








Modelarea piețelor financiare

Seminarul 3: Modele factoriale și evaluarea activelor — aplicații



Daniel Traian PELE
Antoaneta Amza

Academia de Studii Economice din București
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

De la întrebare la aplicații: ce urmărim

- Întrebările seminarului
  - **explică beta diferențele dintre randamentele medii ale activelor?**
  - **cît de precise sînt estimațiile?**
- Livrabile
  - reproduceți exercițiile rezolvate A/B și interpretați rezultatul
  - exercițiile propuse: predați demonstrația, tabelul sau figura cerută
  - codul corectat din C1
  - o frază de interpretare pentru fiecare exercițiu
- Notebook-ul seminarului: capitol3_notebook_seminar.ipynb (Colab); prețuri: data/market
  - fiecare cerință indică secțiunea din notebook
- Pentru exercițiile A1–C1 aveți nevoie de regresie, CAPM (*Capital Asset Pricing Model*, modelul de evaluare a activelor financiare), inferență și testare multiplă
- Analiza componentelor principale (PCA), beta Dimson, regresiile Fama–MacBeth și metoda celor mai mici pătrate generalizate (GLS) servesc exercițiilor de după curs (B3, B4) și pachetelor de extindere din hartă
- Seminar: prof. Daniel Traian Pele, danpele@ase.ro; drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

Ce reprezintă seriile de factori?

ETF	Exchange-traded fund: fond tranzacționat la bursă; SPY urmărește Standard & Poor's 500 (S&P 500).
RF	Risk-free: rata fără risc.
Mkt-RF	Market minus risk-free: randamentul pieței americane de acțiuni minus RF.
B/M	Book-to-market: valoarea contabilă / valoarea de piață.
SMB	Small minus big: firme mici minus mari.
HML	High minus low: B/M mare minus mic.
RMW	Robust minus weak: profitabilitate ridicată minus scăzută.
CMA	Conservative minus aggressive: investiții conservatoare minus agresive.
MOM	Momentum: cîștigători anteriori minus perdanți anteriori.
CAPM	Capital Asset Pricing Model: modelul cu un singur factor, Mkt-RF.
FF3, FF5	Fama–French: modele cu trei, respectiv cinci factori.

Cum citim simbolurile de la Bursa de Valori București?

BVB: Bursa de Valori București; BET (Bucharest Exchange Trading): indicele pieței; BET-TR (total return): varianta cu dividende reinvestite.

TLV	Banca Transilvania	SNP	OMV Petrom
BRD	BRD Groupe Société Générale	TGN	Transgaz
SNG	Romgaz	SNN	Nuclearelectrica
H2O	Hidroelectrica	EL	Electrica
TEL	Transelectrica	FP	Fondul Proprietatea
DIGI	Digi Communications		

Datele folosite

Serie	Sursă, tip de preț	Frecvență	Eșantion
9 ETF-uri sectoriale (XLB, XLE, XLF, XLI, XLK, XLP, XLU, XLV, XLY) și SPY	închidere ajustată	sfîrșit de lună	ian. 1999 – iul. 2026
11 ETF-uri sectoriale (cele 9 + XLRE, XLC), SPY	închidere ajustată	zilnic	20 iun. 2018 – 18 sept. 2026
Mkt-RF, SMB, HML, RMW, CMA, MOM, RF	Kenneth French Data Library, în procente	lunar, zilnic	iul. 1963 – iul. 2026
25 portofolii mărime × B/M	Kenneth French Data Library, ponderate după valoare	lunar	iul. 1963 – iul. 2026
30 portofolii pe industrii, 100 portofolii mărime × B/M, portofolii sortate după o caracteristică	Kenneth French Data Library, ponderate după valoare	lunar	iul. 1963 – iul. 2026
7 ETF-uri factoriale (MTUM, VLUE, QUAL, USMV, IWM, IWD, IWF)	închidere ajustată	sfîrșit de lună; zilnic (MTUM, QUAL)	2013 – iul. 2026
11 blue chips BVB, BET	acțiuni: închidere ajustată; BET: închidere	zilnic	6 ian. 2015 (sau listare) – 18 sept. 2026

- Prețurile zilnice din datele cursului, cîte o serie pe activ (cîmpul folosit: închiderea ajustată, *adjusted close*, pentru ETF-uri și acțiuni; închiderea, *close*, pentru BET)
- Analize pe mai multe active: întîi aliniați *prețurile* pe zilele comune, apoi calculați randamentele; randamentele lunare din prețurile de sfîrșit de lună
- Sectoarele (fondurile *Select Sector SPDR*): XLB (materiale), XLE (energie), XLF (financiar), XLI (industrie), XLK (tehnologie), XLP (bunuri de larg consum), XLU (utilități), XLV (sănătate), XLY (consum discreționar), XLRE (imobiliare), XLC (servicii de comunicații)

Pregătire: de la prețuri la randamente în exces

```
# notebook section "Setup" defines
# prices(), factors(), monthly_excess()
ex, F = monthly_excess(['XLK.US',
                        'SPY.US'])
# month-end adjusted prices -> simple
# returns -> minus RF -> align with factors
print(len(ex), ex.index[0].date(),
      ex.index[-1].date())
print(ex['XLK.US'].head(3))
print(F[['Mkt-RF', 'RF']].head(3))
```

- Rezultat așteptat: $T = 331$ luni, 1999-01-31 – 2026-07-31
- Dacă numărul diferă, verificați cîmpul, regula de sfîrșit de lună și împărțirea la 100

	ian. 1999	**feb. 1999**	**mar. 1999**
$P_{t-\mathbf{1}}$, XLK	12,1103	14,0359	12,6439
P_t, XLK	14,0359	12,6439	13,5835
$R_t = P_t/P_{t-\mathbf{1}} - 1$	15,90%	−9,92%	7,43%
RF (/100)	0,35%	0,35%	0,43%
$R^e_t = R_t - \mathrm{RF}_t$	15,55%	−10,27%	7,00%
Mkt-RF	3,47%	−4,08%	3,48%

- Un rînd calculat: $14{,}0359/12{,}1103 - 1 = 15{,}90\%$; scăzînd 0,35% rezultă 15,55%
- RF (rata fără risc) este deja o rată, iar Mkt-RF este deja un randament în exces

Notații folosite în toate problemele (1/2): randamente și portofolii

- **Randamentul simplu** și **randamentul în exces** al activului i în perioada t

$$R_{i,t} = \frac{P_{i,t}}{P_{i,t-\mathbf{1}}} - 1, \qquad R^e_{i,t} = R_{i,t} - R_{f,t}$$

  - $P_{i,t}$: prețul ajustat (dividende reinvestite); $R_{f,t}$: rata fără risc a perioadei (RF)
  - f_t: un factor, el însuși un randament în exces (Mkt-RF, dacă nu se precizează altfel); R^e_t: vectorul $N \times 1$ al celor N active

- **Randamentul unui portofoliu**: ponderile w_i se aplică randamentelor *simple*

$$R_{p,t} = \sum_{i=\mathbf{1}}^{N} w_i R_{i,t}$$

  - randamentele log $r_t = \ln(1 + R_t)$ se adună în *timp*, cele simple între *active* (Capitolul 0)
  - long–short: $+1/n$ pe n active cumpărate, $-1/n$ pe n active vîndute în lipsă

- Anualizare: $\times 12$ (medii), $\times\sqrt{12}$ (volatilități și rapoarte Sharpe); zilnic 252

Notații folosite în toate problemele (2/2): momente, matrice, secțiuni transversale

- Randamentele în exces așteptate și covarianța celor N active

$$\boldsymbol{\mu} = \mathbb{E}[\mathrm{R}_t^e]\ (N \times 1), \qquad \boldsymbol{\Sigma} = \mathrm{Cov}(\mathrm{R}_t^e)\ (N \times N)$$

  - $\boldsymbol{\Sigma}$ pozitiv definită: orice portofoliu are variață pozitivă, deci $\boldsymbol{\Sigma}^{-1}$ există
  - 1: vectorul cu toate elementele egale cu 1; I: matricea identitate; $^\top$: transpusa; $\mathrm{a}^\top \boldsymbol{\Sigma}^{-1} \mathrm{a}$: o formă pătratică, o distanță la pătrat care ține cont de corelații

- Momentele estimate prin verosimilitate maximă (maximum likelihood, ML) se calculează cu împărțire la T

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T} \hat{\mathrm{E}}^\top \hat{\mathrm{E}}, \qquad \hat{\sigma}_f^2 = \frac{1}{T} \sum_{t=1}^{T} (f_t - \bar{f})^2$$

  - $\hat{\mathrm{E}}$: matricea $T \times N$ a reziduurilor regresiilor; $\bar{f}$: media în timp a factorului; excepție: covarianța factorilor din B8 se calculează cu împărțire la $T - 1$

- Momente **transversale**: între cele N active, la un moment dat, nu în timp
  - $\mathrm{Var}_{cs}(\hat{\beta}) = \frac{1}{N-1} \sum_i (\hat{\beta}_i - \bar{\beta})^2$ (estimatorul nedeplasat), $\bar{\beta} = \frac{1}{N} \sum_i \hat{\beta}_i$; analog Cov_{cs}
  - $\otimes$ (produsul Kronecker, B7): A $\otimes$ B înlocuiește fiecare element a_{jk} al lui A cu blocul a_{jk}B

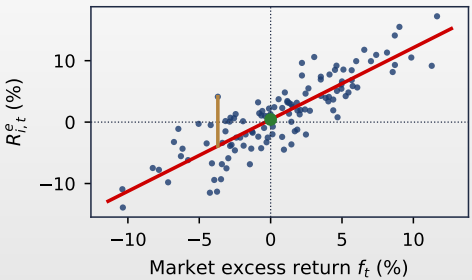
Modelele factoriale

- Detalii: Capitolul 3; manuale: Campbell, Lo & MacKinlay (1997) și, pentru componentele principale, Härdle & Simar (2019)
- **Modelul factorial**: randamentul în exces al unui activ este explicat de cîțiva factori comuni

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

  - $f_{k,t}$: factorul k, un randament în exces (un portofoliu tranzacționat); K: numărul de factori
  - $\beta_{i,k}$: încărcarea (*loading*), adică expunerea activului i la factorul k; $\varepsilon_{i,t}$: șocul idiosincratic, cu media 0
  - α_i: partea din randamentul mediu pe care factorii nu o explică; un model corect are $\alpha_i = 0$ pentru orice activ

Modelul de piață și metoda celor mai mici pătrate



- **Modelul de piață**: un singur factor, f_t = Mkt-RF

$$R^e_{i,t} = \alpha_i + \beta_i f_t + \varepsilon_{i,t}$$

- Estimațiile OLS (ordinary least squares, metoda celor mai mici pătrate), care minimizează $\sum_t \hat{\varepsilon}_t^2$

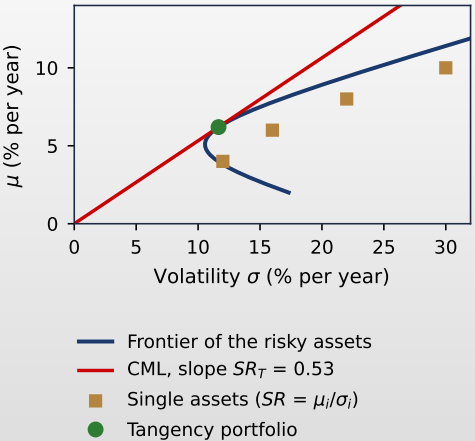
$$\hat{\beta}_i = \frac{\sum_t (f_t - \bar{f})(R^e_{i,t} - \bar{R}^e_i)}{\sum_t (f_t - \bar{f})^2}, \qquad \hat{\alpha}_i = \bar{R}^e_i - \hat{\beta}_i \bar{f}$$

- Interpretare
  - $\hat{\beta}_i$: panta; $\beta > 1$: activul amplifică mișcările pieței
  - $\hat{\alpha}_i$: termenul liber, randamentul mediu în exces într-o lună cu $f_t = 0$
  - $R^2 = 1 - \sum_t \hat{\varepsilon}_t^2 / \sum_t (R^e_{i,t} - \bar{R}^e_i)^2$: partea din varianță explicată de piață

Factorii Fama–French

- Factorii sînt portofolii long–short sortate după o caracteristică (Fama & French (1993); Fama & French (2015); Carhart (1997)); date: Kenneth French Data Library
  - Mkt-RF este ponderat cu capitalizarea; HML (*High Minus Low*, factorul valoare: B/M mare minus mic) opune acțiunile de valoare celor de creștere; RMW (*Robust Minus Weak*, factorul profitabilitate: firme robuste minus slabe) sortează după profitabilitatea operațională, CMA (*Conservative Minus Aggressive*, factorul investiții: firme conservatoare minus agresive) după creșterea activelor
  - MOM: momentum, semnalul 12–1 (slide-ul „Momentum și deplasarea de anticipare")
- Modelele
  - CAPM: Mkt-RF; FF3: Mkt-RF, SMB (*Small Minus Big*, factorul mărime: firme mici minus mari), HML; Carhart: FF3 + MOM; FF5 (*Fama–French five-factor model*, modelul Fama–French cu cinci factori): Mkt-RF, SMB, HML, RMW, CMA
  - regresia FF3: $R^e_{i,t} = \alpha_i + b_i\,\text{Mkt-RF}_t + s_i\,\text{SMB}_t + h_i\,\text{HML}_t + \varepsilon_{i,t}$
- Activele de test
  - 25 de portofolii mărime × B/M: acțiunile sortate în 5 grupe de mărime × 5 grupe de book-to-market
    - sortările se fac în fiecare iunie, independent, cu pragurile chintilelor NYSE (*New York Stock Exchange*, Bursa din New York)
    - fiecare intersecție este un portofoliu ponderat după capitalizare, deținut din iulie pînă în iunie
    - randamentele medii diferă mult între portofolii: 3,9% pe an la firmele mici de creștere (small-growth), 13,3% la firmele mici de valoare (small-value)
  - ETF-urile sectoriale dețin un sector, cele factoriale urmăresc un factor (MTUM: momentum, QUAL: calitate)

Medie–variață și portofoliul tangent

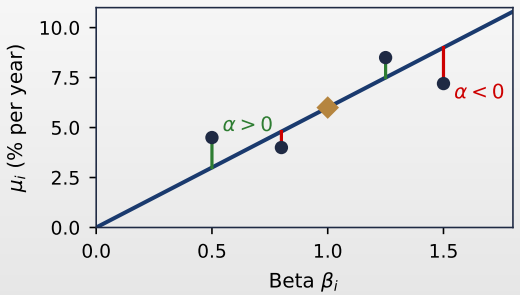


- **Raportul Sharpe**: randamentul mediu în exces pe unitatea de volatilitate, $SR = \mu/\sigma$
  - investitorii medie–variață (Markowitz (1952)) care au acces la un activ fără risc dețin toți portofoliul riscant cu cel mai mare SR: **portofoliul tangent**
- Pentru $\boldsymbol{\Sigma}$ pozitiv definită și $1^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} > 0$, ponderile tangente și pătratul raportului Sharpe

$$w_T = \frac{\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}{1^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}}, \qquad SR_T^2 = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$$

- **CML** (*capital market line*, dreapta pieței de capital): activul fără risc combinat cu portofoliul tangent
  - graficul: patru active, SR între 0,33 și 0,38; $SR_T = 0{,}53$
  - exemplu: două active necorelate, $\boldsymbol{\mu} = (4\%;\ 6\%)$, variațele 0,04 și 0,09: $w_T = (0{,}60;\ 0{,}40)$, $SR_T^2 = 0{,}04 + 0{,}04$, $SR_T = 0{,}283$

CAPM și dreapta pieței titlurilor



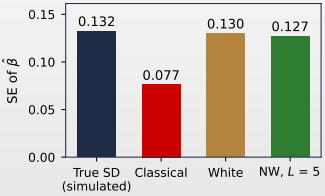
- **CAPM** (Sharpe (1964)): dacă toți dețin portofoliul tangent, acesta este piața și, pentru orice activ

$$\mu_i = \beta_i \mu_m, \qquad \beta_i = \frac{\mathrm{Cov}(R_i^e, R_m^e)}{\mathrm{Var}(R_m^e)}$$

  - $\mu_i = \mathbb{E}[R_i^e]$, $\mu_m = \mathbb{E}[R_m^e]$: randamentele în exces așteptate ale activului și ale pieței
  - SML (security market line, dreapta pieței titlurilor): doar riscul de piață aduce o primă de risc; $\varepsilon_{i,t}$ dispare prin diversificare
- **Alfa**: distanța pe verticală față de SML
  - consecința testabilă: $\alpha_i = 0$ pentru orice activ
  - exemplu: $\beta = 1{,}2$, prima de 6% pe an: CAPM dă 7,2%; un randament realizat de 9% dă $\alpha = 1{,}8\%$
- O respingere poate însemna că aproximarea pieței nu este eficientă (A2)
  - CAPM cu beta zero (Black (1972)): $\mu_i - \gamma = \beta_i(\mu_m - \gamma)$, γ: media unui portofoliu cu beta zero

Erorile standard clasice, White și Newey–West

SD (*standard deviation*): abaterea standard; SE (*standard error*): eroarea standard.



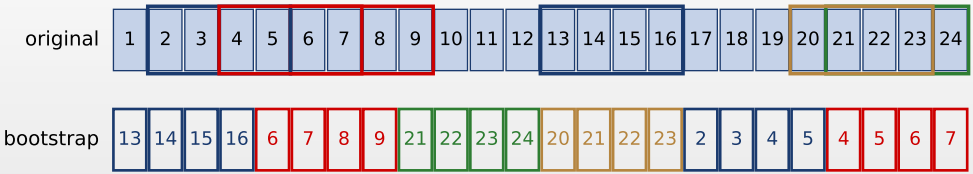
- Graficul: 2 000 de selecții simulate, cu reziduuri autocorelate și mai mari în lunile volatile
  - eroarea standard clasică (0,077) subestimează abaterea standard reală a lui $\hat{\beta}$ (0,132)
  - erorile White și Newey–West sînt apropiate de valoarea reală

- OLS pentru $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\beta})^\top$ cu regresorii $\mathrm{x}_t = (1, f_t)^\top$; X: rîndurile $\mathrm{x}_t^\top$; $\hat{\varepsilon}_t$: reziduurile

$$\widehat{\mathrm{Var}}(\hat{\boldsymbol{\theta}}) = (\mathrm{X}^\top \mathrm{X})^{-1} \hat{\mathrm{S}} (\mathrm{X}^\top \mathrm{X})^{-1}$$

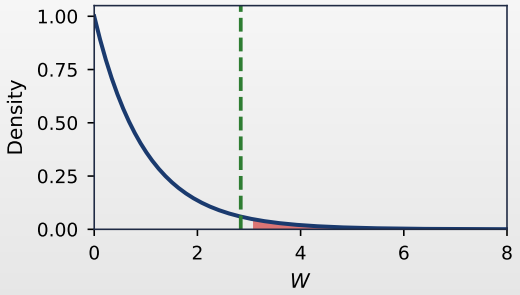
- Variante pentru $\hat{\mathrm{S}}$
  - clasică: $\hat{\sigma}_\varepsilon^2 \mathrm{X}^\top \mathrm{X}$ (varianță constantă, fără autocorelație)
  - White (White (1980)): $\sum_t \hat{\varepsilon}_t^2 \mathrm{x}_t \mathrm{x}_t^\top$, robustă la heteroscedasticitate
  - Newey–West (NW; Newey & West (1987)), estimator HAC (*heteroskedasticity and autocorrelation consistent*, robust la heteroscedasticitate și autocorelație): adaugă $\sum_{l=1}^{L} \left(1 - \frac{l}{L+1}\right) \sum_t \hat{\varepsilon}_t \hat{\varepsilon}_{t-l} (\mathrm{x}_t \mathrm{x}_{t-l}^\top + \mathrm{x}_{t-l} \mathrm{x}_t^\top)$
  - regula pentru laguri $L = \lfloor 4(T/100)^{2/9} \rfloor$; $T = 331$: $L = 5$

Intervale bootstrap



- **Bootstrap pe blocuri suprapuse** (Künsch (1989)): extragem blocuri de b luni consecutive (aici $b = 4$) și le concatenăm pînă la lungimea T
  - reeșantionăm perechile (R_t^e, f_t) împreună; reestimăm; repetăm de B ori; cuantilele de 2,5% și 97,5% formează un interval percentil de 95%
  - dependența din interiorul fiecărui bloc (autocorelația, gruparea volatilității – *volatility clustering*) se păstrează; începuturile blocurilor se pot suprapune
- Variante
  - **bootstrap staționar** (Politis & Romano (1994)): lungimi aleatoare ale blocurilor, cu distribuție geometrică de medie b; bootstrap **i.i.d.** (extrageri independente și identic distribuite): luni izolate ($b = 1$)
  - bootstrap *sub* H_0: reconstruim datele impunînd ipoteza nulă; $p = (1 + \#\{W^* \geq W\})/(B + 1)$

Testul GRS



- **GRS** (testul Gibbons–Ross–Shanken, Gibbons, Ross & Shanken (1989)): H_0: $\alpha_1 = \cdots = \alpha_N = 0$, simultan

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

- Notațiile
  - N: numărul activelor de test; T: numărul de luni; $\hat{\boldsymbol{\alpha}}$: cele N valori alfa; $\hat{\boldsymbol{\Sigma}}$: covarianța ML a reziduurilor
  - $SR_m = \hat{\mu}_m / \hat{\sigma}_m$: raportul Sharpe al factorului; F: distribuția Fisher, exactă pentru erori i.i.d. cu distribuția Normală
- Interpretare
  - W mare: valorile alfa sînt departe de 0 față de zgomotul lor; respingem cînd W depășește cuantila de 95% a lui F
  - exemplu: $T = 121$, $N = 2$, $\hat{\boldsymbol{\alpha}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} = 0{,}05$, $SR_m = 0{,}2$: $W = 59 \times 0{,}05/1{,}04 = 2{,}84 < 3{,}07$, $p = 0{,}062$

Normalitatea reziduurilor

- Distribuția F a statisticii GRS presupune reziduuri cu distribuția Normală (B7)
  - cozile groase schimbă distribuția lui W; bootstrap-ul sub H_0 (slide-ul „Intervale bootstrap") nu cere distribuția Normală, dar presupune reziduuri i.i.d., independente de factor
- **Kurtosis-ul Pearson** al unei serii de reziduuri

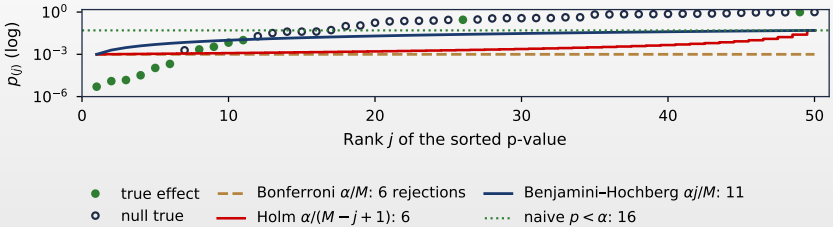
$$\text{Kurt} = \frac{\frac{1}{T}\sum_t \hat{\varepsilon}_t^4}{\hat{\sigma}^4}, \qquad \hat{\sigma}^2 = \frac{1}{T}\sum_t \hat{\varepsilon}_t^2$$

  - Kurt $= 3$ pentru distribuția Normală; Kurt > 3: cozi groase, lunile extreme sînt mai frecvente
- **Graficul Q–Q** (quantile–quantile, cuantilă–cuantilă): reziduurile standardizate ordonate, în funcție de cuantilele distribuției Normale
  - puncte pe prima bisectoare: distribuție Normală; o formă de S: cozi groase (Seminarul 1)
- **Testul Wald**, forma generală: un vector de estimații $\hat{\theta}$ cu covarianța $\hat{V}$

$$W = \hat{\theta}^\top \hat{V}^{-1} \hat{\theta} \to \chi^2_q \text{ sub } H_0 : \theta = 0$$

  - q: numărul restricțiilor; χ^2_q: distribuția hi-pătrat cu q grade de libertate
- **Ce credeți**: două valori alfa au $t_1 = 1{,}5$ și $t_2 = -1{,}5$, iar reziduurile celor două active au corelația $\rho = 0{,}9$; respinge testul Wald comun H_0: $\alpha_1 = \alpha_2 = 0$ la 5%?
  - **Răspuns**: da: $W = (t_1^2 + t_2^2 - 2\rho\, t_1 t_2)/(1-\rho^2) = 8{,}55/0{,}19 = 45{,}0$, mult peste 5,99, cuantila de 95% a lui χ^2_2
  - două active foarte corelate cu valori alfa de semne opuse: diferența valorilor alfa are o eroare standard mică

Testarea multiplă, FWER și FDR



- Cu M teste, unele valori p (*p-values*) sînt mici din întîmplare (sute de factori publicați, Harvey, Liu & Zhu (2016))
  - **FWER** (*family-wise error rate*, rata erorii pe familia de teste): probabilitatea de a avea cel puțin o respingere falsă
  - Bonferroni controlează FWER respingînd cînd $p_i \le \alpha/M$
  - Holm (Holm (1979)) compară, în ordine, $p_{(j)}$ cu $\alpha/(M-j+1)$ și se oprește la prima nerespingere
  - **FDR** (*false discovery rate*, rata descoperirilor false): proporția așteptată a descoperirilor false printre respingeri
  - **Benjamini–Hochberg** (BH; Benjamini & Hochberg (1995)): găsim cel mai mare j cu $p_{(j)} \le \alpha j/M$ și respingem $H_{(\mathbf{1})}, \ldots, H_{(j)}$; valid pentru teste independente sau pozitiv dependente
  - Benjamini–Yekutieli (BY; Benjamini & Yekutieli (2001)): pragurile BH împărțite la $\sum_{i=\mathbf{1}}^{M} 1/i$; valid pentru orice dependență
  - graficul: 12 efecte reale din 50 de teste; BH raportează 11 descoperiri (una falsă), Holm 6

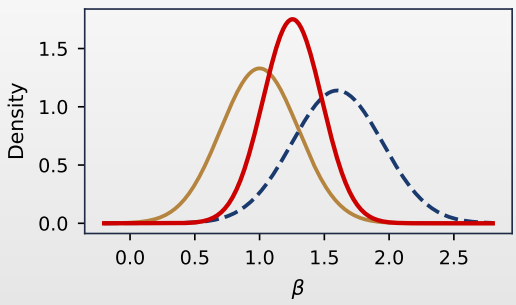
Momentum și deplasarea de anticipare

- **Momentum** (Jegadeesh & Titman (1993)): ordonăm acțiunile după randamentul trecut, cumpărăm cîștigătoarele, vindem perdantele
  - semnalul 12–1: randamentul compus al lunilor $t-12, \ldots, t-2$, cunoscut la sfîrșitul lunii $t-1$

$$s_{i,t} = \prod_{k=2}^{12}(1 + R_{i,t-k}) - 1$$

  - luna $t-1$ se omite din cauza **inversării pe termen scurt** (*short-term reversal*): pe un orizont de o lună, randamentele se inversează (Jegadeesh (1990))
  - factorul MOM (Kenneth French Data Library): sortare 2×3 refăcută lunar, după mărime (mediana NYSE) și după $s_{i,t}$ (percentilele NYSE 30 și 70); MOM $= \frac{1}{2}(S_W + B_W) - \frac{1}{2}(S_L + B_L)$, cu S/B firme mici/mari și W/L cîștigători/perdanți
  - portofoliul se deține în luna t; randamentul lui este media randamentelor simple din luna t (Notații (1/2))
- **Deplasarea de anticipare** (*look-ahead bias*): semnalul folosește informație încă indisponibilă la formarea portofoliului
  - verificare: ponderile nu trebuie să se schimbe dacă se modifică vreun preț de după sfîrșitul lunii $t-1$
- Alte capcane ale unui backtest (C1, C2)
  - **deplasarea de supraviețuire** (*survivorship bias*): blue chips de azi sînt alese retrospectiv; perdantele ieșite din indice lipsesc
  - restricțiile la vînzarea în lipsă, costurile și lichiditatea reduc cîștigul realizabil; cu 3 acțiuni pe fiecare parte, o singură acțiune poate determina rezultatul unei luni

Contracția bayesiană a valorilor beta

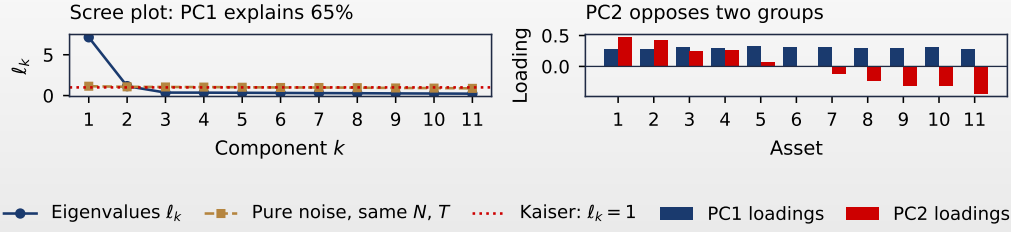


- A priori $\beta_i \sim N(\bar{\beta}, \sigma_\beta^2)$; estimația $\hat{\beta}_i \mid \beta_i \sim N(\beta_i, \text{se}_i^2)$
  - $\bar{\beta}$, σ_β^2: media și varianța reală a valorilor beta între active; se_i: eroarea standard a lui $\hat{\beta}_i$
- **Vasicek** (Vasicek (1973)): media a posteriori apropie $\hat{\beta}_i$ de $\bar{\beta}$ (contracție, *shrinkage*)

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

- Interpretare și estimare
  - $\hat{\beta}_i$ zgomotos (se_i mare): w_i mic, contracție puternică; graficul: $w = 0{,}42$
  - legea varianței totale: $\text{Var}(\hat{\beta}) = \text{Var}(\mathbb{E}[\hat{\beta} \mid \beta]) + \mathbb{E}[\text{Var}(\hat{\beta} \mid \beta)]$, folosită pentru estimarea lui σ_β^2 (A3)
  - **Blume** (Blume (1971)): valorile beta revin spre 1; regula practicienilor $0{,}33 + 0{,}67\hat{\beta}_i$ este o apropiere fixă de 1

Componentele principale



- PCA (*principal component analysis*, analiza componentelor principale); prima și a doua componentă principală (PC1, PC2)
- Standardizăm fiecare serie de randamente (Z: $T \times N$), apoi descompunem matricea de corelație

$$\mathrm{C} = \tfrac{1}{T}\mathrm{Z}^\top \mathrm{Z} = \sum_{k=1}^{N} \ell_k \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\top, \qquad \ell_1 \geq \cdots \geq \ell_N, \quad \textstyle\sum_k \ell_k = N$$

- ℓ_k: valoarea proprie, egală cu varianța componentei k (notată ℓ, ca să nu se confunde cu prima de risc λ); componenta explică ℓ_k / N din varianța totală
- $\boldsymbol{\gamma}_k$: vectorul propriu de normă 1 (încărcările, *loadings*); $\mathrm{Z}\boldsymbol{\gamma}_k$: scorurile componentei, adică seria ei în timp
  - graficul: 11 active simulate, un factor de piață și unul mai slab; PC1: toate încărcările pozitive; PC2 opune două grupe

Numărul de factori

- **Kaiser** (Kaiser (1960)): păstrăm componentele cu $\ell_k > 1$ (mai multă varianță decît un singur activ standardizat)
- **Bai–Ng** (Bai & Ng (2002)): criteriu informațional penalizat

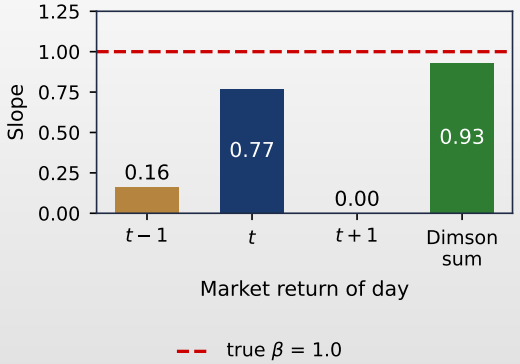
$$IC_{p2}(k) = \ln V(k) + k \frac{N+T}{NT} \ln \min(N, T), \qquad V(k) = \frac{1}{N} \sum_{j>k} \ell_j$$

  - $V(k)$: varianța rămasă după k factori; alegem $\hat{k} = \arg\min_{0 \le k \le k_{\max}} IC_{p2}(k)$
  - construit pentru N și T mari; cu $N = 11$ penalizarea este slabă și $\hat{k}$ tinde spre $k_{\max}$
- **Ahn–Horenstein** (Ahn & Horenstein (2013)): raportul valorilor proprii

$$\hat{k} = \arg \max_{1 \le k \le k_{\max}} \frac{\ell_k}{\ell_{k+1}}$$

  - cea mai mare scădere din graficul scree; în graficul de pe slide-ul „Componentele principale", la $k = 1$
- Zgomotul pur (portocaliu în graficul de pe slide-ul „Componentele principale") are toate valorile proprii aproape de 1: comparați cu el înainte de a numi un factor

Prețuri neactualizate (stale prices) și beta Dimson



- **Prețuri neactualizate (stale prices)**: o acțiune nelichidă nu se tranzacționează în fiecare zi, deci prețul ei reacționează cu întîrziere la știrile pieței
  - beta OLS din aceeași zi este deplasat spre 0
- **Beta Dimson** (Dimson (1979)): adăugăm randamentul pieței din ziua anterioară și din cea următoare și adunăm pantele

$$r_{i,t} = a + b_{-1} r_{m,t-1} + b_0 r_{m,t} + b_{+1} r_{m,t+1} + e_t$$

$$\beta^D = b_{-1} + b_0 + b_{+1}$$

- Graficul (simulare cu $\beta = 1$; acțiunea se tranzacționează în 75% din zile)
  - panta din aceeași zi 0,77; lagul adaugă 0,16; Dimson 0,93, aproape de 1
  - randamentele pieței din ziua anterioară și din cea următoare se formează pe calendarul complet, înainte de eliminarea vreunei zile

Harta exercițiilor

Exercițiul	Întrebarea	Tip	Model
A1	cîți dintre zece factori candidați sînt descoperiri?	Rezolvat	–
A2	de ce evaluează portofoliul tangent orice activ prin beta?	Rezolvat	–
B1	cît de precis se estimează beta și alfa pentru XLK?	Rezolvat	–
B2	este CAPM respins de nouă ETF-uri sectoriale?	Rezolvat	A2, B1
C1	ce erori conține un backtest de momentum scris de inteligența artificială (AI)?	Propus	B1, slide-urile introductive
După curs	A3 [Rezolvat], B3 [Rezolvat], B4 [Propus]: beta Vasicek, PCA și numărul de factori, valorile beta de la BVB	Rezolvat; B4: Propus	B1, A3
Pachete de extindere	A2 extinsă, A4–A8, B5–B11, C2, C3 (opționale, după seminar)	Rezolvat sau Propus (la fiecare cerință)	indicat la fiecare cerință

- **[Rezolvat]**: rezolvare completă în slide-uri și în notebook, un model de urmat; **[Propus]**: o rezolvați individual, rezolvarea se discută la seminar
- Partea A: derivări pe hîrtie; Partea B: estimare și inferență pe date reale, fiecare încheiată cu o întrebare de interpretare; Partea C: întrebări deschise și critica unui răspuns de inteligență artificială (AI)
- Numărul de laguri Newey–West indicat în fiecare exercițiu este convenția de replicare, nu alegerea optimă pentru orice eșantion

A1 [Rezolvat]: zece factori candidați

- **Întrebarea**: cîți dintre zece factori noi sînt descoperiri, dacă ținem cont că testăm zece deodată?
- **Date de intrare**: un exemplu nou (nu vectorul din Capitolul 3); nivel $\alpha = 5\%$, $M = 10$ teste bilaterale ale ipotezei de medie nulă
- **Cerințe**: completați fișa de lucru
  1. Numărați descoperirile naive $p \le 0{,}05$ și descoperirile Bonferroni $p \le 0{,}005$.
  2. Aplicați Holm: comparați $p_{(j)}$ cu $0{,}05/(M - j + 1)$ și opriți-vă la prima nerespingere.
  3. Aplicați Benjamini–Hochberg (BH): găsiți cel mai mare j cu $p_{(j)} \le 0{,}005j$ și respingeți $H_{(1)}, \ldots, H_{(j)}$.
  4. Precizați ce rată de eroare controlează fiecare procedură și sub ce dependență; ca extindere, aplicați și Benjamini–Yekutieli (BY).

j	1	2	3	4	5	6	7	8	9	10
$p_{(j)}$	0,001	0,004	0,009	0,012	0,021	0,035	0,048	0,060	0,20	0,55
Holm $0{,}05/(M-j+1)$	0,0050	…								
BH $0{,}005j$	0,005	…								

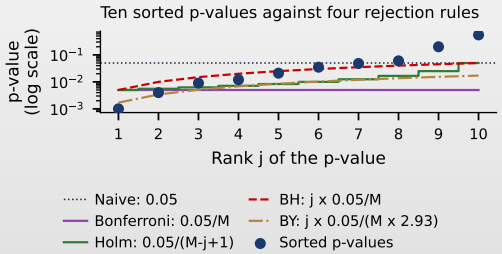
- **De raportat**: fișa completată și cele patru numere de descoperiri
- Notebook: secțiunea A1

Înapoi Harta exercițiilor

A1 [Rezolvat]: rezolvare

j	1	2	3	4	5	6	7	8	9	10
$p_{(j)}$	0,001	0,004	0,009	0,012	0,021	0,035	0,048	0,060	0,20	0,55
Holm	**0,0050**	**0,0056**	0,0063	0,0071	0,0083	0,0100	0,0125	0,0167	0,025	0,05
BH	**0,005**	**0,010**	**0,015**	**0,020**	**0,025**	0,030	0,035	0,040	0,045	0,050

- Naiv **7**; Bonferroni **2**; Holm: $0{,}009 > 0{,}0063$ este prima nerespingere, deci **2**
- BH: $j = 5$ este cel mai mare rang cu $p_{(j)} \leq 0{,}005j$ ($0{,}021 \leq 0{,}025$), deci **5**
  - pentru $j = 6, \ldots, 10$: $0{,}035 > 0{,}030$; $0{,}048 > 0{,}035$; $0{,}060 > 0{,}040$; $0{,}20 > 0{,}045$; $0{,}55 > 0{,}050$
- Rata de eroare controlată
  - Bonferroni și Holm: FWER, pentru orice dependență
  - BH: FDR, sub independență sau dependență pozitivă
  - BY (pragurile BH împărțite la $\sum_{i=1}^{10} 1/i = 2{,}93$): FDR, pentru orice dependență; **1** descoperire
- Interpretare: o căutare de factori tolerează o pondere cunoscută de factori falși, deci FDR este criteriul potrivit



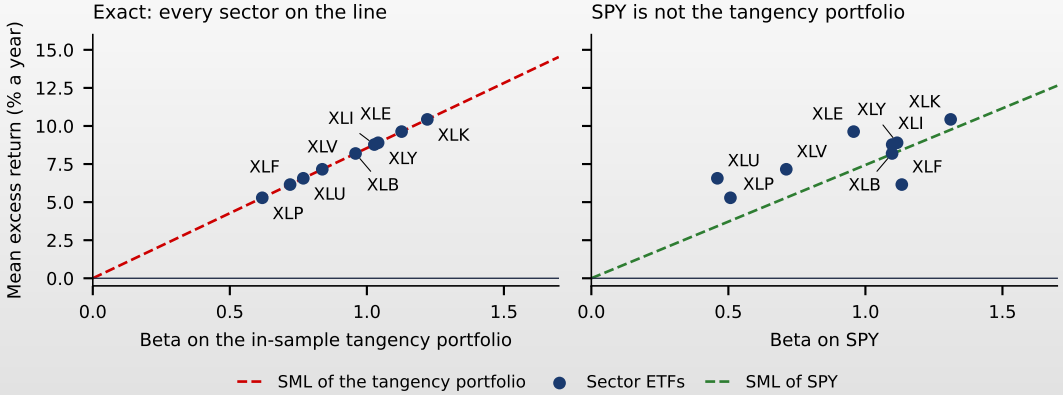
MFM_ch3_seminar

A2 [Rezolvat]: CAPM din portofoliul tangent

- **Întrebarea**: dacă piața este portofoliul tangent, de ce trebuie ca orice activ să satisfacă $\mu_i = \beta_i \mu_m$?
- **Date de intrare**: N active riscante cu randamentele în exces așteptate $\boldsymbol{\mu}$ ($N \times 1$) și covariança pozitiv definită $\boldsymbol{\Sigma}$; un activ fără risc
- **Ipoteza**: piața este portofoliul tangent $\mathrm{w}_m = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}/c$, cu scalarul $c = 1^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} \neq 0$; $R^e_m = \mathrm{w}_m^\top \mathrm{R}^e$
- **Cerințe**
  1. Exprimați vectorul de covariațe $\mathrm{Cov}(\mathrm{R}^e, R^e_m) = \boldsymbol{\Sigma}\mathrm{w}_m$ în funcție de $\boldsymbol{\mu}$ și c.
  2. Calculați $\mu_m = \mathrm{w}_m^\top \boldsymbol{\mu}$ și $\sigma^2_m = \mathrm{w}_m^\top \boldsymbol{\Sigma} \mathrm{w}_m$ și găsiți relația dintre ele.
  3. Combinați pașii 1 și 2 pentru a demonstra că $\mu_i = \beta_i \mu_m$ pentru orice activ, cu $\beta_i = \mathrm{Cov}(R^e_i, R^e_m)/\sigma^2_m$.
  4. Verificați rezultatul pe cele 9 ETF-uri sectoriale (lunar, 1999–2026): înlocuiți $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ cu momentele empirice și comparați $\hat{\mu}_i$ cu $\hat{\beta}_i \hat{\mu}_m$, întîi pentru portofoliul tangent din eșantion, apoi pentru SPY (ETF-ul pe indicele S&P 500).
- **De raportat**: derivarea în trei pași, un grafic al celor două verificări și o frază
- Notebook: secțiunea A2

Înapoi Harta exercițiilor

A2 [Rezolvat]: grafic



MFM_ch3_seminar

A2 [Rezolvat]: pașii demonstrației

1. $\boldsymbol{\Sigma} \mathrm{w}_m = \boldsymbol{\mu}/c$, deci $\mathrm{Cov}(\mathrm{R}^e, R^e_m) = \boldsymbol{\mu}/c$
2. $\sigma^2_m = \mathrm{w}_m^\top \boldsymbol{\Sigma} \mathrm{w}_m = \mu_m/c$, de unde $c = \mu_m/\sigma^2_m$
3. $\boldsymbol{\mu} = c\ \mathrm{Cov}(\mathrm{R}^e, R^e_m) = \boldsymbol{\beta}\,\mu_m$: SML
4. Portofoliul tangent din eșantion (medie 8,55% pe an): cea mai mare abatere $|\hat{\mu}_i - \hat{\beta}_i \hat{\mu}_m| \approx 3 \times 10^{-17}$% pe an (zero numeric); SPY: abateri de pînă la 3,1% pe an (XLU +3,1%, XLE +2,5%, XLF −2,3%)

⊡ Orice portofoliu eficient evaluează exact toate activele; un test CAPM verifică dacă aproximarea pieței este eficientă

B1 [Rezolvat]: beta și alfa pentru XLK

- **Întrebările**
  - care sînt beta și alfa CAPM ale ETF-ului de tehnologie?
  - cît de precis le cunoaștem?
- **Date**: XLK și Mkt-RF, randamente lunare în exces, ianuarie 1999 – iulie 2026 ($T = 331$), ca pe slide-ul de pregătire
- **Cerințe**
  1. Estimați $R^e_{\text{XLK},t} = \alpha + \beta\,\text{Mkt-RF}_t + \varepsilon_t$ prin OLS; anualizați α și erorile lui standard înmulțind cu 12.
  2. Calculați erorile standard (e.s.) clasice, White (robuste la heteroscedasticitate) și Newey–West pentru α și β, cu $L = 5$ laguri.
  3. Testați $H_0 : \alpha = 0$ la 5% cu statistica t Newey–West.
  4. Aplicați un bootstrap pe blocuri suprapuse (Künsch (1989)): reeșantionați blocuri de 12 luni consecutive din perechile (R^e_t, f_t) și reestimați α în 2 000 de extrageri, inițializare 42; raportați intervalul percentil de 95%.
  5. Reprezentați datele cu dreapta estimată, reziduurile în timp și distribuția bootstrap.
- **De raportat**: un tabel (estimație, trei erori standard), $t(\alpha)$, R^2, intervalul și graficul
- **Interpretare**: eroarea HAC a lui beta este dublul celei clasice: ce trăsătură a reziduurilor o cauzează?
- Notebook: secțiunea B1

Înapoi Harta exercițiilor

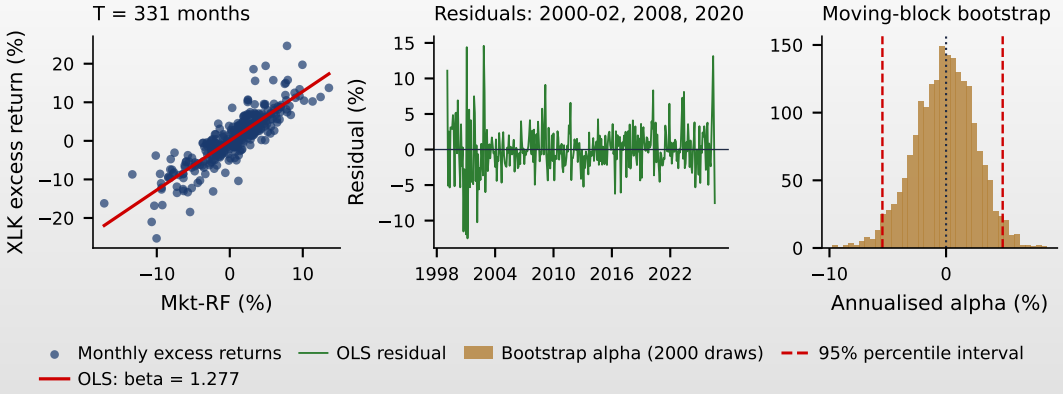
B1 [Rezolvat]: estimații și interpretare

	Estimație	e.s. clasică	e.s. White	e.s. Newey–West
α (anualizat)	0,33%	2,27%	2,32%	2,45%
β	1,277	0,041	0,054	0,082

1. $R^2 = 0{,}742$; $\hat{\alpha}$ lunar $= 0{,}027\%$; anualizăm valoarea nerotunjită: $12 \times 0{,}0273509\% = 0{,}328211\% \approx 0{,}33\%$ pe an
2. Laguri: $L = \lfloor 4 \times 3{,}31^{2/9} \rfloor = \lfloor 5{,}2 \rfloor = 5$; intervalul robust de 95% pentru β: $1{,}277 \pm 1{,}96 \times 0{,}082 = [1{,}12; 1{,}44]$ în loc de $[1{,}20; 1{,}36]$
3. $t(\alpha) = 0{,}33/2{,}45 = 0{,}13$, $p = 0{,}89$: nu respingem H_0; nu există dovezi ale unui alfa CAPM nenul
4. Bootstrap: $\alpha \in [-5{,}5\%; 4{,}9\%]$, abaterea standard a extragerilor 2,61%; intervalul conține 0

- Erorile clasice presupun $\mathrm{Var}(\varepsilon_t \mid f_t)$ constantă și absența autocorelației
  - estimatorul sandwich ponderează fiecare reziduu la pătrat cu $(f_t - \bar{f})^2$
- White mărește e.s. a lui β de la 0,041 la 0,054: heteroscedasticitate
  - reziduurile sînt mari cînd piața înregistrează variații mari (2000–2002, 2008, 2020)
- Newey–West o mărește pînă la 0,082: autocorelație
  - produsele $f_t \varepsilon_t$ sînt autocorelate, deci termenii încrucișați ai estimatorului HAC sînt pozitivi
- Estimațiile punctuale sînt aceleași; se schimbă doar incertitudinea

B1 [Rezolvat]: grafic



MFM_ch3_seminar

B2 [Rezolvat]: testul GRS pe ETF-urile sectoriale

- **Întrebarea**: este CAPM respins de cele nouă ETF-uri sectoriale, testate simultan?
- **Date**: 9 ETF-uri sectoriale, prețuri ajustate de sfîrșit de lună, aliniate pe lunile comune; randamente lunare în exces, ianuarie 1999 – iulie 2026 ($T = 331$, $N = 9$); factorul Mkt-RF
- **Cerințe**
  1. Estimați cele 9 regresii pe Mkt-RF; raportați valorile alfa anuale cu intervale Newey–West și valorile beta.
  2. Construiți matricea reziduurilor $\hat{E}$ ($T \times 9$), $\hat{\Sigma} = \hat{E}^\top \hat{E}/T$ și forma pătratică $\hat{\alpha}^\top \hat{\Sigma}^{-1} \hat{\alpha}$ cu valorile alfa lunare.
  3. Calculați $\hat{\mu}_m$, $\hat{\sigma}_m$ (ML), apoi statistica GRS W (slide-ul „Testul GRS") și valoarea p din $F(9, 321)$.
  4. Repetați testul pe cele 25 de portofolii mărime × B/M, în aceleași 331 de luni.
- **De raportat**: tabelul coeficienților, cele trei elemente ale lui W, W, p, graficul și comparația pe aceeași perioadă
- **Interpretare**: cele 25 de portofolii resping CAPM (GRS = 4,20 în 1963–2026); de ce nu îl resping sectoarele?
- Model: A2 (eficiență), B1 (regresie); Notebook: secțiunea B2

Înapoi Harta exercițiilor

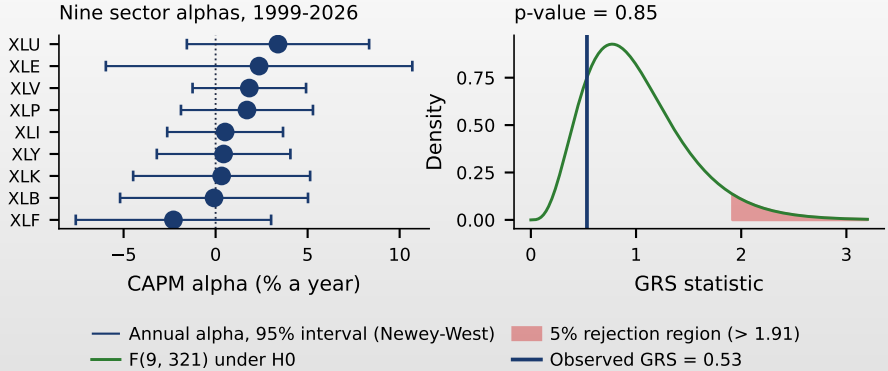
B2 [Rezolvat]: de la reziduuri la W

	XLB	XLE	XLF	XLI	XLK	XLP	XLU	XLV	XLY
$\hat{\alpha}$ (% pe an)	$-0{,}1$	2,4	$-2{,}3$	0,5	0,3	1,7	3,4	1,8	0,4
e.s. NW (% pe an)	2,6	4,2	2,7	1,6	2,5	1,8	2,5	1,6	1,9
$\hat{\beta}$	1,05	0,92	1,07	1,05	1,28	0,45	0,40	0,67	1,07

1. Niciun alfa nu este semnificativ separat; cel mai mare $|t|$ este al XLU, $3{,}39/2{,}53 = 1{,}34$
2. $\hat{\mathrm{E}}$: 331×9; $\hat{\mathbf{\Sigma}}$: 9×9; valorile alfa lunare (%): $-0{,}007; 0{,}197; -0{,}191; 0{,}043; 0{,}027; 0{,}142; 0{,}282; 0{,}153; 0{,}036$; $\hat{\boldsymbol{\alpha}}^\top \hat{\mathbf{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} = 0{,}0153$
3. $\hat{\mu}_m = 0{,}659\%$, $\hat{\sigma}_m = 4{,}51\%$, $SR_m = 0{,}146$; $W = \frac{321}{9} \times \frac{0{,}0153}{1+0{,}146^2} = 35{,}67 \times 0{,}0150 = 0{,}53$, $p = 0{,}85$
4. Aceleași 331 de luni, 25 de portofolii mărime $\times$ B/M: GRS $= 2{,}91$, $p = 9 \times 10^{-6}$

- Interpretare: nerespingerea se datorează activelor de test, nu eșantionului; sectoarele combină firme mici și mari, de valoare și de creștere
- Nerespingerea nu înseamnă acceptarea CAPM: alegeți active de test cu variață mare în caracteristicile suspectate că aduc o primă

B2 [Rezolvat]: grafic și interpretare



- $W = 0{,}53$ se află în corpul distribuției $F(9, 321)$, mult sub valoarea critică de 5%, 1,91: nu respingem H_0: toate $\alpha_i = 0$
- Forma cu raportul Sharpe (A4): sectoarele măresc raportul Sharpe lunar doar de la 0,146 la 0,191 (anual 0,51 → 0,66)

MFM_ch3_seminar

C1 [Propus]: critica unui răspuns de inteligență artificială (AI) — răspunsul

- **Promptul** trimis unui asistent AI: „P conține prețurile ajustate curățate de sfîrșit de lună ale celor 10 blue chips de la BVB, cîte o coloană pe acțiune. Construiește un portofoliu long–short de momentum pe 12 luni, top 3 minus ultimele 3, ponderi egale, și raportează media și volatilitatea anuale”
- **Răspunsul**: codul de mai jos și concluzia

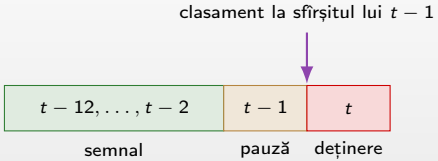
```
R   = np.log(P).diff()                          # monthly returns
mom = P.pct_change(12)                          # past 12-month return
ls  = {}
for t in R.index[13:]:
    s = mom.loc[t].dropna().sort_values()       # rank the stocks
    ls[t] = R.loc[t, s.index[-3:]].mean() - R.loc[t, s.index[:3]].mean()
ls = pd.Series(ls)
print(12 * ls.mean(), np.sqrt(12) * ls.std())  # annual mean, volatility
```

- **Concluzia AI**: „Momentum aduce circa 49% pe an, cu o volatilitate de 16%. Rezultatul concordă cu Fama & French (2012), care documentează momentum la Bursa de Valori București.”

Înapoi Harta exercițiilor

C1 [Propus]: critica unui răspuns de inteligență artificială (AI)

- **Strategia dorită de student** (specificația)
  - semnal momentum 12–1 (slide-ul „Momentum și deplasarea de anticipare"); o acțiune este eligibilă dacă are semnal și preț la sfîrșitul lunii $t-1$
  - long primele 3, short ultimele 3, ponderi egale, cu randamentele simple ale lunii t
- **Date**: 10 blue chips (TLV, SNP, BRD, TGN, FP, SNG, SNN, EL, DIGI, TEL; H2O, listat în iulie 2023, este exclus), prețuri de sfîrșit de lună, decembrie 2015 – august 2026
  - prețurile FP sînt ajustate pentru dividendul special din 7 septembrie 2023 (1,7225 lei pe acțiune)
- **Cerințe**
  1. Verificați momentul semnalului, agregarea randamentelor între acțiuni și citarea; enumerați toate erorile găsite.
  2. Pentru fiecare eroare, precizați cum ați detecta-o: un test, o cifră publicată sau lucrarea însăși.
  3. Raportați codul corectat, media și volatilitatea anuale, numărul de luni și statistica t a mediei (Newey–West, 3 laguri).
  4. Notați fiecare eroare în formatul `AI_ERRORS.md`: cerere, răspuns, eroare, cum ați găsit-o, corecția.
- **De raportat**: codul corectat, patru statistici și `AI_ERRORS.md`



- **Interpretare**
  - se confirmă „circa 49% pe an"?
  - la ce prag este semnificativă prima corectată: 1,96 (un singur test) sau 3 (pragul Harvey, Liu & Zhu (2016) pentru factori noi)?
- Codul AI rulează de obicei fără mesaj de eroare: erorile sînt de fond (financiare), nu de sintaxă
- Model: B1; Notebook: secțiunea C1

A3 [Rezolvat]: Vasicek ca medie a posteriori

- **Întrebarea**: de ce contractă regula Vasicek un beta zgomotos spre media transversală?
- **Date**: 9 ETF-uri sectoriale, randamente lunare în exces pe SPY, ianuarie 1999 – decembrie 2012; beta OLS cu erori Newey–West ($L = 4$ laguri, regula din slide-ul „Erorile standard clasice, White și Newey–West")
  - XLK: $\hat{\beta} = 1{,}415$, $\mathrm{se} = 0{,}139$; beta estimat ulterior, pe 2013–2026: 1,172
  - între cele 9 sectoare: $\bar{\beta} = 0{,}949$, $\mathrm{Var}_{\mathrm{cs}}(\hat{\beta}) = 0{,}1159$, media pătratelor erorilor standard $\overline{\mathrm{se}^2} = 0{,}0077$
- **Ipoteze**: a priori $\beta_i \sim N(\bar{\beta}, \sigma_\beta^2)$; verosimilitate $\hat{\beta}_i \mid \beta_i \sim N(\beta_i, \mathrm{se}_i^2)$, cu se_i cunoscut
- **Cerințe**
  1. Înmulțiți distribuția a priori cu verosimilitatea, completați pătratul în β și scrieți distribuția a posteriori: media β_i^V și varianța ei.
  2. Folosiți legea variației totale pentru a arăta că $\mathrm{Var}_{\mathrm{cs}}(\hat{\beta}) = \sigma_\beta^2 + \mathbb{E}[\mathrm{se}^2]$ și scrieți estimatorul Bayes empiric al lui σ_β^2.
  3. Calculați σ_β^2, ponderea w, β^V și abaterea standard a posteriori pentru XLK, precum și beta ajustat cu regula practicienilor $0{,}33 + 0{,}67\hat{\beta}$ (beta Blume, slide-ul „Contracția bayesiană a valorilor beta").
  4. Comparați erorile absolute ale beta brut, Vasicek și Blume față de estimația ulterioară 1,172.
- **De raportat**: distribuția a posteriori, cifrele de la pasul 3 și cele trei erori
- Notebook: secțiunea A3

Înapoi Harta exercițiilor

A3 [Rezolvat]: distribuția a posteriori

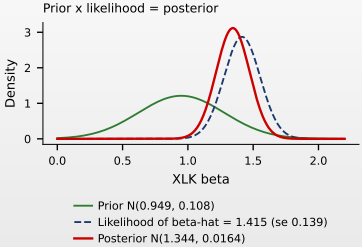
1. A posteriori $\propto \exp\{-\frac{(\beta-\bar{\beta})^2}{2\sigma_{\bar{\beta}}^2}\}\exp\{-\frac{(\hat{\beta}-\beta)^2}{2\,\mathrm{se}^2}\}$; exponentul este

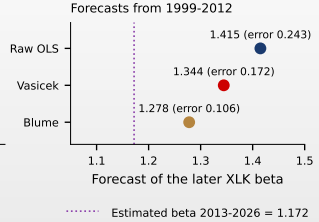
$$-\tfrac{1}{2}\Big[\beta^2\Big(\tfrac{1}{\sigma_{\bar{\beta}}^2}+\tfrac{1}{\mathrm{se}^2}\Big)-2\beta\Big(\tfrac{\bar{\beta}}{\sigma_{\bar{\beta}}^2}+\tfrac{\hat{\beta}}{\mathrm{se}^2}\Big)\Big]+\text{const}$$

   - Completînd pătratul: o distribuție a posteriori Normală cu varianța $v=(1/\sigma_{\bar{\beta}}^2+1/\mathrm{se}^2)^{-1}$
   - și media $v(\bar{\beta}/\sigma_{\bar{\beta}}^2+\hat{\beta}/\mathrm{se}^2)=w\hat{\beta}+(1-w)\bar{\beta}$, cu $w=\sigma_{\bar{\beta}}^2/(\sigma_{\bar{\beta}}^2+\mathrm{se}^2)$

2. Legea varianței totale: $\mathrm{Var}(\hat{\beta})=\mathrm{Var}(\mathbb{E}[\hat{\beta}\mid\beta])+\mathbb{E}[\mathrm{Var}(\hat{\beta}\mid\beta)]=\sigma_{\bar{\beta}}^2+\mathbb{E}[\mathrm{se}^2]$
   - Bayes empiric: $\hat{\sigma}_{\bar{\beta}}^2=\max\{\mathrm{Var}_{cs}(\hat{\beta})-\overline{\mathrm{se}^2},\,0\}$
   - Folosirea lui $\mathrm{Var}_{cs}(\hat{\beta})$ ca atare tratează zgomotul de eșantionare drept varianță reală și dă o contracție prea mică

A3 [Rezolvat]: rezolvare





3. $\hat{\sigma}^2_\beta = 0{,}1159 - 0{,}0077 = 0{,}1082$; $\text{se}^2 = 0{,}0193$; $w = 0{,}849$; $\beta^V = 0{,}849 \times 1{,}415 + 0{,}151 \times 0{,}949 = 1{,}344$; varianța a posteriori 0,0164 (abatere standard 0,128); Blume 1,278
4. Erori față de 1,172: brut 0,243, Vasicek 0,172, Blume 0,106

- Beta XLK este precis, deci Vasicek îl contractă puțin; Blume apropie orice beta de 1 cu o treime din distanță
- Beta ulterior este și el o estimație; A5 (extindere) arată de ce regresia spre medie depășește zgomotul de eșantionare

MFM_ch3_seminar

B3 [Rezolvat]: PCA și numărul de factori

- **Întrebările**
  - cîți factori comuni determină randamentele zilnice ale sectoarelor din SUA (Statele Unite ale Americii)?
  - care sînt acești factori?
- **Date**: 11 ETF-uri sectoriale, randamente logaritmice zilnice în zilele comune, 20 iunie 2018 – 18 septembrie 2026 ($T = 2\,073$); SPY păstrat separat pentru comparație
- **Cerințe**
  1. Standardizați fiecare coloană: Z este $T \times 11$; calculați $\mathrm{C} = \mathrm{Z}^\top \mathrm{Z}/T$, valorile proprii $\ell_1 \geq \cdots \geq \ell_{11}$ și vectorii proprii de normă 1, cu semnul ales astfel încît suma încărcărilor să fie pozitivă.
  2. Raportați ponderile $\ell_k/11$ pentru PC1–PC3 și încărcările PC1 și PC2; construiți scorurile $\mathrm{Z}\gamma_1$ (γ_1: vectorul propriu al PC1) și corelați-le cu SPY.
  3. Estimați numărul de factori, cu $k_{\max} = 5$, prin Kaiser, Bai–Ng (IC_{p2}) și Ahn–Horenstein (raportul valorilor proprii), definite pe slide-ul „Numărul de factori".
- **De raportat**: ponderile, încărcările, corelația cu SPY, cele trei estimații ale lui k și un grafic
- **Interpretare**: PC2 opune utilitățile și bunurile de larg consum tehnologiei
  - are PC2 o primă de risc?
  - ce ar fi necesar pentru a testa o primă asociată PC2?
- Notebook: secțiunea B3

◂ Înapoi ▸ Harta exercițiilor

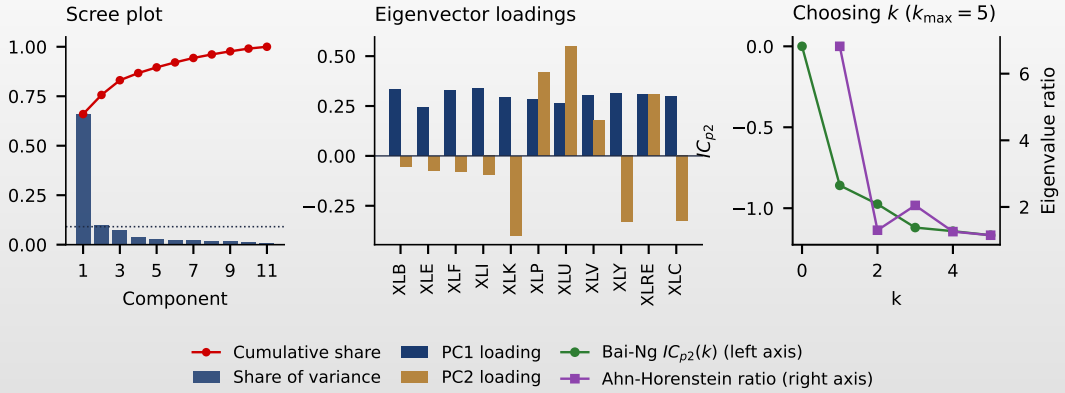
B3 [Rezolvat]: componentele

1. Valorile proprii 7,26; 1,06; 0,82; 0,40; 0,32; ... (suma lor este 11)
2. Ponderi: PC1 66,0%, PC2 9,7%, a treia componentă (PC3) 7,4%; primele trei explică 83,1%
   - Încărcările PC1 sînt toate pozitive (0,24–0,34): un portofoliu aproape egal ponderat; corelația scorurilor PC1 cu SPY $= 0{,}954$: PC1 este puternic legată de piață, dar nu identică cu SPY
   - PC2: XLU $+0{,}55$, XLP $+0{,}42$, XLRE $+0{,}31$ față de XLK $-0{,}40$, XLY $-0{,}33$: sectoare defensive față de sectoare de creștere
   - Vectorii proprii sînt ponderi de portofoliu pe randamente standardizate, nu valori beta; un beta pe PC1 necesită regresia fiecărui sector pe scoruri
3. Numărul de factori: Kaiser 2, Bai–Ng 5, Ahn–Horenstein 1 (calculul: slide-ul „B3: numărul de factori")
   - estimatorii nu concordă: cu $N = 11$ rezultatele asimptotice pentru N mare nu se aplică; raportați intervalul, nu un singur k

⊡ Interpretare: PCA explică covarianța, nu randamentele medii, deci PC2 nu are neapărat o primă de risc

- test: Fama–MacBeth cu valorile beta pe PC2 pe multe active, sau un portofoliu long–short sortat după încărcarea pe PC2, cu media testată cu erori Newey–West

B3 [Rezolvat]: grafic



MFM_ch3_seminar

B3 [Rezolvat]: numărul de factori

- Kaiser: $\hat{k} = 2$
  - două valori proprii depășesc 1: 7,26 și 1,06
- Bai–Ng: $\hat{k} = 5 = k_{\max}$
  - penalizarea este $\frac{N+T}{NT} \ln 11 = 0{,}219$ pe factor
  - $IC_{p2}(k)$ pentru $k = 0, \ldots, 5$: $0; -0{,}86; -0{,}97; -1{,}12; -1{,}14; -1{,}17$; scade pînă la $k_{\max}$
- Ahn–Horenstein: $\hat{k} = 1$
  - rapoartele ℓ_k / ℓ_{k+1} pentru $k = 1, \ldots, 5$: $6{,}82; 1{,}30; 2{,}05; 1{,}26; 1{,}15$; maximul este la $k = 1$
- De raportat: un factor dominant, legat de piață; Kaiser reține o a doua componentă, dar nu stabilește dacă este semnificativă; Bai–Ng selectează limita impusă, $k_{\max}$: rezultat neconcludent fără o analiză a sensibilității la $k_{\max}$

B4 [Propus]: beta CAPM pentru blue chips BVB

- **Întrebările**
  - cît de sensibile sînt blue chips de la BVB la indicele BET?
  - cum deplasează prețurile neactualizate beta zilnic?
- **Date**: 11 blue chips (prețuri de închidere ajustate) și indicele BET, randamente logaritmice zilnice de la 6 ianuarie 2015 (sau de la listare) la 18 septembrie 2026
  - fiecare acțiune se aliniază separat cu BET, pe zilele comune
  - rata fără risc zilnică este tratată ca o constantă, deci nu modifică beta

Exemplu de verificare: FP în jurul datei de 7 septembrie 2023

Data	Închidere	r din închidere	r ajustat	BET_t
5 sept.	1,918	−0,52%	−0,52%	−1,05%
6 sept.	1,984	3,38%	3,38%	1,96%
7 sept.	0,381	−165,0%	5,85%	1,36%
8 sept.	0,386	1,30%	1,30%	1,02%
11 sept.	0,391	1,16%	1,16%	1,34%

- 7 septembrie 2023 este data ex a dividendului special de 1,7225 lei pe acțiune
- Prețul de închidere scade cu 81%, dar acționarul primește dividendul: randamentul total al zilei este +6,0%
- Un eveniment de capital neajustat creează un randament fals: folosiți prețurile ajustate și verificați orice zi cu $|r_t| > 0{,}5$

Înapoi Harta exercițiilor

B4 [Propus]: beta BVB — cerințe și interpretare

- **Cerințe**
  1. Construiți randamentul BET din ziua următoare (lead) și din ziua anterioară (lag) pe calendarul complet al fiecărei perechi, apoi eliminați zilele cu $|r_t| > 0{,}5$ pentru acțiune sau BET; raportați zilele eliminate.
  2. Estimați fiecare beta prin OLS pe BET, cu erori Newey–West ($L = \lfloor 4(T/100)^{2/9} \rfloor$ laguri) și R^2.
  3. Calculați valorile beta Vasicek Bayes empiric (A3), cu cele 11 acțiuni ca secțiune transversală.
  4. Estimați valorile beta Dimson, cu BET la $t-1$, t și $t+1$ (slide-ul „Prețuri neactualizate (stale prices) și beta Dimson").
- **De raportat**: tabel: OLS, Vasicek, Dimson; grafic: OLS cu intervale Newey–West de 95%
- **Interpretare**
  - la ce acțiuni schimbă cel mai mult corecția Dimson valoarea beta?
  - de ce tocmai la aceste acțiuni?
- Model: B1, A3; Notebook: secțiunea B4

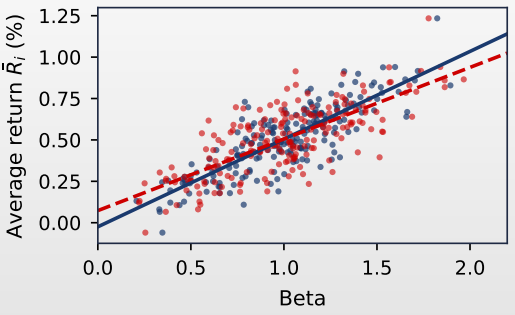
Regresiile în două etape (Fama–MacBeth)

- **Etapa 1** (serii de timp): pentru fiecare activ estimăm valorile beta $\hat{\beta}_i$ ($K \times 1$) prin OLS; $\hat{\mathrm{B}}$: $N \times K$
- **Etapa 2** (Fama & MacBeth (1973); transversal, în fiecare lună t): regresăm cele N randamente pe valorile beta

$$R^e_{i,t} = \lambda_{\mathbf{0},t} + \hat{\boldsymbol{\beta}}_i^\top \boldsymbol{\lambda}_t + e_{i,t}, \qquad \hat{\boldsymbol{\lambda}} = \frac{1}{T}\sum_{t=\mathbf{1}}^{T} \hat{\boldsymbol{\lambda}}_t$$

  - $\boldsymbol{\lambda}$: primele de risc (prețurile riscului) ale factorilor
  - $\lambda_{\mathbf{0}}$: termenul liber; este zero dacă investitorii se pot împrumuta la rata fără risc, iar în modelul Black este excesul ratei cu beta zero față de rata fără risc
  - pentru un factor tranzacționat, modelul implică și $\lambda = \mathbb{E}[f]$
- Erorile standard
  - Fama–MacBeth: $\mathrm{se}(\hat{\lambda}) = \mathrm{sd}(\hat{\lambda}_t)/\sqrt{T}$, abaterea standard a pantelor lunare împărțită la $\sqrt{T}$
  - dacă seria $\hat{\lambda}_t$ este autocorelată, se folosește Newey–West pe această serie
  - corecția Shanken (Shanken (1992)) ține cont că valorile beta sînt estimate: $\mathrm{Var}_{\mathbf{sh}} = (1+c)(\mathrm{Var}_{FM} - \boldsymbol{\Sigma}_f/T) + \boldsymbol{\Sigma}_f/T$, cu $c = \hat{\boldsymbol{\lambda}}^\top \boldsymbol{\Sigma}_f^{-\mathbf{1}} \hat{\boldsymbol{\lambda}}$ și $\boldsymbol{\Sigma}_f$ covarianța factorilor
  - Kan–Robotti–Shanken (Kan, Robotti & Shanken (2013)): erori calculate prin GMM (*generalized method of moments*, metoda generalizată a momentelor), valide și cînd modelul este greșit specificat
  - robust la specificare greșită nu înseamnă HAC: cu momente autocorelate, se adaugă laguri Newey–West în covarianța momentelor (`two_pass_gmm(..., lags=6)`)

Erori în variabile și factori inutili



- true β_i: slope 0.53
- estimated $\hat{\beta}_i$: slope 0.43

- **Erori în variabile**: etapa 2 folosește $\hat{\beta}_i = \beta_i + \eta_i$, cu zgomotul de estimare η_i
  - panta pe un regresor zgomotos este atenuată spre 0 cu factorul $\text{Var}_{cs}(\beta)/[\text{Var}_{cs}(\beta) + \text{Var}(\eta)]$
  - graficul: 200 de active simulate, $T = 60$; panta din simulare 0,53, panta pe estimările $\hat{\beta}_i$ 0,43
  - remedii: portofolii ($\text{Var}(\eta)$ mai mic), eșantioane lungi, corecția Shanken pentru eroarea standard
- **Factor inutil** (Kan & Zhang (1999)): un factor fără legătură cu randamentele are toți parametrii β_i egali cu 0
  - valorile lui beta estimate sînt zgomot pur, totuși etapa 2 poate raporta o primă mare, „semnificativă" (A7)
  - verificați întîi că valorile beta diferă simultan de 0 (etapa 1)

Ajustarea transversală, OLS și GLS

- Ajustarea randamentelor medii $\boldsymbol{\mu}$ ($N \times 1$) prin $\mathrm{X} = [1, \hat{\mathrm{B}}]$

$$R^2_{OLS} = 1 - \frac{\mathrm{e}^\top \mathrm{e}}{\sum_i (\mu_i - \bar{\mu})^2}, \qquad \mathrm{e} = \boldsymbol{\mu} - \mathrm{X}\hat{\boldsymbol{\lambda}}^+$$

  - $\hat{\boldsymbol{\lambda}}^+ = (\mathrm{X}^\top \mathrm{X})^{-1} \mathrm{X}^\top \boldsymbol{\mu}$: termenul liber $\hat{\lambda}_0$ și primele $\hat{\boldsymbol{\lambda}}$; $\bar{\mu}$: media valorilor μ_i
  - cu puține active de test, puternic determinate de factori (25 mărime $\times$ B/M), aproape orice model cu factorii potriviți atinge un R^2_{OLS} mare (Lewellen, Nagel & Shanken (2010))

- **GLS**: ponderăm erorile de evaluare cu inversa covarianței $\hat{\mathrm{V}}^{-1}$ a randamentelor

$$R^2_{GLS} = 1 - \frac{\mathrm{e}_G^\top \hat{\mathrm{V}}^{-1} \mathrm{e}_G}{(\boldsymbol{\mu} - \bar{\mu}_G 1)^\top \hat{\mathrm{V}}^{-1} (\boldsymbol{\mu} - \bar{\mu}_G 1)}$$

  - măsoară ajustarea portofoliilor relevante pentru raportul Sharpe maxim, nu doar a celor 25 de medii
  - $\mathrm{e}_G = \boldsymbol{\mu} - \mathrm{X}\hat{\boldsymbol{\lambda}}^+_{GLS}$, $\hat{\boldsymbol{\lambda}}^+_{GLS} = (\mathrm{X}^\top \hat{\mathrm{V}}^{-1} \mathrm{X})^{-1} \mathrm{X}^\top \hat{\mathrm{V}}^{-1} \boldsymbol{\mu}$; $\bar{\mu}_G = 1^\top \hat{\mathrm{V}}^{-1} \boldsymbol{\mu} / 1^\top \hat{\mathrm{V}}^{-1} 1$: media GLS
  - remediul din Lewellen, Nagel & Shanken (2010): adăugați active de test cu alte caracteristici (industrii) și raportați ambele R^2

Compararea modelelor factoriale prin rapoarte Sharpe

- Pătratul raportului Sharpe maxim atins cu cei K factori ai unui model (Barillas & Shanken (2018))

$$SR^2 = \boldsymbol{\mu}_f^\top \boldsymbol{\Sigma}_f^{-1} \boldsymbol{\mu}_f$$

  - $\boldsymbol{\mu}_f$, $\boldsymbol{\Sigma}_f$: mediile și covarianța factorilor; modelul mai bun evaluează mai multe active și are SR^2 mai mare
- Modele **incluse** unul în altul: o **regresie de spanning**

$$f_{\mathbf{nou},t} = a + \mathrm{b}^\top \mathrm{f}_t + u_t$$

  - $a \neq 0$: noul factor ridică raportul Sharpe maxim; $a = 0$: este acoperit (redundant); testăm a cu Newey–West
- Modele **neincluse** unul în altul (Barillas, Kan, Robotti & Shanken (2020)): bootstrap pentru diferența $SR_A^2 - SR_B^2$
  - folosiți aceleași luni reeșantionate pentru ambele modele, astfel încît diferența să păstreze corelația lor

Valori beta pe ferestre mobile și valori beta istorice

- **Beta pe fereastră mobilă** (*rolling*): estimăm $\beta_{i,t}$ doar pe ultimele H luni pînă la $t-1$ (C2: $H = 36$, cel puțin 18)
  - un portofoliu sortat după beta, format la $t-1$, poate fi astfel implementat, fără look-ahead
- **Anomalia „beta mic"** (Frazzini & Pedersen (2014)): acțiunile cu beta mic au randamente mai mari decît prevede SML, cele cu beta mare mai mici
  - BAB (*betting against beta*, „pariu împotriva lui beta"): mărim expunerea părții cu beta mic și reducem expunerea părții cu beta mare, astfel încît beta portofoliului să fie 0
  - un spread nescalat beta mic minus beta mare are beta negativ: o parte din randamentul lui este expunere la piață
- **Regresia spre medie**: valorile beta estimate într-o perioadă se apropie de 1 în perioada următoare
  - regresia valorilor beta ulterioare pe cele anterioare, între active: o pantă sub 1 măsoară apropierea de medie (A5)

Instrumentele proiectului de replicare (pachetul 8)

- **Teme de factori**: gruparea ierarhică a factorilor (metoda Ward, Ward (1963))
  - distanța $d_{jk} = 1 - \text{corr}(\hat{\varepsilon}_j, \hat{\varepsilon}_k)$ între reziduurile CAPM; Ward unește cele două grupe a căror reuniune mărește cel mai puțin varianța din interiorul grupelor; Ward cere distanțe euclidiene, pentru serii standardizate $\sqrt{2(1-\rho_{jk})}$ (`scipy` cu $1-\rho$ direct nu este echivalent)
  - dendrograma arată unirile; tăiată la o anumită înălțime, dă temele
- **Model bayesian ierarhic** pentru multe valori alfa (Jensen, Kelly & Pedersen (2023))

$$\hat{\boldsymbol{\alpha}} \sim N\left(0,\ \mathbf{M}\mathbf{M}^{\top}\tau_c^2 + \mathbf{I}\tau_w^2 + \hat{\boldsymbol{\Sigma}}/T\right)$$

  - M: matricea 0/1 care arată în ce temă intră fiecare factor
  - τ_c: abaterea standard a valorilor alfa între teme; τ_w: abaterea standard a valorilor alfa ale factorilor din aceeași temă
  - $\hat{\boldsymbol{\Sigma}}/T$: zgomotul de eșantionare al valorilor alfa estimate
  - media a posteriori a fiecărui alfa împrumută informație de la tema sa, așa cum Vasicek împrumută de la secțiunea transversală (slide-ul „Contracția bayesiană a valorilor beta")
- **Scăderea după publicare**: regresăm valorile alfa din afara eșantionului pe cele din eșantion
  - o pantă sub 1: o parte din alfa publicat era zgomot sau a fost eliminată prin arbitraj

Pachete de extindere: cunoștințe necesare

Pachet	Probleme	Cunoștințe necesare
1. Restricții de evaluare	A2 extinsă [Rezolvat], A4, A5	A2, A3, B2
2. Inferența în două etape	A6 [Rezolvat], A7, B5, B6, B7	B1, B2, cursul, Secțiunea 6
3. Compararea modelelor	B8	B1, cursul, Secțiunea 7
4. Testare multiplă pe alfa reale	B9	A1, B1
5. Factori latenți	A8, B10	B3
6. ETF-uri factoriale zilnic	B11	B1
7. Prime factoriale la BVB	C2	B4, C1
8. Proiect de replicare	C3	A3, A1, studiul de caz din curs

- Fiecare pachet are secțiunea lui în notebook și cod de pornire; rezolvați problemele în ordine în cadrul unui pachet
- Utile pentru proiectul de echipă: C2 și C3 sînt posibile puncte de plecare

A2 extinsă [Rezolvat]: alte portofolii eficiente

- **Întrebarea**: are ecuația de evaluare prin beta nevoie de portofoliul pieței sau doar de un reper eficient?
- **Date de intrare**: ca în A2; două repere
  - un portofoliu de pe CML, $a\mathrm{w}_m$ în active riscante ($a > 0$) și $1 - a$ în activul fără risc
  - un portofoliu de pe frontiera activelor riscante (randamente brute R, medii m): condiția de ordinul întîi pentru $\min \mathrm{w}^\top \Sigma \mathrm{w}$ cu restricțiile $\mathrm{w}^\top \mathrm{m} = m^*$, $\mathrm{w}^\top 1 = 1$ dă $\Sigma \mathrm{w} = \kappa_1 \mathrm{m} + \kappa_2 1$, cu κ_1, κ_2 multiplicatorii Lagrange ai celor două restricții
- **Cerințe**
  1. Arătați că portofoliul de pe CML are $\beta_{i,p} = \beta_i / a$ și $\mu_p = a\mu_m$, deci din nou $\mu_i = \beta_{i,p}\mu_p$.
  2. Presupuneți $\kappa_1 \neq 0$ (portofoliul nu este cel de variață minimă globală). Din condiția de ordinul întîi, arătați că $m_i - \gamma = \beta_{i,p}(m_p - \gamma)$ cu $\gamma = -\kappa_2/\kappa_1$, rata cu beta zero.
  3. Precizați ce implică aceasta pentru un test CAPM cu o aproximare a pieței.
- **De raportat**: cele două demonstrații și o frază despre ce verifică un test CAPM
- Model: A2

Înapoi Harta exercițiilor

A2 extinsă [Rezolvat]: rezolvare

1. Randamentul în exces al portofoliului de pe CML este aR_m^e: $\text{Cov}(R_i^e, aR_m^e) = a\,\text{Cov}(R_i^e, R_m^e)$ și $\text{Var}(aR_m^e) = a^2\sigma_m^2$, deci $\beta_{i,p} = \beta_i/a$; cu $\mu_p = a\mu_m$, $\beta_{i,p}\mu_p = \beta_i\mu_m = \mu_i$
2. $\text{Cov}(\mathrm{R}, R_p) = \boldsymbol{\Sigma}\mathrm{w} = \kappa_1\mathrm{m} + \kappa_2\mathrm{1}$ și $\sigma_p^2 = \mathrm{w}^\top\boldsymbol{\Sigma}\mathrm{w} = \kappa_1 m_p + \kappa_2$
   - Deci $\kappa_1(\mathrm{m} - \gamma\mathrm{1}) = \text{Cov}(\mathrm{R}, R_p)$ și $\kappa_1(m_p - \gamma) = \sigma_p^2$; prin împărțire: $\mathrm{m} - \gamma\mathrm{1} = \boldsymbol{\beta}_p(m_p - \gamma)$
   - γ este media oricărui portofoliu cu beta zero față de p: CAPM cu beta zero al lui Black (1972)
3. Orice reper eficient medie–varianță evaluează exact toate activele prin beta; o respingere arată că aproximarea nu este eficientă pentru activele de test, o afirmație despre aproximare în aceeași măsură ca despre model (critica lui Roll, Roll (1977))

A4 [Propus]: GRS măsoară un cîștig de raport Sharpe

- **Întrebarea**: de ce măsoară GRS cît de mult îmbunătățesc activele de test raportul Sharpe al pieței?
- **Notații**: $Y_t = (f_t, \mathrm{R}_t^{e\top})^\top$, $(N+1)\times 1$; momentele ML $\boldsymbol{\mu}_Y = (\mu_f, \boldsymbol{\mu}_R^\top)^\top$ și $\mathrm{V} = \begin{pmatrix} \sigma_f^2 & \boldsymbol{\sigma}_{Rf}^\top \\ \boldsymbol{\sigma}_{Rf} & \boldsymbol{\Sigma}_R \end{pmatrix}$, toate împărțite la T; $\boldsymbol{\beta} = \boldsymbol{\sigma}_{Rf}/\sigma_f^2$, $\boldsymbol{\alpha} = \boldsymbol{\mu}_R - \boldsymbol{\beta}\mu_f$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_R - \boldsymbol{\beta}\boldsymbol{\beta}^\top \sigma_f^2$
- **Identitate datã** (inversa pe blocuri, cu complementul Schur $\boldsymbol{\Sigma}$ al lui σ_f^2)

$$\mathrm{V}^{-1} = \begin{pmatrix} 1/\sigma_f^2 + \boldsymbol{\beta}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} & -\boldsymbol{\beta}^\top \boldsymbol{\Sigma}^{-1} \\ -\boldsymbol{\Sigma}^{-1}\boldsymbol{\beta} & \boldsymbol{\Sigma}^{-1} \end{pmatrix}$$

- **Cerințe**
  1. Dezvoltați $SR^2(f, \mathrm{R}) = \boldsymbol{\mu}_Y^\top \mathrm{V}^{-1}\boldsymbol{\mu}_Y$ în patru termeni și arătați că este egal cu $\mu_f^2/\sigma_f^2 + \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}$.
  2. Scrieți statistica GRS W în funcție de $SR^2(f)$ și $SR^2(f, \mathrm{R})$.
  3. Cele 25 de portofolii mărime $\times$ B/M: $T = 757$, $N = 25$, $SR(f) = 0{,}134$ și $SR(f, \mathrm{R}) = 0{,}405$ lunar; calculați W și decideți la 5% (valoarea critică a lui $F(25, 731)$: 1,52).
  4. Comparați cîștigul observat cu cel așteptat sub H_0, $N(1 + SR^2(f))/(T - N - 3)$.
- **De raportat**: calculul algebric, W, decizia și o frază
- Model: A2, B2; Notebook: secțiunea A4

Înapoi Harta exercițiilor

A5 [Propus]: panta Blume ca pondere de contracție

- **Întrebarea**: este regresia valorilor beta ulterioare pe cele anterioare un estimator Vasicek?
- **Ipoteze simplificatoare**: parametrii β_i sînt constanți în timp, $\hat{\beta}_{i,k} = \beta_i + \eta_{i,k}$ în perioadele $k = 1, 2$; zgomotele sînt independente de β_i și între ele, cu o varianță **comună** s^2; β_i are media $\bar{\beta}$ și varianța σ^2_β
- **Cerințe**
  1. Derivați, pentru $N \to \infty$, limitele pantei OLS b și ale termenului liber a din regresia lui $\hat{\beta}_2$ pe $\hat{\beta}_1$.
  2. Arătați că $a + b\hat{\beta}_1$ este atunci beta Vasicek cu $w = b$.
  3. Cu $\sigma^2_\beta = 0{,}10822$, $s^2 = 0{,}00773$, $\bar{\beta} = 0{,}9495$ din A3, calculați b și a implicite; comparați-le cu 0,7242 și 0,2304 estimate și reprezentați ambele drepte.
  4. Panta estimată are eroarea standard OLS clasică 0,1137 ($N = 9$): calculați $(0{,}7242 - b)/0{,}1137$ ca ilustrare și enumerați explicațiile care rămîn posibile.
- **De raportat**: cele două limite, b și a implicite, graficul celor două drepte și lista explicațiilor
- Model: A3; Notebook: secțiunea A5

	$\hat{\beta}_1$ 1999–2012	e.s.$_1$	$\hat{\beta}_2$ 2013–2026
XLB	1,174	0,083	1,026
XLE	0,904	0,089	1,051
XLF	1,217	0,120	1,029
XLI	1,122	0,057	1,078
XLK	1,415	0,139	1,172
XLP	0,451	0,064	0,578
XLU	0,463	0,086	0,455
XLV	0,716	0,047	0,705
XLY	1,083	0,063	1,169

Înapoi Harta exercițiilor

A6 [Rezolvat]: erori în variabile în etapa 2

- **Întrebarea**: cît de mult atenuează etapa 2 o primă pentru că valorile beta sînt estimate?
- **Cadrul**: un factor, $R_{it} = a_i + \beta_i f_t + \varepsilon_{it}$; ε_{it} i.i.d. în t și i, cu varianța σ_ε^2, independente de f și de valorile β_i; β_i i.i.d. între active, cu varianța $\mathrm{Var}_{cs}(\beta)$
- **Trei prime**: λ, prețul adevărat al riscului; $\mu_f = \mathbb{E}[f_t]$; $\lambda_T = \lambda + \bar{f} - \mu_f$, prima realizată în eșantion. Restricția de evaluare: $a_i = \lambda_0 + \beta_i(\lambda - \mu_f)$
- Lucrăm condiționat de traiectoria factorului $f_1, \ldots, f_T$; etapa 2: regresia OLS a lui $\bar{R}_i$ pe $(1, \hat{\beta}_i)$, $\hat{\beta}_i = \beta_i + \eta_i$, $\eta_i = \sum_t (f_t - \bar{f})\varepsilon_{it} / \sum_t (f_t - \bar{f})^2$
- **Cerințe**
  1. Arătați că $\mathrm{Var}(\eta_i \mid f) = \sigma_\varepsilon^2/(T\hat{\sigma}_f^2)$ și că $\mathrm{Cov}(\eta_i, \bar{\varepsilon}_i \mid f) = 0$.
  2. Arătați că, pentru $N \to \infty$ și T fix, panta din etapa 2 converge la $\lambda_T \ \mathrm{Var}_{cs}(\beta)/[\mathrm{Var}_{cs}(\beta) + \sigma_\varepsilon^2/(T\hat{\sigma}_f^2)]$.
  3. Calculați factorul de atenuare pentru cele două paneluri de mai jos, cu $\mathrm{Var}_{cs}(\beta) = \mathrm{Var}_{cs}(\hat{\beta}) - \bar{\sigma}_\varepsilon^2/(T\hat{\sigma}_f^2)$, și explicați ce anume ajustează corecția Shanken.
- **De raportat**: cele două demonstrații, cei doi factori de atenuare și o frază despre corecția Shanken; Notebook: secțiunea A6

Panel (CAPM)	T	$\mathrm{Var}_{cs}(\hat{\beta})$	$\bar{\sigma}_\varepsilon^2$ (%² pe lună)	$\hat{\sigma}_f^2$ (%² pe lună)
25 mărime × B/M, 1963–2026	757	0,01905	9,33	19,87
9 ETF-uri sectoriale, 1999–2026	331	0,09247	13,45	20,36

Înapoi Harta exercițiilor

A6 [Rezolvat]: derivarea

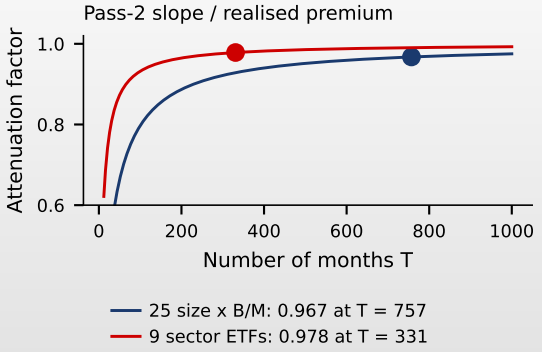
1. Condiționat de f, η_i este o combinație liniară a erorilor ε_{it} cu ponderile $(f_t - \bar{f}) / \sum_s (f_s - \bar{f})^2$
   - $\text{Var}(\eta_i \mid f) = \sigma_\varepsilon^2 \sum_t (f_t - \bar{f})^2 / [\sum_s (f_s - \bar{f})^2]^2 = \sigma_\varepsilon^2 / (T \hat{\sigma}_f^2)$
   - $\text{Cov}(\bar{\varepsilon}_i, \eta_i \mid f) = \sigma_\varepsilon^2 \frac{1}{T} \sum_t (f_t - \bar{f}) / \sum_s (f_s - \bar{f})^2 = 0$, pentru că $\sum_t (f_t - \bar{f}) = 0$
2. Cu restricția, $\bar{R}_i = \lambda_0 + \lambda_T \beta_i + \bar{\varepsilon}_i$: media realizată a factorului $\bar{f}$ este comună tuturor activelor și nu dispare prin mediere cînd $N \to \infty$
   - Condiționat de f, tripletele $(\beta_i, \eta_i, \bar{\varepsilon}_i)$ sînt i.i.d. între active: $\widehat{\text{Cov}}_{cs}(\bar{R}, \hat{\beta}) \to \lambda_T \text{Var}_{cs}(\beta)$ și $\widehat{\text{Var}}_{cs}(\hat{\beta}) \to \text{Var}_{cs}(\beta) + \sigma_\varepsilon^2 / (T \hat{\sigma}_f^2)$
   - Raportul este limita cerută: atenuarea lui λ_T spre zero; $\lambda_T - \lambda = \bar{f} - \mu_f$ dispare doar cînd $T \to \infty$

A6 [Rezolvat]: cifrele

3. 25 de portofolii: zgomotul $9{,}33/(757 \times 19{,}87) = 0{,}00062$; $\text{Var}_{cs}(\beta) = 0{,}01905 - 0{,}00062 = 0{,}01843$; factorul $0{,}01843/0{,}01905 = 0{,}967$

   9 sectoare: zgomotul $13{,}45/(331 \times 20{,}36) = 0{,}00200$; $\text{Var}_{cs}(\beta) = 0{,}09048$; factorul $0{,}978$

- Deplasarea este aici doar de 2–3%: portofoliile diversificate au variață reziduală mică, iar eșantionul este lung; acțiunile individuale au σ_ε^2 mult mai mare, deci atenuarea crește
- Shanken (1992) corectează eroarea standard a lui $\hat{\lambda}$ pentru zgomotul din $\hat{\beta}$ (o corecție de variață); nu elimină deplasarea estimației punctuale în eșantion finit



MFM_ch3_seminar

A7 [Propus]: un factor inutil în etapa 2

- **Întrebarea**: de ce poate primi în etapa 2 o primă aparent semnificativă un factor fără legătură cu randamentele?
- **Ipoteze**: N fix, $T \to \infty$; R_t i.i.d. în timp, cu mediile μ_i, covariață nesingulară $\mathbf{\Sigma}$, momente de ordinul patru finite; un factor g_t i.i.d., cu media 0 și varianța σ_g^2, independent de toate randamentele, deci $\beta_{i,g} = 0$; în demonstrație g este singurul factor din etapa 2
- **Cerințe**
  1. Arătați că $\sqrt{T}\,\hat{\beta}_{i,g} \to N(0, \sigma_i^2/\sigma_g^2)$: $\hat{\beta}_{i,g}$ este de ordinul $T^{-1/2}$ (numărătorul și numitorul separat).
  2. Scrieți $\hat{\lambda}_g = \widehat{\mathrm{Cov}}_{cs}(\bar{R}_i, \hat{\beta}_{i,g})/\widehat{\mathrm{Var}}_{cs}(\hat{\beta}_{i,g})$ cu $z = \sqrt{T}\,\hat{\boldsymbol{\beta}}_g$; arătați că $\hat{\lambda}_g$ este de ordinul $\sqrt{T}$ cînd mediile μ_i diferă.
  3. Cazul mediilor egale: folosind limita comună a lui $(\sqrt{T}(\bar{R} - \boldsymbol{\mu}), z)$, arătați că $\hat{\lambda}_g$ converge la o variabilă aleatoare nedegenerată, nu la 0.
  4. Simularea din Capitolul 3 adaugă g la CAPM pe cele 25 de portofolii (2 000 de replicări): $\lambda_g = 0$ este respins în 59,6% dintre replicări cu erori Fama–MacBeth și în 26,6% cu erori Shanken, iar testul Wald din prima etapă pentru $\beta_{i,g} = 0$ respinge în 6,7%; explicați aceste rate și precizați ce test din prima etapă protejează împotriva lor.
- **De raportat**: cele trei rezultate asimptotice și explicația ratelor de respingere
- Model: A6; Notebook: secțiunea A7

Înapoi Harta exercițiilor

B5 [Propus]: Fama–MacBeth pe ETF-uri

- **Întrebarea**: există prime de piață, de mărime și de valoare în 16 ETF-uri americane în 2013–2026?
- **Date**: randamente lunare în exces, august 2013 – iulie 2026 ($T = 156$), un panel 156×16
  - 9 ETF-uri sectoriale
  - 7 ETF-uri factoriale: MTUM (momentum), VLUE (valoare), QUAL (calitate), USMV (volatilitate minimă), IWM (firme mici, Russell 2000), IWD (firme mari de valoare, Russell 1000 Value), IWF (firme mari de creștere, Russell 1000 Growth)
  - factorii FF3: Mkt-RF, SMB și HML din setul cu trei factori din Kenneth French Data Library(156×3)
- **Cerințe**
  1. Etapa 1: estimați, pe tot eșantionul, regresia fiecărui ETF pe cei trei factori; raportați $\hat{B}$ (16×3).
  2. Etapa 2: în fiecare lună, estimați regresia OLS a celor 16 randamente pe $X = [1, \hat{B}]$; prezentați complet calculul pentru o lună (valori ajustate față de valori observate), apoi calculați media celor 156 de pante; anualizați prin înmulțire cu 12.
  3. Calculați erorile standard Fama–MacBeth, Newey–West pe $\hat{\lambda}_t$ (6 laguri) și Shanken (slide-ul „Regresiile în două etape (Fama–MacBeth)"; pentru termenul liber, bordați covariança factorilor cu o linie și o coloană nule: $\mathrm{Var}_{Sh}(\hat{\lambda}_0) = (1 + c)\,\mathrm{Var}_{FM}(\hat{\lambda}_0)$); ca extindere, calculați erorile Kan–Robotti–Shanken cu funcția `two_pass_gmm` din notebook.
  4. Comparați fiecare $\hat{\lambda}$ cu media în timp a factorului său, într-un grafic, și testați $\lambda = 0$ la 5%.
- **De raportat**: $\hat{B}$, luna calculată complet, tabelul primelor cu erorile lor standard, graficul și deciziile testelor
- **Interpretare**: este o primă nesemnificativă aici dovadă că mărimea și valoarea nu au primă de risc?
- Model: B1, B2 și Capitolul 3, Secțiunea 6; Notebook: secțiunea B5

◂ Înapoi ▸ Harta exercițiilor

B6 [Propus]: o verificare Lewellen–Nagel–Shanken

- **Întrebarea**: se menține R^2 transversal mare al FF3 cu active de test mai variate și cu o măsură GLS (Lewellen, Nagel & Shanken (2010))?
- **Date**: 25 de portofolii mărime × B/M și 30 de portofolii pe industrii (Kenneth French Data Library), randamente lunare în exces, iulie 1963 – iulie 2026 ($T = 757$); factorii FF3
- **Definiții**: μ, X, $\hat{V}$, R^2_{OLS} și R^2_{GLS} ca pe slide-ul „Ajustarea transversală, OLS și GLS"
- **Cerințe**
  1. Punct de plecare: FF3 pe cele 25 de portofolii; reprezentați mediile observate față de cele estimate de model; raportați ambele R^2.
  2. Adăugați cele 30 de industrii și repetați; impuneți și $\lambda = \mathbb{E}[f]$ fără termen liber și raportați R^2_{OLS} cu același numitor.
  3. Incertitudine: reeșantionați lunile i.i.d. (500 de extrageri, inițializare 42), reestimați ambele etape și $\hat{V}$ la fiecare extragere; raportați intervale de 95%.
  4. Extindere: repetați analiza pentru CAPM și FF5 pe ambele mulțimi de active.
- **De raportat**: tabelul celor două R^2 cu intervalele lor, pe ambele mulțimi de active, și graficul
- **Interpretare**: se menține R^2 al FF3 după adăugarea industriilor?
- Model: B2, Capitolul 3, Secțiunea 6; Notebook: secțiunea B6

Înapoi Harta exercițiilor

B7 [Propus]: GRS fără normalitate

- **Întrebarea**: este respingerea CAPM pe cele 25 de portofolii un artefact al ipotezei că reziduurile au distribuția Normală?
- **Date**: 25 de portofolii mărime × B/M, CAPM cu Mkt-RF, lunar, iulie 1963 – iulie 2026 ($T = 757$, $N = 25$)
- **Cerințe**
  1. Diagnostic: estimați cele 25 de regresii; calculați kurtosis-ul Pearson al fiecărei serii de reziduuri (slide-ul „Normalitatea reziduurilor") și media lui; desenați un grafic Q–Q.
  2. Testul exact: calculați GRS cu $\hat{\boldsymbol{\Sigma}}$ ML și valoarea p din $F(25, 731)$.
  3. Bootstrap sub H_0: reeșantionați rînduri întregi de reziduuri (păstrînd corelația lor transversală), reconstruiți $\mathrm{R}_t^* = \hat{\boldsymbol{\beta}} f_t + \boldsymbol{\varepsilon}_t^*$ (deci $\alpha = 0$), reestimați și recalculați GRS de 2 000 de ori, inițializare 42; raportați $p = (1 + \#\{W^* \geq W\})/2001$ și valoarea critică bootstrap de 95%.
  4. Extindere: calculați testul Wald robust cu momentele $(1, f_t)^\top \otimes \boldsymbol{\varepsilon}_t$, parametrii ordonați $(\boldsymbol{\alpha}, \boldsymbol{\beta})$; $W = \hat{\boldsymbol{\alpha}}^\top \hat{\mathrm{V}}_{\alpha}^{-1} \hat{\boldsymbol{\alpha}} \to \chi^2_{25}$ cu sandwich-ul $\mathrm{D}^{-1}\hat{\mathrm{S}}\mathrm{D}^{-1}/T$, $\mathrm{D} = (\mathrm{Z}^\top \mathrm{Z}/T) \otimes \mathrm{I}_{25}$, $\mathrm{Z} = [1, \mathrm{f}]$, iar $\hat{\mathrm{S}}$ White sau Newey–West (6 laguri).
  5. Explicați de ce GRS nu se poate calcula pentru cele 100 de portofolii mărime × B/M în august 2018 – iulie 2026 ($T = 96$).
- **De raportat**: kurtosis-ul mediu, graficul Q–Q, W cu valorile p ale celor trei metode și valoarea critică bootstrap
- **Interpretare**: schimbă abaterea reziduurilor de la distribuția Normală decizia testului GRS?
- Model: B2; Notebook: secțiunea B7

◂ Înapoi ▸ Harta exercițiilor

B8 [Propus]: compararea modelelor prin rapoarte Sharpe

- **Întrebările**
  - ce model factorial atinge cel mai mare raport Sharpe?
  - este diferența semnificativă (Barillas & Shanken (2018); Barillas, Kan, Robotti & Shanken (2020))?
- **Date**: Mkt-RF, SMB, HML, RMW, CMA, MOM lunari (Kenneth French Data Library), iulie 1963 – iulie 2026 ($T = 757$) și ianuarie 2000 – iulie 2026 ($T = 319$)
- **Componența**: CAPM (Mkt-RF); FF3 și Carhart folosesc SMB din setul cu trei factori, FF5 și FF5 + MOM pe cel din setul cu cinci; covariațele se împart la $T - 1$
- **Cerințe**
  1. Calculați pătratul raportului Sharpe maxim, SR^2, pentru fiecare model (slide-ul „Compararea modelelor factoriale prin rapoarte Sharpe"), anualizați ($\times 12$) și raportați $\sqrt{SR^2}$.
  2. Modele incluse unul în altul, FF5 + MOM față de FF5: regresați MOM pe cei cinci factori FF5 și testați termenul liber (Newey–West).
  3. Modele neincluse unul în altul, FF5 față de Carhart: aplicați un bootstrap staționar (Politis & Romano (1994)) cu bloc mediu de 6 luni și 2 000 de extrageri, inițializare 42, cu aceiași indici pentru ambele modele; raportați intervalul de 95% al diferenței SR^2.
  4. Extindere: calculați t de spanning pentru fiecare dintre cei șase factori pe ceilalți cinci, în ambele perioade.
- **De raportat**: rapoartele Sharpe în ambele perioade, t de spanning și intervalul bootstrap
- **Interpretare**
  - de ce nu poate testa egalitatea intervalul percentil al unei diferențe între modele incluse unul în altul?
  - ce factori nu adaugă un alfa semnificativ după 2000?
- Model: B1; Notebook: secțiunea B8

◂ Înapoi ▸ Harta exercițiilor

B9 [Propus]: testarea multiplă a 25 de valori alfa

- **Întrebarea**: cîte dintre cele 25 de portofolii mărime × B/M au un alfa FF3 real?
- **Date**: 25 de portofolii mărime × B/M ponderate după valoare (setul „25 Portfolios Formed on Size and Book-to-Market" din Kenneth French Data Library) și factorii FF3 (Fama & French (1993)), randamente lunare în exces, iulie 1963 – iulie 2026 ($T = 757$)
- **Cerințe**
  1. Pentru un portofoliu (small-growth), prezentați calculul complet: regresia FF3, t Newey–West al lui α (6 laguri), p bilateral $= 2(1 - \Phi(|t|))$, Φ: funcția de repartiție a distribuției Normale standard.
  2. Repetați pentru toate cele 25; desenați o hartă 5×5 a valorilor alfa anuale, împreună cu statisticile t corespunzătoare.
  3. Ordonați valorile p și aplicați fișa din A1: naiv, Bonferroni, Holm, Benjamini–Hochberg; extindere: Benjamini–Yekutieli.
  4. Numărați portofoliile cu $|t| > 3$ (Harvey, Liu & Zhu (2016)), un prag euristic, nu o procedură cu o rată de eroare garantată.
- **De raportat**: calculul pentru small-growth, harta, numărul de descoperiri al fiecărei proceduri și numărul de portofolii cu $|t| > 3$
- **Interpretare**: cele 25 de teste sînt puternic corelate
  - ce metode rămîn valide?
  - ce test folosește direct corelația?
- Model: A1; Notebook: secțiunea B9

◂ Înapoi ▸ Harta exercițiilor

A8 [Propus]: active cu corelații egale

- **Întrebarea**: cum depinde ponderea primei componente de corelația medie și de N?
- **Etapa 1**, $N = 3$, $\rho = 0{,}5$: $\mathrm{C} = \begin{pmatrix} 1 & 0{,}5 & 0{,}5 \\ 0{,}5 & 1 & 0{,}5 \\ 0{,}5 & 0{,}5 & 1 \end{pmatrix}$
  1. Verificați că $\mathbf{v_1} = (1,1,1)/\sqrt{3}$, $\mathbf{v_2} = (1,-1,0)/\sqrt{2}$, $\mathbf{v_3} = (1,1,-2)/\sqrt{6}$ sînt vectori proprii ortonormați și găsiți valorile proprii.
  2. Calculați ponderea PC1, $\ell_1/3$.
- **Etapa 2**, N active, $\mathrm{C} = (1-\rho)\mathrm{I} + \rho 1 1^\top$
  3. Arătați că $\ell_1 = 1 + (N-1)\rho$ și că toate celelalte valori proprii sînt egale cu $1-\rho$; găsiți limita ponderii PC1 cînd $N \to \infty$.
  4. Inversați formula ponderii pentru a obține ρ echivalent pentru o pondere PC1 s; aplicați-o la acțiunile blue chip de la BVB ($N = 10$, $s = 0{,}425$, B10) și la sectoarele din SUA ($N = 11$, $s = 0{,}660$, B3).
  5. Extindere: pentru o C oarecare cu corelația medie în afara diagonalei $\bar{\rho}$, arătați că $1 + (N-1)\bar{\rho} \leq \ell_1$.
- **De raportat**: valorile proprii, formula ponderii PC1 cu limita ei și cele două valori ρ
- Model: B3; Notebook: secțiunea A8

Înapoi Harta exercițiilor

B10 [Propus]: PCA pe blue chips BVB

- **Întrebarea**: este primul factor al acțiunilor blue chip românești la fel de puternic ca la sectoarele din SUA?
- **Date**: 10 blue chips (TLV, SNP, BRD, TGN, FP, SNG, SNN, EL, DIGI, TEL; H2O exclus, listat în 2023) și BET, aliniate pe zilele comune de la 7 iunie 2017 la 18 septembrie 2026; randamente logaritmice zilnice; eliminați zilele cu vreun $|r_t| > 0{,}5$ și raportați cîte; BET rămîne în afara PCA
- **Cerințe**
  1. Aplicați procedura din B3: standardizare ($T \times 10$), valori proprii, ponderi de variantă, încărcările PC1.
  2. Corelați scorurile PC1 cu randamentele BET.
  3. Estimați numărul de factori prin Kaiser, Bai–Ng și Ahn–Horenstein, cu $k_{\max} = 5$.
  4. Comparați cu sectoarele din SUA pe aceeași perioadă (de la 20 iunie 2018): ponderile PC1 și graficele scree alăturate.
- **De raportat**: numărul de zile eliminate, ponderile, încărcările PC1, corelația cu BET, cele trei estimații ale lui k și graficele scree
- **Interpretare**
  - ce implică un prim factor mai slab pentru diversificarea în cadrul BVB?
  - cît din diferență provine din compararea unor acțiuni cu portofolii sectoriale?
- Model: B3; Notebook: secțiunea B10

◂ Înapoi ▸ Harta exercițiilor

B11 [Propus]: alfa ETF-urilor factoriale

- **Întrebarea**: obțin MTUM și QUAL un alfa peste expunerile lor la cei șase factori?
- **Date**: prețurile ajustate ale MTUM (18 apr. 2013 – 31 iul. 2026) și QUAL (18 iul. 2013 – 31 iul. 2026); FF5 (Fama & French (2015)), MOM (Carhart (1997)) și RF zilnici, aceeași dată de final
- **Regula calendarului**: păstrați un randament al ETF-ului doar dacă prețul lui anterior datează din ziua de tranzacționare imediat anterioară din calendarul factorilor
  - Exemplu: MTUM are prețuri pe 18 și 23 aprilie 2013, dar nu pe 19 și 22 aprilie; randamentul 18 → 23 aprilie acoperă trei zile de factori și se elimină
- **Cerințe**
  1. Construiți randamentele zilnice în exces cu regula calendarului; raportați n înainte și după aliniere.
  2. Estimați $R_t^e = \alpha + b\,\text{Mkt-RF}_t + s\,\text{SMB}_t + h\,\text{HML}_t + r\,\text{RMW}_t + c\,\text{CMA}_t + m\,\text{MOM}_t + \varepsilon_t$; anualizați α prin înmulțire cu 252.
  3. Calculați eroarea standard Newey–West (10 laguri) și t pentru α; aplicați un bootstrap pe blocuri suprapuse ale perechilor, cu blocuri de 20 de zile și 1 000 de extrageri, inițializare 42, și raportați intervalul de 95%.
  4. Reprezentați cele două intervale pentru fiecare ETF.
- **De raportat**: n înainte și după aliniere, cele șase încărcări, α cu e.s. și t, cele două intervale și graficul
- **Interpretare**: de ce nu trebuie regresat niciodată un randament ETF calculat pe 3 zile în funcție de un randament factorial de o zi?
- Model: B1; Notebook: secțiunea B11

◂ Înapoi ▸ Harta exercițiilor

C2 [Propus]: prime factoriale la BVB — cadrul

- **Întrebarea**: generează momentum (Jegadeesh & Titman (1993)) și beta mic prime în rîndul acțiunilor blue chip de la BVB?
- **Date**: prețuri ajustate de sfîrșit de lună ale celor 10 blue chips din C1 și BET, decembrie 2015 – august 2026 (septembrie 2026 incomplet); randamente lunare simple; prețurile FP sînt ajustate pentru dividendul special din 7 septembrie 2023 (1,7225 lei pe acțiune)
- **Reguli fixate dinainte**
  - semnalul și beta folosesc informația pînă la sfîrșitul lui $t-1$ (cronologia din C1)
  - cel puțin 6 acțiuni disponibile; egalitățile se departajează după simbol
  - ponderile rămîn cele stabilite la formare; un randament de deținere lipsă (suspendare, cotație lipsă, delistare) nu se exclude din media părții: la delistare se include valoarea terminală efectivă; altfel luna se marchează incompletă și se raportează cît de des apare

Clasamentul pentru august 2026

Acțiunea	Semnal	Parte
EL	129,9%	long
TGN	127,3%	long
DIGI	113,2%	long
SNG	88,9%	
BRD	87,8%	
SNN	75,5%	
TEL	66,1%	
FP	51,0%	short
TLV	48,5%	short
SNP	46,6%	short

- Semnalul: randamentul compus din august 2025 – iunie 2026, cunoscut la sfîrșitul lui iulie 2026

Înapoi Harta exercițiilor

C2 [Propus]: prime factoriale la BVB — cerințe

- **Cerințe**
  1. Construiți momentum 12–1 (specificația din C1): long primele 3, short ultimele 3, ponderi egale.
  2. Construiți portofoliul beta mic: estimați beta pe BET pe ferestre mobile de 36 de luni (cel puțin 18 luni, pînă la $t-1$); long cele 3 cu beta cel mai mic, short cele 3 cu beta cel mai mare; un spread nescalat, nu BAB-ul neutru față de beta al lui Frazzini & Pedersen (2014).
  3. Raportați perioadele, T, media anuală, volatilitatea și t Newey–West (3 laguri); comparați t cu 1,96 și 3; reprezentați spread-urile cumulate.
  4. Discutați deplasarea de supraviețuire (blue chips de azi, alese retrospectiv), restricțiile la vînzarea în lipsă, costurile și lichiditatea.
  5. Extindere: identificați surse publice pentru capitalul propriu contabil și numărul de acțiuni, necesare unui factor de valoare la BVB.
- **Livrabile**: un folder Quantlet în repository-ul propriu (script, sursa datelor și descriere scurtă) și un slide cu concluzia și principala ei limită
- **Criterii de evaluare**
  - semnale formate doar cu informația disponibilă la sfîrșitul lunii $t-1$
  - statistica t Newey–West comparată cu 1,96 și cu 3, cu decizia precizată
  - limitele discutate: deplasarea de supraviețuire, vînzarea în lipsă, costurile
  - critica AI din C1 documentată în `AI_ERRORS.md`
- Model: B4, C1; Notebook: Partea C

C3 [Propus]: replicarea factorilor americani, un proiect în etape (1/2)

- **Context**: studiul de caz din Capitolul 3, Jensen, Kelly & Pedersen (2023): 82,4% dintre factorii americani păstrează un alfa CAPM semnificativ într-un model ierarhic Bayes empiric, față de 75,6% cu Benjamini–Yekutieli și 35% la Hou, Xue & Zhang (2020) pe randamente brute (Figura 1)
- **Datele**: portofoliile sortate după o singură caracteristică, ponderate după capitalizare, din Kenneth French Data Library, iulie 1963 – iulie 2026, și Mkt-RF din setul cu trei factori
- Fiecare factor cumpără un capăt al sortării și îl vinde pe celălalt, pentru 15 caracteristici
  - 30% la fiecare capăt: mărime, book-to-market, profitabilitate operațională, investiții, profit/preț, cash flow/preț, dividend/preț
  - 20% la fiecare capăt: acumulări, emisiuni nete de acțiuni, beta de piață, variantă, variantă reziduală
  - 10% la fiecare capăt: randamentele anterioare din lunile 12–2 (momentum 12–1, slide-ul „Momentum și deplasarea de anticipare"), 1–0 și 60–13
- **Etapa 1, construcția și ratele frecventiste**
  1. Construiți cei 15 factori, cu poziție long pe partea cu randamentul mai mare în articolul original; citați acel articol pentru fiecare semn și notați anul final al eșantionului și data publicării în tabelul factorilor.
  2. Regresați fiecare factor pe Mkt-RF, rescalați-l la o volatilitate reziduală lunară de $10\%/\sqrt{12}$ (Secțiunea II.A a articolului), raportați $\hat{\alpha}$, t și rata de replicare OLS ($t \geq 1{,}96$).
  3. Aplicați Benjamini–Yekutieli la 5% și raportați $|t|$ critic și rata.

Înapoi Harta exercițiilor

C3 [Propus]: replicarea factorilor americani, un proiect în etape (2/2)

- **Etapa 2, teme și modelul ierarhic**
  4. Construiți dendrograma Ward pe distanța $1-$ corelația reziduurilor CAPM (autorii: R `hclust(method="ward.D")` cu $1-\rho$; în SciPy, `linkage(method="ward")` cu distanțele condensate $\sqrt{2(1-\rho)}$, echivalentul R `ward.D2`; precizați perechea distanță–algoritm); comparați-o cu temele articolului (`Cluster Labels.csv`, codul autorilor).
  5. Construiți M ($15 \times J$), matricea apartenenței la teme; estimați τ_c, τ_w prin verosimilitate maximă pentru $\hat{\boldsymbol{\alpha}} \sim N(0, \mathrm{MM}^\top \tau_c^2 + \mathrm{I}\tau_w^2 + \hat{\boldsymbol{\Sigma}}/T)$, raportați dacă estimația atinge frontiera $\tau_c = 0$; calculați mediile și variațele a posteriori (Propoziția 4) și rata Bayes empiric (z a posteriori $\geq 1{,}96$).
- **Etapa 3, sensibilitate**
  6. Calculați rata Bayes empiric cu o singură temă.
  7. Reprezentați rata în funcție de τ_c și calculați FDR bayesian (Propoziția 5).
- **Etapa 4, scăderea după eșantion**
  8. Estimați valorile alfa în anii eșantionului original (de la iulie 1963) și după, cu cel puțin 60 de luni în fiecare.
  9. Regresați prin GLS fezabil, cu inversa covarianței factorilor, valorile alfa ulterioare pe cele din eșantion și comparați panta cu 0,26 (Figura 8).
- **De raportat**: un tabel al factorilor (an final al eșantionului, data publicării, alfa, t, temă, media a posteriori, z), ratele, alături de cele din Figura 1, $\hat{\tau}$, alături de valorile din Tabelul I, dendrograma și panta; păstrați separat cifrele articolului și cele ale cursului
- Model: A3, A1 și studiul de caz din Capitolul 3

Acronime folosite în acest material (1/2)

- **AI**: Artificial Intelligence — *inteligență artificială*
- **BAB**: Betting Against Beta — *strategia „pariu împotriva lui beta"*
- **BET**: Bucharest Exchange Trading (index) — *indicele de referință al Bursei de Valori București*
- **BH**: Benjamini–Hochberg (procedure) — *procedura Benjamini–Hochberg*
- **BY**: Benjamini–Yekutieli (procedure) — *procedura Benjamini–Yekutieli*
- **BVB**: Bursa de Valori București
- **CAPM**: Capital Asset Pricing Model — *modelul de evaluare a activelor financiare*
- **CMA**: Conservative Minus Aggressive (investment factor) — *factorul investiții: firme conservatoare minus agresive*
- **CML**: Capital Market Line — *dreapta pieței de capital*
- **ETF**: Exchange-Traded Fund — *fond tranzacționat la bursă*
- **FDR**: False Discovery Rate — *rata descoperirilor false*
- **FF3**: Fama–French three-factor model — *modelul Fama–French cu trei factori*
- **FF5**: Fama–French five-factor model — *modelul Fama–French cu cinci factori*
- **FM**: Fama–MacBeth (regressions) — *regresiile Fama–MacBeth*
- **FWER**: Family-Wise Error Rate — *probabilitatea de cel puțin o eroare de tip I în familia de teste*
- **GLS**: Generalised Least Squares — *metoda celor mai mici pătrate generalizate*
- **GMM**: Generalized Method of Moments — *metoda generalizată a momentelor*
- **GRS**: Gibbons–Ross–Shanken (test) — *testul Gibbons–Ross–Shanken*

Acronime folosite în acest material (2/2)

- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors) — *erori standard robuste la heteroscedasticitate și autocorelație*
- **HML**: High Minus Low (value factor) — *factorul valoare: B/M mare minus mic*
- **KRS**: Kan–Robotti–Shanken (standard errors) — *erorile standard Kan–Robotti–Shanken*
- **ML**: Maximum Likelihood — *verosimilitate maximă*
- **MOM**: Momentum (factor) — *factorul momentum*
- **NW**: Newey–West (standard errors) — *erori standard Newey–West*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **PC1**: first Principal Component — *prima componentă principală*
- **PC2**: second Principal Component — *a doua componentă principală*
- **PC3**: third Principal Component — *a treia componentă principală*
- **PCA**: Principal Component Analysis — *analiza componentelor principale*
- **RF**: Risk-Free rate — *rata fără risc*
- **RMW**: Robust Minus Weak (profitability factor) — *factorul profitabilitate: firme robuste minus slabe*
- **SMB**: Small Minus Big (size factor) — *factorul mărime: firme mici minus mari*
- **SML**: Security Market Line — *dreapta pieței titlurilor*
- **SUA**: Statele Unite ale Americii

Bibliografie (1/4)

- Ahn, S.C., Horenstein, A.R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3), 1203–1227.
- Bai, J., Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.
- Barillas, F., Kan, R., Robotti, C., Shanken, J. (2020). Model comparison with Sharpe ratios. *Journal of Financial and Quantitative Analysis*, 55(6), 1840–1874.
- 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.
- 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.

Bibliografie (2/4)

- Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2), 197–226.
- 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. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105(3), 457–472.
- 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.
- 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.
- Härdle, W.K., Simar, L. (2019). Applied Multivariate Statistical Analysis, ed. a 5-a. Springer.
- 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.

Bibliografie (3/4)

- Hou, K., Xue, C., Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies*, 33(5), 2019–2133.
- 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, T.I., Kelly, B., Pedersen, L.H. (2023). Is there a replication crisis in finance? *Journal of Finance*, 78(5), 2465–2518.
- 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.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lewellen, J., Nagel, S., Shanken, J. (2010). A skeptical appraisal of asset pricing tests. *Journal of Financial Economics*, 96(2), 175–194.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7(1), 77–91.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.

Bibliografie (4/4)

- 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.
- 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.
- Ward, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838.
- Date: Kenneth R. French Data Library, mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.