most sharply
- The rule of Kaiser (1960) ($\ell_k > 1$) has no theoretical justification
  - even pure noise produces eigenvalues scattered around 1

How Many Factors? (2/2)

Panel	N	T	Kaiser	Bai–Ng	Ahn–Horenstein
11 sector ETFs, daily 2018–2026	11	2073	2	5	1
100 size $\times$ B/M portfolios, monthly 1963–2026	95	757	3	4	1

- Standardised returns, with at most $k = 5$ factors for sectors and $k = 8$ for portfolios
  - 5 portfolios with missing months are dropped
- On the 95 portfolios, Ahn–Horenstein picks only the market factor (ratio 14.6)
  - the second peak is at $k = 3$ (ratio 3.5), as in FF3
- Bai–Ng on 11 sectors picks the largest number allowed, $k = 5$
  - its penalty is designed for many assets, so with only 11 the answer is not reliable (Appendix)

Instrumented PCA (1/2)

- **IPCA**: Instrumented Principal Component Analysis, Kelly, Pruitt & Su (2019)
- Loadings are functions of L observed characteristics $\mathrm{z}_{i,t}$ (size, B/M, momentum, ...)

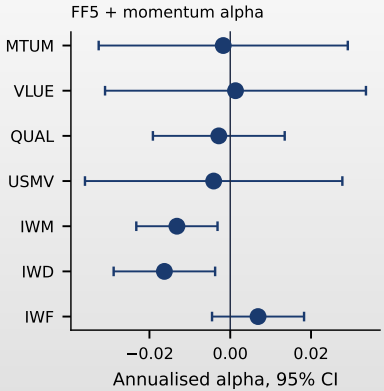
$$R^e_{i,t+1} = \alpha_{i,t} + \boldsymbol{\beta}^\top_{i,t}\mathrm{f}_{t+1} + \varepsilon_{i,t+1}, \qquad \boldsymbol{\beta}^\top_{i,t} = \mathrm{z}^\top_{i,t}\boldsymbol{\Gamma}_\beta$$

- $\boldsymbol{\Gamma}_\beta$ ($L \times K$) and the factors f_{t+1} minimise $\sum_{i,t}(R^e_{i,t+1} - \mathrm{z}^\top_{i,t}\boldsymbol{\Gamma}_\beta\mathrm{f}_{t+1})^2$
  - by alternating steps
  - with $\boldsymbol{\Gamma}_\beta$ fixed, f_{t+1} comes from a cross-sectional regression
  - with the factors fixed, $\boldsymbol{\Gamma}_\beta$ comes from a panel regression

Instrumented PCA (2/2)

- The question tested: why do characteristics predict returns?
  - because they measure risk exposure (through beta)?
  - or directly, as alpha ($\alpha_{i,t} = \mathbf{z}_{i,t}^{\top}\mathbf{\Gamma}_{\alpha}$)?
- Result on US stocks: a few IPCA factors explain average returns
  - better than models with as many observed factors
- $\mathbf{\Gamma}_{\alpha}$ is not significant, hence the paper's title: "characteristics are covariances"
- The motivation: IPCA removes two limits of PCA from the earlier slide
  - constant loadings and the need for a complete panel
  - extensions: Appendix

Alphas of Factor ETFs (1/3)



Factor loadings

	Mkt-RF	SMB	HML	RMW	CMA	MOM
MTUM	1.03	-0.12	0.01	-0.22	-0.05	0.36
VLUE	0.98	0.08	0.34	-0.06	0.09	-0.03
QUAL	0.97	-0.08	-0.04	0.15	0.03	-0.04
USMV	0.74	-0.11	-0.02	0.20	0.21	-0.02
IWM	1.04	0.85	0.08	-0.10	-0.02	0.03
IWD	0.94	-0.04	0.29	0.09	0.19	-0.04
IWF	1.03	-0.11	-0.28	-0.03	-0.12	0.00

MFM_ch3_etf_alphas

Alphas of Factor ETFs (2/3)

- The question: do the ETFs earn more than their exposures explain?
  - $R_t^e = \alpha + \boldsymbol{\beta}^\top \mathrm{f}_t + \varepsilon_t$, with the six factors (FF5 + MOM)
- Daily regressions, Newey–West errors
  - each ETF on its own trading days in 2000–2026
  - the chart: annualised alpha with a 95% interval ($0.02 = 2\%$) and the loadings
- The loadings match the labels
  - MTUM on MOM 0.36, VLUE on HML 0.34
  - IWM on SMB 0.85, IWF on HML -0.28
- USMV (minimum volatility) has a market beta of 0.74

MFM_ch3_etf_alphas

Alphas of Factor ETFs (3/3)

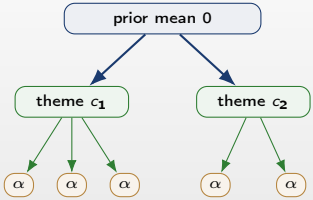
- Alphas close to zero
  - MTUM -0.2% ($t = -0.11$)
  - QUAL -0.3% ($t = -0.34$)
- IWM and IWD have negative alphas, individually significant at 5%; IWF has a positive, non-significant alpha
  - IWM -1.3% ($t = -2.57$), IWD -1.6% ($t = -2.54$)
  - fees (below 0.2% a year) explain only a small part
  - the periodic Russell reconstitution and construction differences with the Fama–French portfolios are possible explanations; the regression does not quantify them separately
- R^2 between 0.81 and 0.97
  - factor ETFs mostly sell exposure to known factors (Seminar 3, problem B11)

MFM_ch3_etf_alphas

Timing, Crowding and Costs

- **Factor timing**: moving capital between factors based on signals
  - for example how cheap a factor is or how it has done recently
  - Premia are uncertain and histories short, so it is hard to show that timing works out of sample (Chapter 13)
- **Crowding**: many investors hold the same factor positions
  - Premia can shrink (McLean & Pontiff (2016))
  - crowding can amplify losses during forced liquidations
  - Momentum crash: after a market fall the momentum portfolio has a negative beta
  - so it loses heavily when the market rebounds sharply, as in 2009 (Daniel & Moskowitz (2016))
- **Implementability**: academic factor returns do not deduct trading costs
  - nor the cost of borrowing the shorted stocks

153 Factors Grouped into 13 Themes



each factor's alpha borrows strength from the other factors of its theme

- **Data** (Section II of the paper): 153 factors in 93 countries
  - stocks are split into thirds by each characteristic, weighted by market capitalisation, with a cap
  - the factor buys the third with the higher return in the original study and sells the other extreme third
  - each factor is scaled to a specific volatility (uncorrelated with the market) of 10% a year, so that alphas are comparable
- **Themes**: 13 clusters of US factors
  - hierarchical clustering: factors whose CAPM residuals are strongly correlated fall into the same theme
- **Model** (Section I): $f_{i,t} = \alpha_i + \beta_i r_t^m + \varepsilon_{i,t}$
  - $f_{i,t}$: return of factor i; r_t^m: market excess return; β_i: market exposure; $\varepsilon_{i,t}$: noise
  - $\alpha_i = c_j + s_n + w_i$: the component of theme j, of signal n (the characteristic, for example B/M) and the component specific to factor i (signal n built in a given country), all with prior mean 0

Shrinking Alphas and the Replication Rate

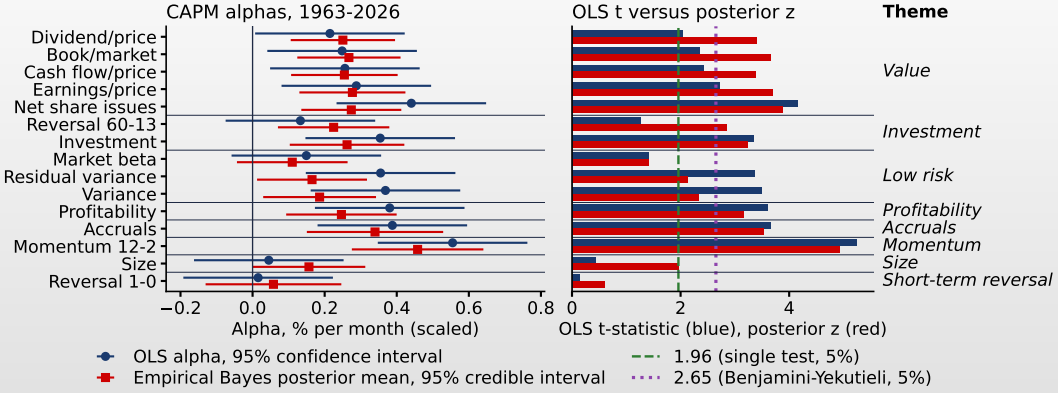
- The single-factor case: the estimated alpha is pulled towards zero

$$\mathbb{E}(\alpha \mid \hat{\alpha}) = \kappa \hat{\alpha}, \qquad \kappa = \frac{1}{1 + \sigma^2/(\tau^2 T)}$$

  - equation 5 of the paper; full hierarchy: Appendix
  - σ: residual volatility
  - τ: dispersion of the true alphas
  - T: number of months
  - short, noisy samples are shrunk more, as in Vasicek (Section 3)
  - σ^2/T plays the role of $\mathrm{se}(\hat{\beta}_i)^2$ and τ^2 that of $\sigma_{\bar{\beta}}^2$

- **Replication rate**: among the factors significant in the original studies
  - the share whose alpha is also significant in the new data
  - in the Bayesian version, a factor is "significant" if its posterior mean divided by its posterior standard deviation is at least 1.96

Do 15 US Factors Replicate? (1/2)



MFM_ch3_jkp_replication

Do 15 US Factors Replicate? Reading the Chart

- Left panel: the CAPM alpha of each factor, in % a month
  - scaled to 10% idiosyncratic volatility a year
- Blue: OLS alpha with its 95% confidence interval
- Red: the empirical-Bayes posterior mean with its 95% credible interval
- Right panel: OLS t (blue) and posterior z (red)
  - z = posterior mean divided by posterior standard deviation
- Dashed lines: 1.96 (a single test, 5%) and 2.65
  - 2.65: the Benjamini–Yekutieli hurdle, 5%
- 15 US long–short factors from value-weighted sorts of the Kenneth French Data Library
  - July 1963 – July 2026 ($T = 757$)
  - grouped by the themes of the paper (authors' cluster labels)

Do 15 US Factors Replicate? (2/2)

- Replication rates
  - OLS: 73.3% (11 of 15 factors)
  - Benjamini–Yekutieli: 53.3% (8 factors; critical $|t| = 2.65$)
  - Empirical Bayes: 86.7% (13 factors)
- Shrinkage pulls alpha towards the theme mean, not only towards zero
- Example: long-term reversal (investment theme) rises from 0.13% to 0.23% a month
  - its OLS t is 1.25, but its posterior z is 2.85
- Not replicated: market beta and short-term reversal ($z \leq 1.41$)
- Size only just passes ($z = 1.96$)
  - in the paper too, size is one of the two themes with a rate below 50%
- The rate here uses all 15 signals
  - without checking whether each was significant in its original study (Appendix)

MFM_ch3_jkp_replication

What the Course Replication Cannot Show

- **Limits of the course version**
  - 15 US signals instead of 153 in 93 countries
  - the portfolios of the Kenneth French Data Library, not thirds built without the smallest firms (micro-caps)
  - a single country, so each signal appears once and the signal component cannot be estimated (its dispersion is set to zero)
  - the result depends on the grouping into themes: if all factors were in one theme, the empirical-Bayes rate would be 100%
- **Limits stated by the authors** (Appendix)
  - the CAPM is the only benchmark
  - all characteristics are built by the same method
  - the prior ignores unpublished factors
- **Open questions** (the paper's data end in 2020)
  - did the alphas and the replication rate hold in 2021–2026?
  - can US factors, used as a prior, improve the estimates of BVB factors, whose histories are short?
- Data and code of the paper: jkpfactors.com

Project Design: Sample Dates and Deliverables

- **Deliverables**
  - first, an event table: factor, original paper with DOI, end of its sample, publication date, each read from the paper (model row: MOM, Jegadeesh & Titman (1993))
  - a pre-registration note (H_0, H_1, dates from the table, list of factors, corrections), saved in Git (a commit) before seeing the results
  - code as a Quantlet: one folder with the script, its data source and a short description, as in the chapter's Quantlets
  - a report with Newey–West errors, placebo periods and the power of the test
  - power: the probability of rejecting $b_2 = 0$ at 5% if the true decline were, for example, the 58% of McLean & Pontiff (2016)
  - power by simulation or approximately $\Phi(|\delta|/s - 1.96) + \Phi(-|\delta|/s - 1.96)$, $s = \mathrm{se}(\hat{b}_2)$ (δ: the assumed decline in the units of b_2, e.g. $-0.58\times$ the pre-publication mean; Φ: the distribution function of the standard Normal distribution)
- **Logs**: prompts in `AI_USE.md`, errors caught in `AI_ERRORS.md`

GRS under the Null Hypothesis

- Adding N assets always raises the in-sample Sharpe ratio
  - the test asks whether the rise exceeds what noise would produce
- Sharpe-ratio form: $W = \frac{T-N-1}{N} \frac{\widehat{SR}^2(f,\mathrm{R}) - \widehat{SR}^2(f)}{1+\widehat{SR}^2(f)}$
- Under H_0, with i.i.d. Normal residuals, $W \sim F(N, T-N-1)$, so $\mathbb{E}[W] = \frac{T-N-1}{T-N-3}$
- Multiplying by $\frac{N(1+\widehat{SR}^2(f))}{T-N-1}$: $\mathbb{E}\left[\widehat{SR}^2(f,\mathrm{R}) - \widehat{SR}^2(f)\right] \approx \frac{N\,(1+\widehat{SR}^2(f))}{T-N-3}$
- 25 portfolios, $T = 757$: $25 \times 1.018/729 = 0.035$ (monthly)
  - against $0.405^2 - 0.134^2 = 0.146$ observed
- 9 sector ETFs, $T = 331$: observed gain $0.191^2 - 0.146^2 = 0.015$
  - below the expected $9 \times 1.021/319 = 0.029$

Back

Tests with Many Assets

- $\hat{\boldsymbol{\Sigma}} = \hat{\mathbf{E}}^\top \hat{\mathbf{E}}/T$ has rank at most $T - 2$ (two parameters per regression)
  - $\hat{\mathbf{E}}$: the $T \times N$ matrix of residuals
  - Example: 100 size $\times$ B/M portfolios on 8 years of monthly data ($T = 96$)
  - then $\hat{\boldsymbol{\Sigma}}$ is singular, GRS does not exist (Seminar 3, problem B7)
- Tests for many assets that avoid $\hat{\boldsymbol{\Sigma}}^{-1}$: standardised sums of squared t-statistics, Pesaran & Yamagata (2024)
- Power enhancement by screening large alphas: Fan, Liao & Yao (2015)

Back

Misspecification-Robust Standard Errors

- Fama–MacBeth and Shanken errors assume a **correct** model
  - all pricing errors $e = \mu - X\gamma$ are zero ($X = [1, B]$)
- GRS rejects here the restrictions of the model with traded factors ($\alpha = 0$)
  - the cross-sectional relation with free premia may be only approximate: we estimate the pseudo-true value, which minimises the pricing errors
- The noise in $\hat{B}$ multiplies e and adds a variance term (Kan, Robotti & Shanken (2013))
- Write both passes as one exactly identified GMM system (Further Topics)
  - moments $(R_t - a - Bf_t) \otimes (1, f_t)$ and $X^\top(R_t - X\gamma)$
  - the sandwich variance with the full Jacobian is the robust one
  - Setting $e = 0$ in the Jacobian gives the correct-model, heteroskedasticity-robust version of Jagannathan & Wang (1998)
  - a moment matrix S without autocovariances is robust to misspecification, not to autocorrelation; with Newey–West lags it is also HAC
- FF5 on the 25 portfolios (Section 6), no lags: the SE of CMA rises from 2.2% to 2.7% (with 6 lags: RMW 3.0%, CMA 3.1%)
  - that of the intercept from 3.3% to 3.8%; for HML the change is below 0.05 percentage points
- Advanced course: Generalized Method of Moments

Back

Hansen–Jagannathan Distance and Bayesian Selection

- **Hansen–Jagannathan distance** (Hansen & Jagannathan (1997))

$$\delta^2 = \min_{\mathrm{b}} \; \mathrm{e(b)}^\top \, \mathbb{E}[\mathrm{R}^e \mathrm{R}^{e\top}]^{-1} \, \mathrm{e(b)}, \qquad \mathrm{e(b)} = \mathbb{E}[m_t(\mathrm{b}) \, \mathrm{R}_t^e]$$

  - how far the model's SDF is from the nearest SDF that prices all assets correctly
  - Excess returns R_t^e with the normalisation $\mathbb{E}[m] = 1$
  - For gross returns R_t: $\mathrm{e} = \mathbb{E}[m\mathrm{R}] - 1$, and the SDF needs a free scale ($\mathbb{E}[m] = 1/(1 + R_f)$)
  - The same weighting matrix for every model, so distances are comparable
  - The test statistic has a weighted χ^2 limit
- Selection among very many models: Bryzgalova, Huang & Julliard (2023)
  - they average over about 2×10^{15} (two quadrillion) factor models
  - robust to weak factors
- Kozak, Nagel & Santosh (2020) shrink the SDF coefficients
  - a few principal components of anomaly returns are enough
- Common lesson: compare all models with the same metric and report its uncertainty

Back

Reducing the Factor Zoo with Double-Selection LASSO

- Feng, Giglio & Xiu (2020): does a new factor add information given the existing zoo?
- **Double-selection LASSO** (Least Absolute Shrinkage and Selection Operator)
  - LASSO: a regression whose penalty sets the coefficients of unimportant variables exactly to zero, so it selects variables
  - Step 1: among hundreds of existing factors, LASSO selects those that explain the average returns of the test assets
  - Step 2: a second LASSO selects those correlated with the new factor, so that omitting them does not distort the result
  - Step 3: the SDF coefficient of the new factor is tested, controlling for all factors selected in steps 1 and 2
- Result: most new factors are redundant
- Only a few new factors explain returns that the old ones do not
- The case study reaches a compatible conclusion
  - the zoo is highly redundant, and a few themes carry the information

Back

Why the Factor-Number Estimators Disagree

- Onatski (2010): a factor is real if its eigenvalue stands clearly outside the range where the noise eigenvalues fall
- For standardised i.i.d. noise, the asymptotic spectral support (Marchenko–Pastur) is approximately $[(1-\sqrt{N/T})^2, (1+\sqrt{N/T})^2]$: an illustrative benchmark, not the Onatski test
  - $= [0.86, 1.15]$ for $N = 11,\ T = 2073$
  - the second eigenvalue, 1.06, lies inside
  - the benchmark assumes that every series is noise; here the first eigenvalue is 7.26, so it does not classify the second component directly
  - the Kaiser rule ($\ell_k > 1$) keeps two components, but it does not establish whether the second is significant
- Bai & Ng (2002): with $N = 11,\ T = 2073$, the IC_{p2} penalty is 0.219 per factor
- On the sectors, the criterion $IC_{p2}(k)$ for $k = 0, \ldots, 5$ is
  - 0, -0.86, -0.97, -1.12, -1.14, -1.17
- It falls all the way to the maximum allowed, $k_{\max} = 5$
  - the penalty is too small to stop weak factors from being added
- Report the range of estimates and the sensitivity to $k_{\max}$, not a single number
  - Seminar 3, problem B3

Back

Risk-Premium PCA and Three-Pass Premia

- **RP-PCA** (Risk-Premium Principal Component Analysis; Lettau & Pelger (2020)): principal components of $\frac{1}{T}\mathbf{X}^\top\mathbf{X} + \gamma\bar{\mathbf{x}}\bar{\mathbf{x}}^\top$
  - $\mathbf{X}$: the $T \times N$ matrix of excess returns, $\bar{\mathbf{x}}$: their means, γ: the weight on average returns
  - $\gamma = -1$ gives covariance PCA
  - $\gamma > -1$ also favours factors that explain average returns
  - In their data it finds factors with higher Sharpe ratios and smaller pricing errors than PCA
- **Three-pass estimator** (Giglio & Xiu (2021)): the premium of an observed factor g_t when the model omits factors
  - (1) PCA of the test-asset returns
  - (2) cross-sectional regression of the means on the latent betas
  - (3) regression of g_t on the latent factors; the premium of g_t follows from the latent premia and the coefficients of this regression
- Both keep the statistical basis of PCA but aim to explain average returns, not only variance
- IPCA also adds the characteristics of the assets

Back

IPCA and Crypto-Asset Factors

- IPCA links PCA (statistics) with characteristic sorts (Fama–French): time-varying betas, few factors
- Crypto-assets: market, size and momentum factors (Liu, Tsyvinski & Wu (2022))
  - they price the cross section of cryptocurrencies
- Dimension reduction on the return features of crypto and traditional assets: Pele et al. (2023), Jheng et al. (2026)
- Advanced course: Multivariate Statistical Analysis

Back

Building the 153 Factors

- **Data** (Section II of the paper)
  - 138 from the list of Hou, Xue & Zhang (2020), plus 15 from the wider literature
  - US returns from 1926, most developed markets from 1986
  - all in US dollars, in excess of the US Treasury bill
- **Construction**, each country and month
  - thirds by the characteristic, with breakpoints computed without the smallest firms (micro-caps)
  - "capped" value weights: market equity winsorised at the NYSE 80th percentile
  - one-month holding period
  - at least 5 stocks per leg and 60 monthly observations
- **Themes**: 13 clusters of US factors (Section II.B)
  - hierarchical clustering, with distance $1-$ correlation of CAPM residuals and Ward's method (it merges the groups that least increase the variance within groups)
  - Ward needs Euclidean distances: for standardised series, $d_{ij} = \sqrt{2(1-\rho_{ij})}$; authors: R `ward.D` on $1-\rho$; in SciPy, `ward` on $\sqrt{2(1-\rho)}$ (R equivalent: `ward.D2`)

Back

The Bayesian Hierarchy of Alphas

- **The hierarchy** (Section I of the paper)
  - $\alpha_i = c_j + s_n + w_i$ (theme j, signal n, factor i; eq. 21)
  - prior variance $\mathbf{\Omega} = \mathrm{M}\mathrm{M}^\top \tau_c^2 + \mathrm{Z}\mathrm{Z}^\top \tau_s^2 + \mathrm{I}\tau_w^2$ (eq. 23)
  - posterior mean $(\mathbf{\Omega}^{-1} + T\hat{\mathbf{\Sigma}}^{-1})^{-1} T\hat{\mathbf{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ (Proposition 4)
  - M, Z: 0/1 matrices linking each factor to its theme and signal
  - $\tau_c^2, \tau_s^2, \tau_w^2$: variances of the three alpha components
  - $\hat{\mathbf{\Sigma}}$: residual covariance
  - empirical Bayes: τ_c, τ_s, τ_w by maximum likelihood of $\hat{\boldsymbol{\alpha}} \sim N(0, \mathbf{\Omega} + \hat{\mathbf{\Sigma}}/T)$ (Section I.C.2)

Back

Complete Results

- With OLS, 98 of the 119 US factors replicate
- The Benjamini–Yekutieli correction requires $|t| > 2.7$ (Figure 4)
- Bayesian FDR: 0.1%, with 95% interval $[0, 1\%]$
- Expected share of true factors: 94% (Section III.C)
- Replication rate for mega and large stocks: 77.3% and 79.8% (Figure 5)
- 11 of 13 themes have a rate of at least 50%
  - the exceptions are low leverage and size
- The high rate among large firms shows that the result does not come only from the smallest firms
- Tangency Sharpe ratio (monthly), US 1972–2020: market 0.13, all factors 0.80 (Figure 14)

Back

The Course Replication in Detail

- Benjamini–Yekutieli gives the lowest rate, as in the US results of the paper (Figure 1)
- Net share issues are pulled towards the mean of their theme: from 0.44% to 0.27% a month
- The posterior standard deviation is 0.07%–0.10%, below the OLS SE of 0.11%
  - each factor borrows information from the other factors of its theme
- **Prior dispersion**: $\hat{\tau}_c = 0.28\%$, $\hat{\tau}_w = 0.04\%$ a month
  - against 0.35% and 0.21% in Table I of the paper
  - with 15 signals little factor-specific dispersion is left
- **Themes**: the 15 signals fall into 8 of the paper's 13 themes
  - 5 in value, 3 in low risk, 2 in investment

Back

Limits Stated by the Authors and Extensions

- **Limits stated by the authors**
  - The CAPM is the only benchmark model: low-leverage factors have no significant CAPM alpha, but contribute when combined with the other themes (Section III.A)
  - All characteristics are built by the same rule, sometimes differently from the original paper (Section II.C)
  - Factors tried but never published would shift the prior: in the scenarios of Harvey, Liu & Zhu (2016) the US rate falls to 81.5% and 79.8% (Figure 10)
- **Complete the course replication** (Seminar 3, C3)
  - justify each sign from the original paper
  - compare the dendrogram with the paper's themes
  - add the Bayesian FDR and the post-sample slope (Figure 8)
- Data and code of the paper: jkpfactors.com, github.com/bkelly-lab/ReplicationCrisis

Back

The Discovery Loop, Reusable Prompts (1/2)

- **Literature**
  - `Summarise the evidence since 2016 on post-publication decay and replication of anomaly returns. For each paper: DOI, sample, number of factors, headline estimate, identification of the decay. Flag entries you are unsure of.`
  - Resolve every DOI on Crossref
  - Check each headline number against the abstract
- **Hypothesis**
  - `Contrast data mining, publication-informed arbitrage and time-varying risk as explanations of decay. For each, describe the predicted pattern across the in-sample, post-sample and post-publication periods, and across high vs low limits-to-arbitrage stocks.`
- **Data and code**
  - `Load the monthly momentum factor from the Kenneth French Data Library; from 1965, regress MOM on a post-sample dummy (1990-01 to 1993-03) and a post-publication dummy (from 1993-04), with Newey-West errors.`
  - Replicate a known number first: momentum $t = 5.53$ before 2000 and 0.75 after (split-sample chart)

Back

The Discovery Loop, Reusable Prompts (2/2)

- **Robustness**
  - `Write a pre-registered robustness plan: one publication date per factor, a Holm correction across factors, excluding the 2009 momentum crash, the working-paper date instead of the journal date, and 1927-1962 as a placebo period for Mkt-RF, SMB, HML and MOM (RMW and CMA start in July 1963).`
- **Critique**
  - `Referee this: one factor, one publication date, t = -1.45. What is the power to detect a 58% decay? Why does pooling many anomalies change it?`
- **Report**
  - In-sample mean, both dummies with Newey–West errors, the placebo, the data vintage (download date)
  - Pooled factors: errors robust to cross-factor correlation within a month and to autocorrelation
  - Prompts that shaped a decision go into `AI_USE.md`
- The AI assistant proposes the robustness variants
- The researcher chooses in advance which ones are reported

Back

What the Researcher Must Check

- **Wrong event date**: sample end, working paper and journal issue are different dates
  - test: read them from the paper itself
- **Choosing the factor after seeing it decay**: fix the list of factors in advance
- Apply Holm or BH across all factors
- **Serially correlated, crash-prone returns**: Newey–West errors
- Re-estimate without 2009
- **Data vintages**: the factor library is revised retroactively
  - test: record the download date
  - reproduce the chapter's split-sample t-statistics
- **Low power**: with a single factor, H_0 is rarely rejected
- Report the confidence interval of the decay, not only its sign

Back

Answers to the Quick Checks (1/2)

Beta, error and shrinkage

- Classical: $0.90 \pm 1.96 \times 0.05 = [0.802, 0.998]$: the upper bound is very close to 1, but below it
- Newey–West: $0.90 \pm 1.96 \times 0.09 = [0.724, 1.076]$, contains 1: no evidence of a lower market sensitivity (total risk also includes the idiosyncratic part)
- Report Newey–West: residuals are heteroskedastic and autocorrelated; HAC changes the estimated variance, not the OLS estimate
- $\sigma^2_\beta = 0.10 - 0.02 = 0.08$; SE 0.2: $w = 0.08/0.12 = 0.67$, $\beta^V = 0.67 \times 1.6 + 0.33 \times 1.0 = 1.40$; SE 0.1: $w = 0.08/0.09 = 0.89$, $\beta^V = 1.53$
- The noisier estimate is shrunk more; Blume gives $0.33 + 0.67 \times 1.6 = 1.40$ whatever the precision
- Dimson beta: $0.25 + 0.55 + 0.00 = 0.80$; the stock price absorbs the market return with a one-day delay, so the same-day slope misses the 0.25 part

Reading a GRS statistic

- $W = \frac{60}{4} \times (0.25^2 - 0.15^2)/(1 + 0.15^2) = 15 \times 0.040/1.0225 = 0.59 < 2.53$: not rejected
- Expected gain $4 \times 1.0225/58 = 0.071$, above the observed 0.040: the gap is within noise
- $T = 650$: $W = \frac{645}{4} \times 0.0391 = 6.31 > 2.39$: rejected; the gap is the same, the evidence grows with T

Back

Answers to the Quick Checks (2/2)

Premia, Shanken and a negative R^2

- $\text{se}_{FM} = 3\%/\sqrt{400} = 0.15\%$, $t = 3.33$
- $c = 0.5^2/1^2 = 0.25$; $\text{Var}_{\text{Sh}} = 1.25 \times (0.0225 - 0.0025) + 0.0025 = 0.0275$, SE 0.166%, $t = 3.02$: only just above 3
- Traded factor: $c = 0.25/4.5^2 = 0.012$, so the correction is negligible
- Mean 4%: SST $= 4 + 0 + 4 = 8$, SSE $= 16 + 0 + 16 = 32$, $R^2 = 1 - 32/8 = -3$; with a free intercept, zero slopes already give SSE = SST, and OLS can only lower SSE

Counting discoveries

- Naive 4; Bonferroni (0.01) 1; Holm: $0.004 \leq 0.01$, $0.012 \leq 0.0125$, $0.020 > 0.0167$ stop: 2; BH thresholds $0.01j$: $0.030 \leq 0.040$ is the last success: 4
- BY: the sum is 2.283, thresholds $0.0044j$: only 0.004 passes: 1
- $1 - 0.95^5 = 0.23$
- A search over candidates: BH (FDR)
- Holm or Bonferroni (FWER) limit to 5% the probability of at least one false rejection

Back

Further Study: Courses and Code

Advanced courses on quantinar.com

- Generalized Method of Moments
- Multivariate Statistical Analysis

Code and data

- All charts and tables: github.com/danpele/MFM/tree/main/Quantlets/Ch_03
- Factor data: Kenneth French Data Library

Back

References (1/6)

- Ahn, S.C., Horenstein, A.R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3), 1203–1227.
- Anatolyev, S., Mikusheva, A. (2022). Factor models with many assets: strong factors, weak factors, and the two-pass procedure. *Journal of Econometrics*, 229(1), 103–126.
- Ang, A., Liu, J., Schwarz, K. (2020). Using stocks or portfolios in tests of factor models. *Journal of Financial and Quantitative Analysis*, 55(3), 709–750.
- Bai, J., Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.
- Barillas, F., Shanken, J. (2018). Comparing asset pricing models. *Journal of Finance*, 73(2), 715–754.
- Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B*, 57(1), 289–300.
- Benjamini, Y., Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29(4), 1165–1188.
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, 45(3), 444–455.
- Blume, M.E. (1971). On the assessment of risk. *Journal of Finance*, 26(1), 1–10.
- Bryzgalova, S., Huang, J., Julliard, C. (2023). Bayesian solutions for the factor zoo: we just ran two quadrillion models. *Journal of Finance*, 78(1), 487–557.

References (2/6)

- Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997). The Econometrics of Financial Markets. Princeton University Press.
- Carhart, M.M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82.
- Chen, A.Y., Zimmermann, T. (2022). Open source cross-sectional asset pricing. *Critical Finance Review*, 11(2), 207–264.
- Cochrane, J.H. (2011). Presidential address: discount rates. *Journal of Finance*, 66(4), 1047–1108.
- Daniel, K., Moskowitz, T.J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2), 221–247.
- Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2), 197–226.
- Ehsani, S., Linnainmaa, J.T. (2022). Factor momentum and the momentum factor. *Journal of Finance*, 77(3), 1877–1919.
- Fama, E.F., French, K.R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2), 427–465.
- Fama, E.F., French, K.R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- Fama, E.F., French, K.R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1–22.
- Fama, E.F., MacBeth, J.D. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy*, 81(3), 607–636.
- Fan, J., Liao, Y., Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica*, 83(4), 1497–1541.

References (3/6)

- Feng, G., Giglio, S., Xiu, D. (2020). Taming the factor zoo: a test of new factors. *Journal of Finance*, 75(3), 1327–1370.
- Frazzini, A., Pedersen, L.H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1), 1–25.
- Gibbons, M.R., Ross, S.A., Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5), 1121–1152.
- Giglio, S., Xiu, D. (2021). Asset pricing with omitted factors. *Journal of Political Economy*, 129(7), 1947–1990.
- Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4), 1029–1054.
- Hansen, L.P., Jagannathan, R. (1997). Assessing specification errors in stochastic discount factor models. *Journal of Finance*, 52(2), 557–590.
- Harvey, C.R., Liu, Y., Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.

References (4/6)

- Hou, K., Xue, C., Zhang, L. (2015). Digesting anomalies: an investment approach. *Review of Financial Studies*, 28(3), 650–705.
- Hou, K., Xue, C., Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies*, 33(5), 2019–2133.
- Jagannathan, R., Wang, Z. (1998). An asymptotic theory for estimating beta-pricing models using cross-sectional regression. *Journal of Finance*, 53(4), 1285–1309.
- Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance*, 45(3), 881–898.
- Jegadeesh, N., Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48(1), 65–91.
- Jensen, M.C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance*, 23(2), 389–416.
- Jensen, T.I., Kelly, B., Pedersen, L.H. (2023). Is there a replication crisis in finance? *Journal of Finance*, 78(5), 2465–2518.
- Jheng, S.-L., Găman, Ș., Conda, A., Pele, D.T., Härdle, W.K., Teng, H.-W., Osterrieder, J. (2026). Crypto currency returns. SSRN Working Paper.
- Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141–151.
- Kan, R., Robotti, C., Shanken, J. (2013). Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance*, 68(6), 2617–2649.
- Kan, R., Zhang, C. (1999). Two-pass tests of asset pricing models with useless factors. *Journal of Finance*, 54(1), 203–235.
- Kelly, B.T., Pruitt, S., Su, Y. (2019). Characteristics are covariances: a unified model of risk and return. *Journal of Financial Economics*, 134(3), 501–524.

References (5/6)

- Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics*, 149(2), 149–173.
- Kozak, S., Nagel, S., Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2), 271–292.
- Lettau, M., Pelger, M. (2020). Factors that fit the time series and cross-section of stock returns. *Review of Financial Studies*, 33(5), 2274–2325.
- Lewellen, J., Nagel, S., Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, 96(2), 175–194.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1), 13–37.
- Liu, Y., Tsyvinski, A., Wu, X. (2022). Common risk factors in cryptocurrency. *Journal of Finance*, 77(2), 1133–1177.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- McLean, R.D., Pontiff, J. (2016). Does academic research destroy stock return predictability? *Journal of Finance*, 71(1), 5–32.
- Moskowitz, T.J., Ooi, Y.H., Pedersen, L.H. (2012). Time series momentum. *Journal of Financial Economics*, 104(2), 228–250.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Newey, W.K., West, K.D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, 61(4), 631–653.
- Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *Review of Economics and Statistics*, 92(4), 1004–1016.

References (6/6)

- 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.
- Pesaran, M.H., Yamagata, T. (2024). Testing for alpha in linear factor pricing models with a large number of securities. *Journal of Financial Econometrics*, 22(2), 407–460.
- Politis, D.N., Romano, J.P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313.
- Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2), 129–176.
- Ross, S.A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341–360.
- Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies*, 5(1), 1–33.
- Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3), 425–442.
- Vasicek, O.A. (1973). A note on using cross-sectional information in Bayesian estimation of security betas. *Journal of Finance*, 28(5), 1233–1239.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838.
- 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
- **AMEX**: American Stock Exchange
- **APT**: Arbitrage Pricing Theory
- **BAB**: Betting Against Beta
- **BET**: Bucharest Exchange Trading (index)
- **BET-TR**: BET Total Return (index)
- **B/M**: Book-to-Market (ratio)
- **BH**: Benjamini–Hochberg (procedure)
- **BM**: Book-to-Market (B/M quintile in the size–value chart)
- **BVB**: Bursa de Valori București — *the Bucharest Stock Exchange*
- **BY**: Benjamini–Yekutieli (procedure)
- **CAGR**: Compound Annual Growth Rate
- **CAPM**: Capital Asset Pricing Model
- **CI**: Confidence Interval
- **CMA**: Conservative Minus Aggressive (investment factor)
- **CML**: Capital Market Line
- **DOI**: Digital Object Identifier
- **ETF**: Exchange-Traded Fund
- **FDR**: False Discovery Rate
- **FF3**: Fama–French three-factor model
- **FF5**: Fama–French five-factor model

Acronyms used in this material (2/3)

- **FWER**: Family-Wise Error Rate
- **GICS**: Global Industry Classification Standard
- **GLS**: Generalised Least Squares
- **GMM**: Generalized Method of Moments
- **GRS**: Gibbons–Ross–Shanken (test)
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **HML**: High Minus Low (value factor)
- **IPCA**: Instrumented Principal Component Analysis
- **LASSO**: Least Absolute Shrinkage and Selection Operator
- **ME**: Market Equity (market capitalisation)
- **MOM**: Momentum (factor)
- **NASDAQ**: National Association of Securities Dealers Automated Quotations
- **NYSE**: New York Stock Exchange
- **OLS**: Ordinary Least Squares
- **OMV**: Österreichische Mineralölverwaltung — *Austrian Mineral Oil Administration (OMV group)*
- **PC1**: first Principal Component
- **PC2**: second Principal Component
- **PC3**: third Principal Component
- **PCA**: Principal Component Analysis

Acronyms used in this material (3/3)

- **PRDS**: Positive Regression Dependence on a Subset (of p-values)
- **RF**: Risk-Free rate
- **RMSE**: Root Mean Squared Error
- **RMW**: Robust Minus Weak (profitability factor)
- **ROE**: Return On Equity
- **RP-PCA**: Risk-Premium Principal Component Analysis
- **SDF**: Stochastic Discount Factor
- **SE**: Standard Error
- **SMB**: Small Minus Big (size factor)
- **SML**: Security Market Line
- **SPDR**: Standard & Poor's Depositary Receipts
- **SSE**: Sum of Squared Errors
- **SST**: Total Sum of Squares
- **TSFM**: Time-Series Factor Momentum
- **UMD**: Up Minus Down (momentum factor)
- **US**: United States