








Modelarea piețelor financiare

Seminarul 10: Microstructura pieței — aplicații



Daniel Traian PELE
Drd. Antoaneta Amza, antoaneta.amza@csie.ase.ro

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

Cît de precis măsurăm costurile de tranzacționare?

- Întrebarea seminarului
  - **cît de precis pot fi măsurate costurile de tranzacționare și lichiditatea din prețuri și volume publice?**
- Azi, în această ordine: A1, A2, B1, B2, B3, B4, C1
- Livrabile
  - răspunsurile la exercițiile A
  - cîte o figură și un tabel pentru fiecare exercițiu B
  - codul corectat din C1
  - o frază de interpretare pentru fiecare exercițiu
- Notebook-ul de azi: `chapter10_seminar_notebook.ipynb` (Colab)

Ce observăm în fiecare bară?

- OHLC înseamnă Open, High, Low, Close: deschiderea, maximul, minimul și închiderea unui interval. Volumul numără acțiunile tranzacționate; valoarea tranzacționată depinde și de preț.
- UTC înseamnă Coordinated Universal Time, timpul universal coordonat. Convertim orele fondului la ora New York înainte de selectarea ședinței; Bitcoin păstrează UTC.
- Un tick este pasul minim de cotare. Un punct de bază (pb; bp în graficele englezești) este o sutime de punct procentual: 25 pb înseamnă 0,25 la sută. Tick-ul este în unități de preț, iar pb exprimă o variație relativă.
- Eroarea standard (SE, *Standard Error*) a unei estimări: cît ar varia estimarea de la un eșantion la altul.

Ce piețe comparăm?

- SUA (Statele Unite ale Americii; US în graficele englezești), USD (dolarul american). Un ETF (*Exchange-Traded Fund*) este un fond tranzacționat la bursă; SPY (SPDR S&P 500 ETF Trust) este ETF-ul pe indicele S&P 500 (*Standard & Poor's* 500).
- Acțiunile americane, după simbolul bursier (Nasdaq sau NYSE): AAPL (Apple), MSFT (Microsoft), NVDA (NVIDIA), JPM (JPMorgan Chase), GME (GameStop), MSTR (Strategy, fosta MicroStrategy).
- BVB (Bursa de Valori București), BET (indicele de referință al BVB, *Bucharest Exchange Trading*), BET-TR (indicele BET cu dividendele reinvestite, *Total Return*), TVBETETF (ETF-ul BET Patria–Tradeville, listat la BVB, care replică indicele BET), BNR (Banca Națională a României), RON (leul românesc).
- VIX (indicele de volatilitate al bursei Cboe, fosta CBOE, *Chicago Board Options Exchange*) se calculează din cotațiile opțiunilor: este volatilitatea implicită, nu cea realizată în eșantionul nostru. Cripto: BTC (Bitcoin), ETH (Ether, rețeaua Ethereum), SOL (Solana), XRP (tokenul registrului XRP Ledger).

Ce companii românești identifică simbolurile?

- Simbolurile bursiere BVB: TLV (Banca Transilvania), BRD (BRD–Groupe Société Générale), SNP (OMV Petrom), H2O (Hidroelectrica), SNG (Romgaz).
- SNN (Nuclearelectrica), TGN (Transgaz), TEL (Transelectrica), EL (Electrica), DIGI (Digi Communications). OMV vine de la *Österreichische Mineralölverwaltung*, grupul austriac OMV.
- Comparăm lichiditatea după conversia valorilor tranzacționate în aceeași monedă și după verificarea tratării dividendelor și divizărilor de acțiuni.

Datele folosite

Seria	Conținut	Fereastra	Unități
SPY, bare de 5 minute	OHLC și volum; ora UTC convertită la ora New York, barele 09:30–15:55, doar zilele cu toate cele 78 de bare	12 oct. 2020 – 18 sep. 2026, 1 477 de zile	USD; volum în acțiuni
Bitcoin, bare de 5 minute	OHLC, ora UTC	31 aug. 2024 – 18 sep. 2026	USD
SUA, zilnic (7)	SPY, AAPL, MSFT, NVDA, JPM, GME, MSTR	19 sep. 2024 – 18 sep. 2026	USD; volum în acțiuni
BVB, zilnic (11)	TLV, SNP, H2O, BRD, SNG, SNN, TGN, EL, DIGI, TEL și ETF-ul pe BET TVBETETF	la fel	lei; convertiți în USD la cursul BNR al zilei
Cripto, zilnic (7)	BTC, ETH, SOL, XRP, Dogecoin, Cardano, Litecoin	la fel, 7 zile din 7	volumul este deja în USD
VIX; BET-TR	VIX (indicele CBOE al volatilității implicite) și BET-TR (indicele BET cu reinvestirea dividendelor), închideri zilnice	ca în B4, B8	puncte de indice

- Seriile zilnice: deschidere, maxim, minim, închidere, închidere ajustată și volum; OHLC scalate cu raportul închidere ajustată / închidere (dividende și divizări); valoarea tranzacționată = închiderea ajustată pentru divizări × volumul; zilele cu maxim = minim sau volum nul eliminate
- Cursul de referință BNR (lei pentru un USD), fixingul zilnic oficial; 1 pb = 0,01%

Pregătire: rulați și verificați

```
# functiile care incarca datele sint in sectiunea
      Date a notebook-ului
s = intraday_spy()          # SPY 5-minute bars
print(len(s), s['date'].nunique())
print(s.index[0], s.index[-1])
b = intraday_btc()          # Bitcoin, UTC
print(len(b), b['date'].nunique())
r = returns('TLV', '2024-09-19')
dv = dollar_volume('TLV', '2024-09-19')
print(len(r), dv.median() / 1e6)
d = ohlc('SPY', '2024-09-19')
print(len(d))  # daily Open, High, Low, Close
     bars
```

- Rezultatul așteptat
  - 115 206 bare SPY, 1 477 de zile
  - prima bară 12 oct. 2020 09:30, ultima 18 sep. 2026 15:55 (ora New York)
  - 213 754 de bare Bitcoin în 749 de zile
  - 492 de randamente zilnice pentru Banca Transilvania; valoarea mediană tranzacționată 3,46 milioane USD
- SPY zilnic: 501 rînduri OHLC
- Dacă un număr diferă, opriți-vă și verificați seria, fusul orar și data de început

Notația folosită în toate exercițiile

- p_t: prețul tranzacției (sau închiderea barei); m_t, q_t, c, $s = 2c$: slide-urile „Bid, ask și spread-ul” și „Autocovariațele și modelul Roll”; spread-urile relative în pb
- $\hat{\gamma}_0$, $\hat{\gamma}_1$: autocovariațele de selecție ale variațiilor de preț, centrate pe media din eșantion (slide-ul „Autocovariațele și modelul Roll”); T = numărul variațiilor de preț (78 de închideri dau 77 de variații)
- Randamentele intrazilnice: variațiile logaritmului închiderii în cursul zilei; prima bară pornește de la deschidere; $|r|$ în pb
- Raportul de ilichiditate Amihud (ILLIQ): slide-ul „Ilichiditatea: raportul Amihud”; bootstrap-ul pe zile și pe blocuri: slide-ul „Reeșantionare: bootstrap pe zile, pe blocuri, Monte Carlo”; toate reeșantionările folosesc sămînța generatorului aleator (seed) 42
- Un exemplu de calcul: 78 de închideri, $\hat{\gamma}_1 = -1{,}0 \cdot 10^{-8}$ în unități logaritmice dă $\hat{s} = 2\sqrt{10^{-8}} = 2 \cdot 10^{-4} = 2$ pb

Bid, ask și spread-ul

- Cotațiile: **bid** b_t este cel mai bun preț la care puteți vinde, **ask** a_t cel mai bun preț la care puteți cumpăra

$$m_t = \frac{a_t + b_t}{2}, \qquad c = \frac{a_t - b_t}{2}, \qquad s = 2c = a_t - b_t, \qquad S_t^Q = \frac{a_t - b_t}{m_t}$$

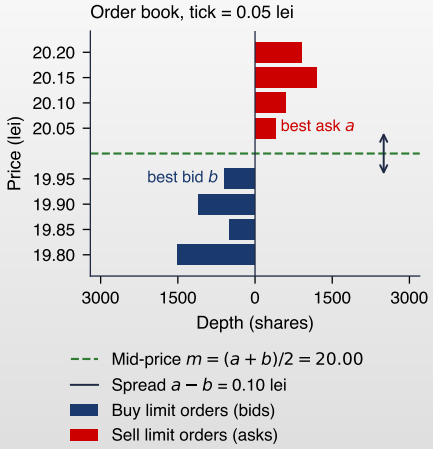
  - m_t: prețul de mijloc (mid); c: jumătatea de spread; S_t^Q: spread-ul relativ cotat, în puncte de bază (1 pb $= 0{,}01\%$)
  - un spread cotat nu este niciodată sub un **tick**, pasul minim de preț

- **Spread-ul efectiv**: cît costă efectiv o tranzacție la prețul p_t

$$S_t^E = \frac{2q_t\,(p_t - m_t)}{m_t}, \qquad q_t = \begin{cases} +1 & \text{tranzacție inițiată de cumpărător} \\ -1 & \text{tranzacție inițiată de vînzător} \end{cases}$$

- Spread-ul îl remunerează pe furnizorul de lichiditate pentru procesarea ordinelor, riscul de stoc și selecția adversă (tranzacții cu parteneri mai bine informați)
- **Tranzacția dus-întors**: cumpărați la a, vindeți la b: costul este $a - b$, **un singur** spread, jumătate pe fiecare sens

Registrul de ordine



- **Ordin limită**: un preț și o cantitate, care așteaptă în registru; **ordin la piață**: se execută imediat la cele mai bune cotații opuse
- Cel mai bun bid b: cel mai mare preț limită de cumpărare; cel mai bun ask a: cel mai mic preț limită de vînzare; adîncimea: acțiunile care așteaptă la fiecare preț
- Exemplu (graficul): $b = 19{,}95$, $a = 20{,}05$ lei, tick 0,05 lei
  - $m = 20$, $s = 0{,}10$ lei $= 0{,}5\% = 50$ pb; spread-ul are doi pași de preț
  - un sens costă $c/m = 25$ pb, tranzacția dus-întors 50 pb
- O cumpărare la piață mai mare decît adîncimea la a preia și nivelurile următoare: prețul ei mediu crește cu mărimea ordinului (impactul asupra prețului)

Autocovariațele și modelul Roll

- Autocovariațele de selecție ale variațiilor de preț $\Delta p_t = p_t - p_{t-1}$, $t = 1, \ldots, T$

$$\hat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T}\left(\Delta p_t - \overline{\Delta p}\right)^2, \qquad \hat{\gamma}_1 = \frac{1}{T-1}\sum_{t=2}^{T}\left(\Delta p_t - \overline{\Delta p}\right)\left(\Delta p_{t-1} - \overline{\Delta p}\right), \qquad \hat{\rho}_1 = \frac{\hat{\gamma}_1}{\hat{\gamma}_0}$$

  - T: numărul variațiilor de preț; $\hat{\gamma}_1$ este media celor $T-1$ produse; $\overline{\Delta p}$: media de selecție (centrarea)
  - γ_0, γ_1: valorile din populație; accentul circumflex: o estimare
- **Oscilația bid-ask** (*bid-ask bounce*): tranzacțiile cad aleator la bid sau la ask, deci variațiile de preț tind să se inverseze
- **Modelul Roll** Roll (1984): un preț eficient m_t de tip mers aleator plus o alternanță de $\pm c$

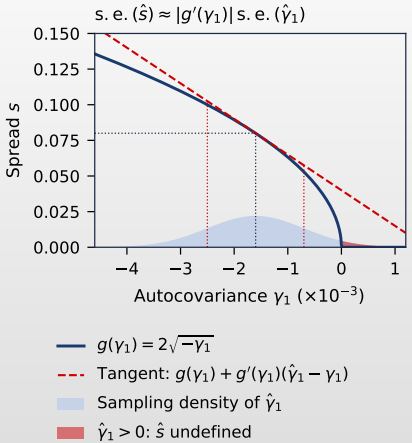
$$m_t = m_{t-1} + \varepsilon_t, \qquad p_t = m_t + c\, q_t, \qquad q_t = \pm 1 \text{ cu probabilitatea } 1/2$$

  - ε_t: știri independente și identic distribuite (i.i.d.), media 0, dispersia σ^2; modelul implică $\gamma_1 = -c^2$ (dedus în A1)

$$\hat{s} = 2\sqrt{-\hat{\gamma}_1} \text{ (definit doar dacă } \hat{\gamma}_1 < 0)$$

  - exemplu: $\hat{\gamma}_1 = -0{,}0004$ lei^2: $\hat{c} = 0{,}02$, $\hat{s} = 0{,}04$ lei; dacă $\hat{\gamma}_1 \geq 0$, raportați „nedefinit” sau trunchiați la 0 și precizați convenția

Precizia estimatorului Roll



- Dispersia de selecție a lui $\hat{\gamma}_1$ în modelul Roll, cu semne $q_t = \pm 1$ independente Harris (1990)

$$T \operatorname{Var}(\hat{\gamma}_1) \to \gamma_0^2 + \gamma_1^2$$

- **Metoda delta**: SE a unei funcții netede g de o estimare (dezvoltare Taylor de ordinul întîi)

$$\operatorname{es}\left(g(\hat{\theta})\right) \approx |g'(\theta)| \operatorname{es}(\hat{\theta})$$

  - graficul: $g = 2\sqrt{-\gamma_1}$ lîngă $\gamma_1 = -0{,}0016$; tangenta transformă ± 1 SE a lui $\hat{\gamma}_1$ în $\hat{s}$
  - lîngă 0 curba este abruptă și aproximarea liniară nu mai funcționează; coada roșie ($\hat{\gamma}_1 > 0$) nu are $\hat{s}$

- **Deplasarea din centrare**: scăderea mediei de selecție împinge $\hat{\gamma}_1$ în jos, deci Roll găsește un spread chiar cînd $c = 0$ (B2)

$$\mathbb{E}[\hat{\gamma}_1] \approx \gamma_1 - \frac{\gamma_0 + 2\gamma_1}{T}$$

Spread-ul din maxime și minime: Corwin–Schultz

- Corwin–Schultz Corwin & Schultz (2012): maximul unei bare este de obicei o cumpărare la ask, minimul o vînzare la bid, deci amplitudinea logaritmică = volatilitate + spread
  - H_t, L_t: maximul și minimul barei (sau zilei) t; $H_{t,t+1}$, $L_{t,t+1}$: maximul și minimul pe ambele bare
  - partea de volatilitate crește cu lungimea intervalului, partea de spread nu
- β: suma pătratelor celor două amplitudini pe o bară; γ: pătratul amplitudinii pe două bare

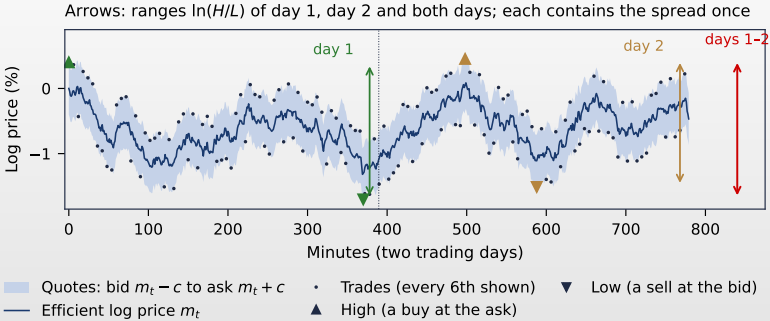
$$\beta = \sum_{j=0}^{1} \left[\ln \frac{H_{t+j}}{L_{t+j}} \right]^2, \qquad \gamma = \left[\ln \frac{H_{t,t+1}}{L_{t,t+1}} \right]^2$$

- α: partea din amplitudinea logaritmică rămasă pentru spread; $\hat{S}$: spread-ul relativ ($\hat{S} \approx \alpha$ pentru α mic)

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \qquad \hat{S} = \frac{2(e^{\alpha} - 1)}{1 + e^{\alpha}}$$

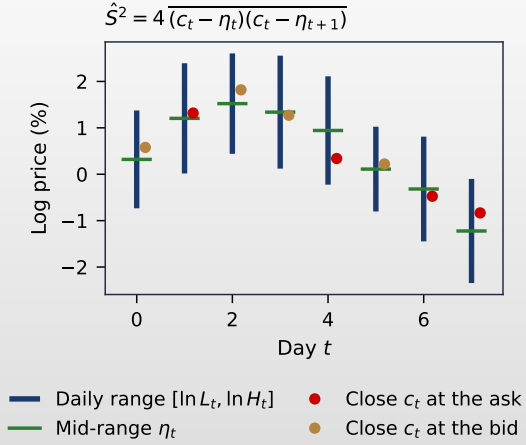
  - o estimare negativă pe două bare se înlocuiește cu 0, apoi se face media pe toate perechile
  - exemplu: amplitudinile zilnice 2%, amplitudinea pe două zile 2,5%: $\beta = 0{,}0008$, $\gamma = 0{,}000625$, $\alpha = 0{,}0079$, $\hat{S} \approx 79$ pb; fără spread, amplitudinea pe două zile ar fi $\sqrt{2} \times 2\% = 2{,}83\%$

Maxime, minime și spread-ul pe o traiectorie simulată



- Două zile de tranzacționare simulate: log-prețul eficient m_t, cotațiile $m_t \pm c$ cu $c = 40$ pb, tranzacții la bid sau la ask
- Fiecare maxim (triunghi în sus) este o cumpărare la ask, fiecare minim o vînzare la bid: orice amplitudine $\ln(H/L)$ conține o dată întregul spread
- Amplitudinea pe două zile este mai mare decît cele zilnice doar prin partea de volatilitate: compararea lui γ cu β izolează spread-ul

Spread-ul din închideri și amplitudini: Abdi–Ranaldo



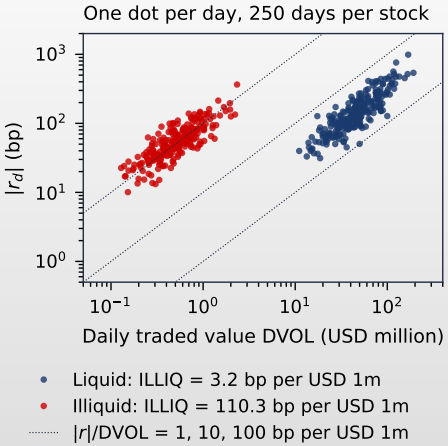
- Abdi–Ranaldo Abdi & Ranaldo (2017): log-închiderea c_t este la bid sau la ask; mijlocul amplitudinii aproximează log-prețul de mijloc

$$\eta_t = \frac{\ln H_t + \ln L_t}{2}$$

$$\hat{S}^2 = 4\,\overline{(c_t - \eta_t)(c_t - \eta_{t+1})}, \qquad \hat{S} = \sqrt{\max(\hat{S}^2, 0)}$$

  - bara de deasupra: media pe zile; $c_t - \eta_t$: închiderea față de mijlocul amplitudinii de azi, $c_t - \eta_{t+1}$: față de cel de mîine
  - știrile mișcă η_{t+1} independent de alternanța din c_t: în medie rămîne doar termenul de spread $S^2/4$
  - exemplu: un produs mediu de 0,002 · 0,001: $\hat{S}^2 = 8 \cdot 10^{-6}$, $\hat{S} = 28$ pb
- Ambii estimatori cu amplitudini amestecă spread-ul și volatilitatea: cu cît volatilitatea este mai mare față de spread, cu atît măsoară mai puțin spread-ul

Ilichiditatea: raportul Amihud



- **Raportul Amihud** Amihud (2002): randamentul absolut pe unitatea de valoare tranzacționată (DVOL, *Dollar Volume*), mediat pe D zile

$$\mathrm{ILLIQ} = \frac{1}{D}\sum_{d=\mathbf{1}}^{D}\frac{|r_d|}{\mathrm{DVOL}_d}, \qquad \mathrm{DVOL}_d = P_d \times V_d$$

  - r_d: randamentul logaritmic zilnic al prețului ajustat; P_d: prețul; V_d: numărul de acțiuni tranzacționate; DVOL în USD
  - unități: pb la 1 milion USD; măsoară **impactul asupra prețului**, nu spread-ul
  - exemplu: $|r_d| = 100$ pb la 2 milioane USD dă 50 pb la 1 milion USD
- Graficul: acțiunea ilichidă se mișcă la fel de mult la 0,5 milioane USD ca acțiunea lichidă la 50 de milioane USD: $\mathrm{ILLIQ} = 110{,}3$ față de 3,2
- Numitorul trebuie să fie o valoare, nu un număr de acțiuni (altfel raportul depinde de nivelul prețului)

Reeșantionare: bootstrap pe zile, pe blocuri, Monte Carlo

- **Bootstrap-ul pe zile** Efron (1979): extrageți D zile întregi cu întoarcere, recalculați statistica $\hat{\theta}_b^*$, repetați de B ori; intervalul de încredere (CI, *Confidence Interval*) de 95%:

$$\text{CI}_{\mathbf{95\%}} = \left[q^*_{\mathbf{2,5\%}},\ q^*_{\mathbf{97,5\%}} \right]$$

  - q_p^*: cuantila p a valorilor $\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*$; zilele întregi păstrează dependența dintre barele unei zile; zilele sînt tratate ca independente
- **Bootstrap-ul pe blocuri suprapuse** Künsch (1989) pentru seriile zilnice: reeșantionați blocuri de L zile consecutive (de exemplu $L = 20$) și lipiți-le; păstrează dependența mai scurtă decît L
- **Referința Monte Carlo**: simulați modelul (de exemplu Roll cu c cunoscut) de multe ori, calculați statistica pe fiecare eșantion
  - cuantilele de 2,5% și 97,5% ale valorilor simulate: cum arată statistica dacă modelul este adevărat
- **Corelația Spearman**: corelația Pearson a rangurilor $R(x_i)$, $R(y_i)$ (1 = cea mai mică valoare)

$$r_S = \text{Corr}\left(R(x_i),\ R(y_i)\right) = 1 - \frac{6\sum_i d_i^{\mathbf{2}}}{n(n^{\mathbf{2}}-1)}, \qquad d_i = R(x_i) - R(y_i)$$

  - a doua formă este valabilă fără valori egale; $r_S = 1$: aceeași ordine; robustă la valori extreme și la transformări monotone (de exemplu logaritmi)
- Volatilitatea anualizată: abaterea standard a randamentelor logaritmice zilnice înmulțită cu $\sqrt{A}$, $A = 252$ (acțiuni) sau 365 (cripto)

Regresia în logaritmi: elasticitatea și R^2

- **Regresia log-log** estimată prin metoda celor mai mici pătrate (OLS, *Ordinary Least Squares*) pe n zile

$$\ln y_d = a + b \ln x_d + e_d, \qquad \hat{b} = \frac{\sum_d (\ln x_d - \overline{\ln x})(\ln y_d - \overline{\ln y})}{\sum_d (\ln x_d - \overline{\ln x})^2}$$

  - b: **elasticitatea**: un x mai mare cu 1% se asociază cu un y mai mare cu b%; e_d: eroarea
  - exemplu: $b = 0{,}9$ și x crește cu 10%: y crește cu circa 9%
- **Coeficientul de determinare**: ponderea din dispersia lui $\ln y$ explicată de regresie

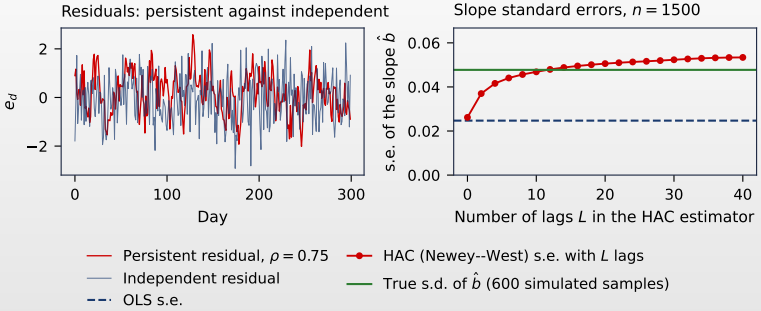
$$R^2 = 1 - \frac{\sum_d \hat{e}_d^2}{\sum_d \left(\ln y_d - \overline{\ln y}\right)^2} \in [0, 1]$$

- Eroarea standard OLS obișnuită presupune erori independente, cu dispersie constantă

$$\mathrm{SE}_{\mathrm{OLS}}(\hat{b}) = \frac{\hat{\sigma}_e}{\sqrt{\sum_d (\ln x_d - \overline{\ln x})^2}}, \qquad \hat{\sigma}_e^2 = \frac{1}{n-2} \sum_d \hat{e}_d^2$$

  - verificați-o cu funcția de autocorelație (ACF) a reziduurilor $\hat{e}_d$: o scădere lentă înseamnă erori persistente și o SE OLS prea mică

Erori HAC: dependența schimbă incertitudinea

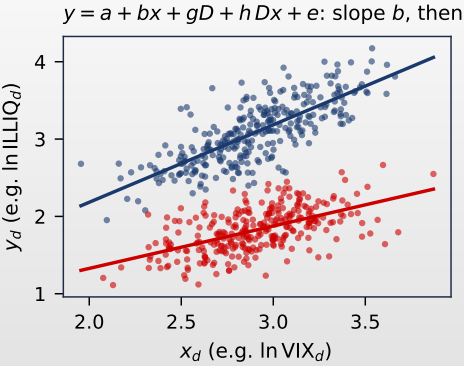


- Erorile standard **HAC** (*Heteroskedasticity and Autocorrelation Consistent*), robuste la heteroscedasticitate și autocorelație Newey & West (1987): cu $g_d = x_d\,\hat{e}_d$ (x_d: regresorii zilei d), estimatorul de tip sandwich

$$\hat{V} = (X'X)^{-1}\,\hat{S}\,(X'X)^{-1}, \qquad \hat{S} = \sum_d g_d g_d' + \sum_{\ell=1}^{L} \left(1 - \frac{\ell}{L+1}\right) \sum_d \left(g_d g_{d-\ell}' + g_{d-\ell} g_d'\right)$$

- L: numărul de laguri; $1 - \ell/(L+1)$: ponderile Bartlett; graficul: reziduuri ($\rho = 0{,}75$) și regresor persistente: SE OLS 0,025, HAC ($L = 20$) 0,051, abaterea standard adevărată 0,048

O ruptură a pantei și testul Wald



- **Variabila dummy** $D_d = 1$ de la o dată fixată dinainte, 0 înainte

$$y_d = a + b\,x_d + g\,D_d + h\,D_d x_d + e_d$$

  - panta b înainte de ruptură, $b + h$ după; g: deplasarea interceptului
- **Testul Wald** pentru H_0: $h = 0$ (panta nu se schimbă)

$$W = \left(\frac{\hat{h}}{\mathrm{SE}(\hat{h})}\right)^2 \sim \chi^2_1 \text{ sub } H_0$$

  - respingem la 5% dacă $W > 3{,}84$; $p = P(\chi^2_1 > W)$; folosiți SE HAC cînd reziduurile sînt persistente
  - nerespingerea nu dovedește stabilitatea: poate însemna doar o putere mică
- O **tendință** liniară t (ani de la prima zi), ca regresor suplimentar, preia o evoluție lentă a lui y

Pachetele de extindere: trei instrumente

- Pachetele de extindere (după curs) folosesc trei instrumente; modelele (Glosten–Milgrom, Kyle, Almgren–Chriss și probabilitatea tranzacționării informate, PIN, *probability of informed trading*) sînt prezentate în Capitolul 10
- **Regula lui Bayes**: actualizarea probabilității unei valori V_H după observarea unei cumpărări

$$P(V_H \mid \text{cump.}) = \frac{P(\text{cump.} \mid V_H)\, P(V_H)}{P(\text{cump.} \mid V_H)\, P(V_H) + P(\text{cump.} \mid V_L)\, P(V_L)}$$

- **Teorema proiecției**: pentru (v, y) cu distribuție Normală bivariată, cea mai bună prognoză a lui v dat y este liniară

$$\mathbb{E}[v \mid y] = \mathbb{E}[v] + \frac{\text{Cov}(v,y)}{\text{Var}(y)}\,(y - \mathbb{E}[y])\,, \qquad \text{Var}(v \mid y) = \text{Var}(v) - \frac{\text{Cov}(v,y)^2}{\text{Var}(y)}$$

- **Verosimilitatea Poisson** a unui număr k cu media λ și log-sum-exp pentru amestecuri

$$\ln P(k) = k \ln \lambda - \lambda - \ln k!, \qquad \ln \sum_j e^{z_j} = z_{\max} + \ln \sum_j e^{z_j - z_{\max}}$$

  - scăderea lui $z_{\max}$ evită depășirea numerică atunci cînd valorile z_j sînt mari (mii de tranzacții pe zi)

Harta exercițiilor

Exercițiul	Întrebarea	Tipul	Exemplu de urmat
A1	ce efect are oscilația bid-ask asupra variațiilor de preț?	Rezolvat	–
A2	cît de zgomotoasă este o estimare Roll?	Propus	A1
B1	este forma de U intrazilnică sistematică?	Rezolvat	–
B2	ce frecvență dă un spread credibil?	Rezolvat	–
B3	par acțiunile BVB la fel de lichide ca cele americane?	Propus	B2
B4	cît de strîns urmează ilichiditatea indicele VIX?	Rezolvat	–
C1	ce este greșit într-un răspuns de inteligență artificială (AI)?	Propus	A1, B2
Pachete de extindere	A3–A10 (modele informaționale, execuție, probabilitatea tranzacționării informate (PIN)), B5–B8, C2, C3 (opționale, după seminar)	ambele	indicat la fiecare cerință

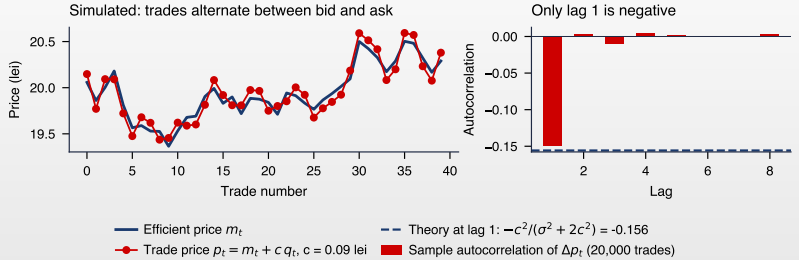
- **[Rezolvat]**: rezolvarea completă în slide-uri și în notebook, un model de urmat; **[Propus]**: îl rezolvați dumneavoastră, iar rezolvarea se discută la seminar
- Partea A: derivări pe hîrtie; Partea B: estimare și inferență pe date reale, fiecare cu o întrebare de interpretare; Partea C: întrebări deschise și critica unui răspuns AI (*Artificial Intelligence*, inteligență artificială)

A1 [Rezolvat]: Modelul Roll — cerință

- **Întrebările**
  - ce efect are oscilația bid-ask asupra momentelor variațiilor de preț?
  - cum recuperăm spread-ul din ele?
- **Date și ipoteze**
  - Modelul Roll din slide-ul „Autocovariațele și modelul Roll" Roll (1984): $p_t = m_t + c\, q_t$, cu m_t un mers aleator cu știri ε_t i.i.d. (media 0, dispersia σ^2)
  - Semnele $q_t = \pm 1$, i.i.d. cu probabilitatea 1/2, independente de ε
  - Cifrele: o acțiune BVB de circa 20 de lei, zilnic $\mathrm{Var}(\Delta p) = 0{,}0520$, $\mathrm{Cov}(\Delta p_t, \Delta p_{t-1}) = -0{,}0081$ (lei^2)
- **Cerințe**
  1. Scrieți Δp_t în funcție de ε_t, q_t și q_{t-1}, apoi deduceți $\mathrm{Var}(\Delta p_t)$ și $\mathrm{Cov}(\Delta p_t, \Delta p_{t-1})$.
  2. Din cele două cifre, calculați c, s în lei și în % din preț, σ^2, ponderea oscilației bid-ask $2c^2/\mathrm{Var}(\Delta p)$ și autocorelația de ordinul 1 ρ_1.
  3. Extindere: semnele formează un lanț Markov staționar cu $\mathrm{Corr}(q_t, q_{t-k}) = \rho^k$, $\mathrm{Var}(q_t) = 1$, tot independent de ε; dezvoltați cei patru termeni de covariață și găsiți limita lui $2\sqrt{-\hat{\gamma}_1}$ pentru $\rho = 0{,}3$.
- **Raportați**: cele două momente și un tabel: c, s (lei), s (%), σ^2, ponderea oscilației bid-ask, ρ_1 și limita din extindere
- **Interpretare**: fluxul de ordine persistent face spread-ul Roll prea mare sau prea mic?

◂ Înapoi ▸ Lista exercițiilor

A1 [Rezolvat]: Oscilația bid-ask, înaintea calculelor



- Simulare cu cifrele din A1 ($c = 0{,}09$ lei, $\sigma^2 = 0{,}0358$), 20 000 de tranzacții: prețul tranzacției sare între $m_t + c$ și $m_t - c$ în jurul prețului eficient
- Autocorelația de ordinul 1 de selecție a lui Δp_t: -0,150, apropiată de valoarea teoretică dedusă în rezolvare; nicio autocorelație la laguri mai mari; notebook: secțiunea A1

MFM_ch10_seminar

A1 [Rezolvat]: Momente și cifre

- Pasul 1. $\Delta p_t = \varepsilon_t + c(q_t - q_{t-1})$: inovația eficientă plus schimbarea părții (bid sau ask) la care se face tranzacția
- Pasul 2. Independența: $\mathrm{Var}(\Delta p_t) = \sigma^2 + c^2(1+1) = \sigma^2 + 2c^2$
- Pasul 3. $\mathrm{Cov}(\Delta p_t, \Delta p_{t-1}) = c^2\mathrm{Cov}(q_t - q_{t-1}, q_{t-1} - q_{t-2})$; singurul termen comun este q_{t-1}, cu semnul minus: $-c^2$
- Pasul 4. $c = \sqrt{0{,}0081}$; $\sigma^2 = 0{,}0520 - 2c^2$; $\rho_1 = -0{,}0081/0{,}0520$
- Interpretare: circa o treime din dispersia zilnică este zgomot bid-ask, deci volatilitatea din prețurile tranzacțiilor supraestimează volatilitatea eficientă

Mărimea	Valoarea
c	0,090 lei
$s = 2c$	0,18 lei
$s/20$	0,9%
$2c^2$	0,0162
σ^2	0,0358
ponderea oscilației bid-ask	31,2%
ρ_1	-0,156

A1 [Rezolvat]: Semne corelate

- Pasul 5. Cu $\text{Cov}(q_t, q_{t-k}) = \rho^k$, dezvoltăm cei patru termeni ai $\text{Cov}(q_t - q_{t-1}, q_{t-1} - q_{t-2})$
- $\text{Cov}(q_t, q_{t-1}) - \text{Cov}(q_t, q_{t-2}) - \text{Var}(q_{t-1}) + \text{Cov}(q_{t-1}, q_{t-2}) = \rho - \rho^2 - 1 + \rho = -(1-\rho)^2$
- Pasul 6. $\text{Cov}(\Delta p_t, \Delta p_{t-1}) = -c^2(1-\rho)^2 = -0{,}0081 \cdot 0{,}49 = -0{,}0040$
- Pasul 7. Estimatorul jumătății de spread converge la $c(1-\rho) = 0{,}063$ lei în loc de 0,090; cel al spread-ului la 0,126 în loc de 0,18 lei: ambele subestimate cu 30%
- Interpretare: fluxul persistent (ordine împărțite în mai multe tranzacții) face inversările mai rare, ascunde o parte din oscilația bid-ask și deplasează spread-ul Roll **în jos**

A2 [Propus]: Precizia estimatorului Roll — cerință

- **Exemplu de urmat**: A1
- **Întrebarea**: cît contează mărimea eșantionului T pentru o singură estimare Roll, cu momentele populației fixate?
- **Date și ipoteze**
  - Momentele populației din A1: $\gamma_0 = 0{,}0520$, $\gamma_1 = -0{,}0081$, deci $c = 0{,}090$, $s = 0{,}18$ lei; estimatorul Roll $\hat{s}$ (slide-ul „Autocovariațele și modelul Roll") Roll (1984); precizia lui: Harris (1990)
  - T = numărul variațiilor de preț din $T+1$ prețuri; dispersia lui $\hat{\gamma}_1$ pentru semne binare și metoda delta: slide-ul „Precizia estimatorului Roll"; Φ: funcția de repartiție a distribuției Normale standard
- **Cerințe**
  1. Deduceți $g'(\gamma_1)$ pentru $g(\gamma_1) = 2\sqrt{-\gamma_1}$ și scrieți $\mathrm{SE}(\hat{s})$ prin metoda delta.
  2. Completați tabelul de mai jos pentru $T = 78$ și $T = 250$ de variații de preț.
  3. Aproximați $P(\hat{\gamma}_1 > 0)$ cu distribuția Normală centrată în covarianța populației, folosind eroarea standard de la pasul 2; explicați sensul efectului centrării pe media eșantionului.
- **Raportați**: formula $\mathrm{SE}(\hat{s})$, tabelul completat și o frază despre sensul efectului centrării
- **Interpretare**: dacă modelul Roll este adevărat, cît de des am observa o covarianță pozitivă într-un eșantion de 78 de variații?

T	$\mathrm{SE}(\hat{\gamma}_1)$	$\mathrm{SE}(\hat{s})$ (lei)	$\mathrm{SE}(\hat{s})/s$	$P(\hat{\gamma}_1 > 0)$
78	...	...	...	...
250	...	...	...	...

◂ Înapoi ▸ Lista exercițiilor

B1 [Rezolvat]: Forma de U intrazilnică pentru SPY — cerință

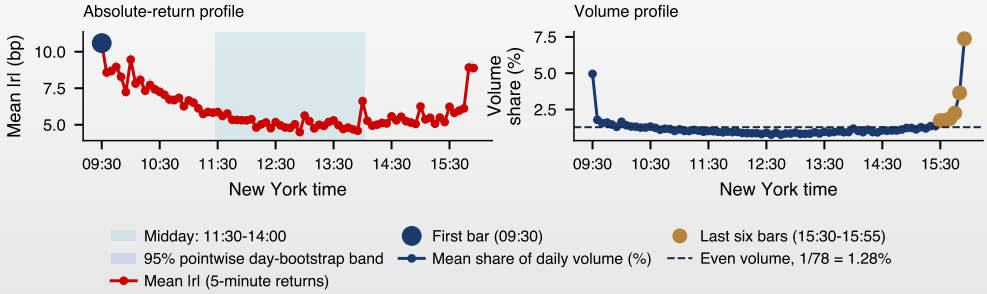
- **Întrebările**
  - sînt volatilitatea și volumul sistematic mai mari la deschidere și la închidere?
  - este tiparul determinat de cîteva zile extreme?
- **Date și ipoteze**
  - Barele SPY de 5 minute, 1 477 de zile complete (slide-ul „Datele folosite"); $|r|$ în pb; barele sînt etichetate cu ora de început
  - Un rînd pe zi: $|r|$ al primei bare (09:30), media $|r|$ a celor 31 de bare de la prînz care încep între 11:30 și 14:00 inclusiv și ponderea volumului zilei în ultimele șase bare
- **Cerințe**
  1. Calculați raportul mediilor pe zile, $\overline{|r|}_{\textbf{deschidere}}/\overline{|r|}_{\textbf{prînz}}$, și ponderea medie a volumului din ultimele 30 de minute.
  2. Construiți intervale bootstrap pe zile de 95% pentru ambele statistici (1 000 de reeșantionări ale rîndurilor complete, seed 42).
  3. Reprezentați media $|r|$ și ponderea medie a volumului pe bare, cu benzi bootstrap punctuale pe zile.
- **Raportați**: cele două estimări cu intervale și cele două profile
- **Interpretare**
  - de ce reeșantionăm zile, nu bare?
  - ce implică tiparul pentru un investitor care trebuie să cumpere azi 1 milion de acțiuni?

◂ Înapoi ▸ Lista exercițiilor

B1 [Rezolvat]: Reeșantionarea zilelor, pas cu pas

- Pasul 1. Estimările punctuale: $\overline{|r|}_{\text{deschidere}} = 10{,}6$ pb, $\overline{|r|}_{\text{prînz}} = 5{,}1$ pb, raportul 2,06; ultimele 30 de minute 18,6% din volumul zilnic, față de $6/78 = 7{,}7\%$ la o repartizare uniformă
- Pasul 2. O reeșantionare: extragem 1 477 de rînduri-zi cu întoarcere (aici apar 933 de zile distincte), păstrăm fiecare rînd întreg, recalculăm
  - Prima reeșantionare: 10,7 pb și 5,0 pb, raportul 2,13
- Pasul 3. Repetăm de 1 000 de ori; percentilele 2,5% și 97,5%: raportul $[1{,}98; 2{,}15]$, ponderea volumului $[18{,}3; 19{,}0]\%$
- Motivul reeșantionării zilelor întregi: barele unei zile au în comun volatilitatea și știrile ei, deci reeșantionarea zilelor păstrează această dependență; zilele sînt totuși tratate ca independente, deși volatilitatea persistă

B1 [Rezolvat]: Profilele volatilității și volumului



- Ambele intervale sînt înguste și exclud reperele uniforme (raportul 1, 7,7%): un tipar sistematic, nu cîteva zile extreme; benzile sînt punctuale, nu o bandă simultană pentru toată curba
- Interpretare: un cumpărător mare găsește cel mai mult volum la deschidere și la închidere; deschiderea este însă și cea mai volatilă, deci momentul optim depinde și de spread-uri, impact, urgență și riscul așteptării

MFM_ch10_seminar

B2 [Rezolvat]: Estimatori de spread la două frecvențe — cerință (1/2)

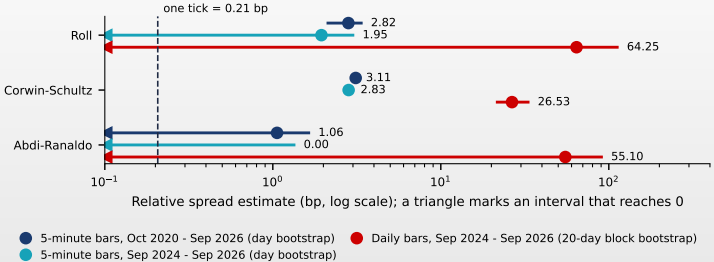
- **Întrebarea**: ce frecvență de eșantionare conduce la un spread bid-ask credibil pentru SPY, cel mai tranzacționat ETF?
- **Date și ipoteze**
  - Intrazilnic: cele 1 477 de zile de bare SPY de 5 minute; pe zi, 78 de închideri dau 77 de randamente logaritmice
  - Zilnic: bare SPY zilnice, 19 sep. 2024 – 18 sep. 2026 (501 zile); un pas = 0,01 USD. Roll zilnic folosește o singură autocovarianță de ordinul 1, centrată pe media întregii ferestre
- Rețetele (prețuri în logaritm; toate spread-urile relative)
  - Roll: pe zi, autocovarianța de ordinul 1 centrată pe medie a celor 77 de randamente, mediată pe zile, apoi $2\sqrt{-\bar{\gamma}_1}$, pusă la 0 dacă $\bar{\gamma}_1 \geq 0$ Roll (1984)
  - Corwin–Schultz: estimarea pe două bare pentru fiecare pereche de bare consecutive (negativele puse la 0), mediată; pentru barele zilnice se adaugă ajustarea pentru noapte Corwin & Schultz (2012)
  - Abdi–Ranaldo: media lui $4(c_t - \eta_t)(c_t - \eta_{t+1})$, apoi rădăcina pătrată, 0 dacă este negativă Abdi & Ranaldo (2017)

◂ Înapoi ▸ Lista exercițiilor

B2 [Rezolvat]: Estimatori de spread la două frecvențe — cerință (2/2)

- **Cerințe**
  1. Calculați cei trei estimatori din barele de 5 minute, cu intervale bootstrap pe zile de 95% (reeșantionați rîndurile zilnice de momente deja calculate, nu prețurile).
  2. Repetați pe aceeași fereastră de doi ani ca barele zilnice, astfel încît să difere doar frecvența.
  3. Calculați cei trei estimatori din barele zilnice cu bootstrap pe blocuri suprapuse (blocuri de 20 de zile, 1 000 de reeșantionări) și raportați cît de des este trunchiat fiecare la 0.
  4. Comparați fiecare estimare cu un pas raportat la preț, o limită inferioară pentru spread-ul cotat (nu pentru cel efectiv).
  5. Simulați modelul Roll cu dispersia barelor SPY, 78 de prețuri pe zi, 1 477 de zile pe eșantion, 200 de eșantioane, pentru un spread de un tick ($c =$ jumătate de tick) și $c = 0$; calculați ponderea zilelor cu $\hat{\gamma}_1 > 0$ și estimarea agregată Harris (1990).
  6. Corectați estimarea Roll agregată pe 5 minute pentru deplasarea din centrarea pe medie din slide-ul „Precizia estimatorului Roll", cu $T = 77$.
- **Raportați**: un grafic cu estimările și intervalele, ponderile trunchierii și rezultatele Monte Carlo lîngă cele observate
- **Interpretare**
  - ce frecvență dă un spread credibil?
  - ce măsoară estimatorii zilnici?
  - este ponderea covarianțelor pozitive o dovadă împotriva modelului Roll?

B2 [Rezolvat]: Estimări și intervale



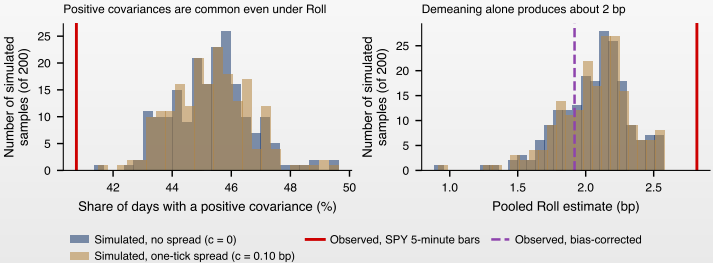
- 5 minute (2020–2026): Roll 2,82 [2,09; 3,43], Corwin–Schultz 3,11 [3,03; 3,20], Abdi–Ranaldo 1,06 [0,00; 1,67] pb; aceiași doi ani (496 de zile): 1,95; 2,83; 0,00 pb
- Zilnic: Roll 64,3 [0,0; 114,4], Corwin–Schultz 26,5 [21,3; 33,7], Abdi–Ranaldo 55,1 [0,0; 92,2] pb; trunchiat la 0 în 9,7% (Roll) și 3,4% (Abdi–Ranaldo) din reeșantionări; Abdi–Ranaldo pe 5 minute în 11,2%
- Un pas înseamnă 0,21 pb: estimările zilnice sînt cu două ordine de mărime peste el, fiindcă abaterea standard zilnică este 103,4 pb; chiar și estimarea Corwin–Schultz pe 5 minute reprezintă 15 pași

MFM_ch10_seminar

Un spread de un pas și semnul autocovarianței

- **Ce credeți**: dacă spread-ul SPY este exact un pas, în ce pondere din zile este pozitivă autocovarianța de ordinul 1 a randamentelor de 5 minute, deși modelul Roll prezice $\gamma_1 = -c^2 < 0$?
- **Răspuns**: aproape jumătate: 45,3% din zile în simulare, la fel ca fără spread ($c = 0$)
- Semnalul este mic: $c^2 \approx 0{,}01$ pb^2, față de dispersia pe bară $9{,}2^2 \approx 85$ pb^2
- Abaterea standard a lui $\hat{\gamma}_1$ pe o zi cu 77 de randamente este de circa $85/\sqrt{77} \approx 9{,}6$ pb^2: semnul zilnic este aproape aleator
- Centrarea pe medie deplasează $\hat{\gamma}_1$ cu circa $-\gamma_0/77 \approx -1{,}1$ pb^2, de aceea ponderea rămîne sub 50%: $\Phi(-1{,}1/9{,}6) \approx 0{,}45$ (Φ: funcția de repartiție a distribuției Normale standard)

B2 [Rezolvat]: Monte Carlo sub modelul Roll



- Un pas: $c = 0{,}10$ pb, față de abaterea standard pe bară de 9,2 pb; ponderea zilelor pozitive 45,3% [43,0; 48,1], cu $c = 0$ 45,3%: nu se deosebesc; observat 40,8%, sub ambele intervale de referință simulate
- Roll agregat: 2,07 pb (un pas), 2,06 pb ($c = 0$), observat 2,82 pb; corecția (formula deplasării rezolvată în raport cu γ_1, cu $T = 77$ și mediile pe zile $\bar{\gamma}_0$, $\bar{\gamma}_1$) $\hat{\gamma}_1^c = (\bar{\gamma}_1 + \bar{\gamma}_0/77)/(1 - 2/77)$ dă 1,92 pb
- Interpretare: doar barele intrazilnice dau ordine de mărime credibile; covariațele pozitive sînt așteptate sub Roll, deci nu îl resping; după corecția aproximativă a centrării rămîne o autocovariață negativă neexplicată de reperul Roll cu spread de un pas; corecția nu separă oscilația bid-ask, heteroscedasticitatea, dependența semnelor sau discretizarea. Acestea sînt intervale de referință simulate, nu CI bootstrap

MFM_ch10_seminar

B3 [Propus]: Estimări ale spread-ului pentru 25 de active — cerință

- **Exemplu de urmat**: B2
- **Întrebările**
  - plasează aproximările zilnice ale spread-ului acțiunile BVB la același nivel de lichiditate cu cele americane?
  - confirmă raportul Amihud acest lucru?
- **Date și ipoteze**
  - Cele 25 de active de pe slide-ul Datele folosite: 7 SUA, 10 acțiuni BVB plus ETF-ul pe BET (11), 7 cripto; bare zilnice, 19 sep. 2024 – 18 sep. 2026
  - Estimatorii ca în B2 (rețetele zilnice); volatilitatea anualizată cu 252 de zile (acțiuni) sau 365 (cripto); Amihud în pb la 1 milion USD, volumul cripto deja în USD, BVB convertit la cursul BNR
- **Cerințe**
  1. Calculați Corwin–Schultz și Abdi–Ranaldo pentru fiecare activ și mediana fiecărui estimator pe grupuri.
  2. Numărați, pe grupuri, activele pentru care Abdi–Ranaldo este pus la 0.
  3. Calculați corelațiile Spearman ale Corwin–Schultz cu volatilitatea, ale Amihud cu valoarea tranzacționată mediană și ale Corwin–Schultz cu Amihud și desenați cele trei grafice de dispersie.
- **Raportați**: un tabel cu rîndurile SUA, BVB, cripto și coloanele mediana Corwin–Schultz, mediana Abdi–Ranaldo, zerourile Abdi–Ranaldo; cele trei corelații și grafice
- **Interpretare**
  - în ce indicator ați avea încredere pentru a compara piețele?
  - ce date v-ar trebui pentru validare?

◂ Înapoi ▸ Lista exercițiilor

B4 [Rezolvat]: Ilichiditatea și VIX — cerință (1/2)

Întrebările

- cu cît crește ilichiditatea intrazilnică a SPY cînd crește volatilitatea implicită?
- cît de precisă este elasticitatea?

Date și ipoteze

- $\mathrm{ILLIQ}_d = \frac{1}{K_d} \sum_k 10^4 |r_{d,k}| / (\mathrm{DVOL}_{d,k}/10^8)$: pb la 100 de milioane USD, din barele de 5 minute ale zilei d, $k = 1, \ldots, K_d$ ($r_{d,k}$: randamentul logaritmic, $\mathrm{DVOL}_{d,k}$ = închiderea $\times$ volumul; 10^4 convertește în pb, 10^8 în 100 de milioane USD); K_d: numărul barelor cu valoare nenulă, de regulă 78
- Seriile sînt aliniate pe zilele de tranzacționare cu închiderea indicelui VIX: 1 477 de zile, 2020-10-12 – 2026-09-18; media ILLIQ 1,83

B4 [Rezolvat]: Ilichiditatea și VIX — cerință (2/2)

Cerințe

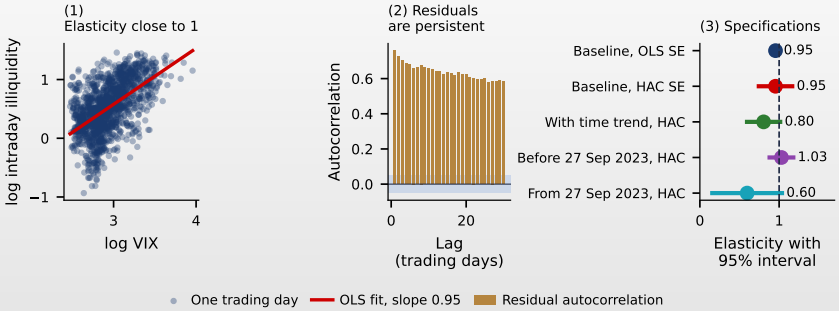
1. Estimați prin OLS $\ln \mathrm{ILLIQ}_d = a + b \ln \mathrm{VIX}_d + e_d$ (regresia log-log din slide-ul „Regresia în logaritmi: elasticitatea și R^2") și desenați dreapta.
2. Desenați autocorelațiile reziduurilor; raportați erorile standard OLS și HAC (slide-ul „Erori HAC: dependența schimbă incertitudinea", 20 de laguri) și intervalul HAC de 95% Newey & West (1987).
3. Extindere: adăugați o tendință liniară t (în ani de la prima zi) și reestimați b.
4. Extindere: cu $D_d = 1$ de la 2023-09-27 (mijlocul eșantionului, fixat înainte de estimare), adăugați $gD_d + hD_d \ln \mathrm{VIX}_d$ și testați $h = 0$ cu testul Wald HAC din slide-ul „O ruptură a pantei și testul Wald" la 5%.

Raportați: un tabel al specificațiilor ($\hat{b}$, SE, interval) și graficul coeficienților

Interpretare

- de ce diferă cele două erori standard?
- este elasticitatea stabilă?
- este relația cauzală?

B4 [Rezolvat]: Estimarea și reziduurile



- (1) $\hat{b} = 0{,}95$, $R^2 = 0{,}30$: un VIX mai mare cu 10% se asociază cu o ilichiditate mai mare cu circa $0{,}95 \times 10\%$; (2) autocorelația reziduurilor la lagul 1 este 0,76 și scade lent: SE OLS 0,038, SE HAC 0,121 (de 3,2 ori mai mare), CI HAC $[0{,}72; 1{,}19]$

MFM_ch10_seminar

B4 [Rezolvat]: Specificațiile și testul de ruptură

Specificația	$\hat{b}$	SE	interval 95%
De bază, SE OLS	0,95	0,038	[0,88; 1,03]
De bază, SE HAC	0,95	0,121	[0,72; 1,19]
Cu tendință (HAC)	0,80	0,120	[0,57; 1,04]
Înainte de 2023-09-27 (HAC)	1,03	0,090	[0,85; 1,21]
De la 2023-09-27 (HAC)	0,60	0,238	[0,13; 1,06]

- Tendința -0,062 pe an (SE 0,021): ilichiditatea scade în timp, legătura cu VIX rămîne; ruptura: $\hat{h} = -0{,}43$, Wald $(\hat{h}/\mathrm{SE})^2 = (-0{,}433/0{,}254)^2 = 2{,}90$, $p = 0{,}089$: nerespins la 5%
- Interpretare: reziduurile persistente fac erorile OLS mult prea mici; elasticitatea este aproape de 1; nerespingerea la 5% nu dovedește stabilitatea; relația nu este cauzală, fiindcă ambele reacționează la știri

C1 [Propus]: Critica unui răspuns AI — răspunsul

- **Promptul**: „Am închiderile zilnice `P` (lei) și volumele `V` (număr de acțiuni, pozitive în fiecare zi) pentru o acțiune BVB. Estimează spread-ul Roll și ilichiditatea Amihud. Cît mă costă o tranzacție dus-întors?”
- **Răspunsul, text**: Roll: $\mathrm{Cov}(\Delta p_t, \Delta p_{t-1}) = -c^2$ și $s = 2c$; dacă covarianța iese pozitivă, luați valoarea ei absolută, formula dă în continuare spread-ul; Amihud: media lui $|r_d|$ împărțit la volumul zilei d, la 1 milion; rezultat: $s = 0{,}05$ lei la o acțiune de 20 de lei, 0,25% din preț, deci o tranzacție dus-întors costă $2s$, 0,50%
- **Răspunsul, codul**

```
import numpy as np
dp = P.diff().dropna().values
g = np.cov(dp[1:], dp[:-1])[0, 1]
s_roll = 2 * np.sqrt(abs(g))
s_pct = 100 * s_roll / P.mean()
r = np.log(P).diff().dropna()
illiq = 1e6 * (r.abs() / V.loc[r.index]).mean()
```

◂ Înapoi ▸ Lista exercițiilor

C1 [Propus]: Critica unui răspuns AI — cerință

- **Exemplu de urmat**: A1, B2
- **Întrebările**
  - ce părți ale răspunsului AI sînt greșite?
  - cum ați detecta fiecare eroare înainte de a folosi cifrele?
- **Date și ipoteze**
  - Un audit conceptual: nu este nevoie de un fișier de date; `P` și `V` sînt ipotetice
  - Verificați fiecare formulă și fiecare unitate din răspuns cu notațiile introduse azi
- **Cerințe**
  1. Răspunsul conține trei erori; identificați-le în text și, unde este cazul, în cod.
  2. Pentru fiecare eroare, precizați cum ați detecta-o: un test, o cifră publicată sau o formulă din noțiunile introductive de azi.
  3. Corectați fiecare eroare: dați formula, linia de cod și cifra corectată acolo unde informația o permite; precizați unitățile raportului Amihud.
  4. Consemnați erorile găsite efectiv în `AI_ERRORS.md`: cererea, răspunsul, eroarea, cum ați găsit-o, corecția.
- **Raportați**: cele trei erori cu detectarea și corecția
- **Interpretare**: care eroare ar schimba o decizie de investiții?

Proiectul de echipă: alegeți unul

- C2: se reflectă ilichiditatea în prețurile de pe BVB? C3: costul implicit al competiției pentru viteză pentru SPY; sau: Roll, Corwin–Schultz și Abdi–Ranaldo pentru trei active nediscutate aici (cu datele folosite documentate)
- Livrabile: un Quantlet reproductibil și un raport de cel mult două pagini
- Criterii de evaluare: întrebarea, datele și convențiile, metoda, incertitudinea (intervale, robustețe), limitele
- Declarați folosirea AI în `AI_USE.md`; consemnați fiecare eroare AI găsită efectiv în `AI_ERRORS.md`

Pachetele de extindere: cunoștințe necesare

Pachetul	Exercițiile	Cunoștințe necesare
1. Modele informaționale	A3 [Rezolvat], A4, A5 [Rezolvat], A6	cursul, secțiunea *Informația și prețurile*; regula lui Bayes, proiecții
2. Execuția optimă	A7 [Rezolvat], A8	cursul, secțiunea *Impactul asupra prețului și execuția optimă*
3. Tranzacționarea informată (PIN)	A9 [Rezolvat], A10	verosimilitatea Poisson, log-sum-exp
4. Ilichiditatea între piețe	B5	B2, B3
5. Volum și impact	B6 [Rezolvat]	B1, estimarea cu variabile instrumentale (IV)
6. Ceasul Bitcoin	B7	B1
7. Descoperirea prețului	B8 [Rezolvat]	cointegrare, modelul vectorial cu corecția erorilor (VECM)
8. Proiecte	C2, C3	B4, B5; studiul de caz din curs

- Fiecare pachet are secțiunea lui în notebook; rezolvați-le în ordine în cadrul pachetului; C2 și C3 sînt posibile proiecte de echipă

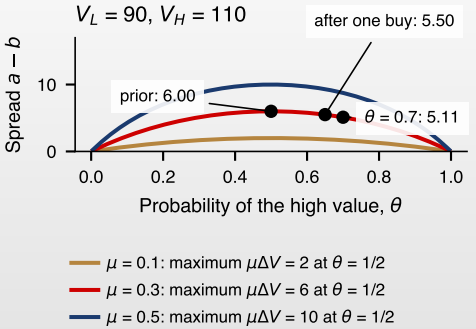
A3 [Rezolvat]: Glosten–Milgrom pentru orice probabilitate inițială — cerință

- **Întrebările**
  - cît de larg este spread-ul cînd formatorul de piață (*market maker*) învață din fluxul de ordine?
  - cînd este maxim?
- **Date și ipoteze**
  - $V \in \{V_L, V_H\}$, $\Delta V = V_H - V_L$; $\mathcal{F}_t$ = tranzacțiile observate pînă acum, $\theta = P(V = V_H \mid \mathcal{F}_t)$ Glosten & Milgrom (1985)
  - Verosimilitățile: $P(\text{cump.} \mid V_H) = (1+\mu)/2$, $P(\text{cump.} \mid V_L) = (1-\mu)/2$ (ponderea investitorilor informați μ; investitorii neinformați cumpără sau vînd cu probabilitatea 1/2)
  - Cotațiile înaintea următoarei tranzacții: $a = \mathbb{E}[V \mid \mathcal{F}_t, \text{cump.}]$, $b = \mathbb{E}[V \mid \mathcal{F}_t, \text{vînz.}]$; $V_L = 90$, $V_H = 110$, $\mu = 0{,}3$
- **Cerințe**
  1. La $\theta = 0{,}5$, calculați $P(V_H \mid \text{cump.})$, ask-ul, bid-ul și spread-ul.
  2. După o cumpărare, probabilitatea a posteriori devine noul θ: calculați cotațiile și spread-ul pentru tranzacția următoare.
  3. Deduceți spread-ul pentru orice $0 < \theta < 1$ și arătați că este maxim la $\theta = 1/2$; verificați-l la $\theta = 0{,}7$.
  4. Extindere: arătați că $\mathbb{E}[\text{prețul următoarei tranzacții} \mid \mathcal{F}_t] = \mathbb{E}[V \mid \mathcal{F}_t]$ și că probabilitățile formează un martingal.
- **Raportați**: cotațiile, formula spread-ului și argumentul de martingal
- **Interpretare**: de ce este spread-ul maxim cînd formatorul de piață este cel mai nesigur?

◂ Înapoi ▸ Lista exercițiilor

A3 [Rezolvat]: Cotațiile și spread-ul general

- Pasul 1 ($\theta = 0{,}5$).
  $P(V_H \mid \text{cump.}) = 0{,}65 \cdot 0{,}5/(0{,}65 \cdot 0{,}5 + 0{,}35 \cdot 0{,}5) = 0{,}6500$:
  $a = 90 + 20 \cdot 0{,}6500 = 103{,}00$, $b = 97{,}00$, spread $6{,}00 = \mu \Delta V$
- Pasul 2. După o cumpărare $\theta = 0{,}6500$: $a = 105{,}50$,
  $b = 100{,}00$, spread 5,50
- Pasul 3. Cu $x = 2\theta - 1$: $P(V_H \mid \text{cump.}) = \theta(1+\mu)/(1+\mu x)$,
  $P(V_H \mid \text{vînz.}) = \theta(1-\mu)/(1-\mu x)$
- Pasul 4. $a - b = \mu \Delta V (1 - x^2)/(1 - \mu^2 x^2) \leq \mu \Delta V$, cu egalitate doar la $x = 0$; la $\theta = 0{,}7$: 5,11
- Pasul 5. $P(\text{cump.})a + P(\text{vînz.})b = \mathbb{E}[V \mid \mathcal{F}_t]$ (așteptarea totală); așteptările iterate dau martingalul
- Interpretare: la $\theta = 1/2$ o tranzacție modifică cel mai mult probabilitățile estimate de formatorul de piață, deci selecția adversă costă cel mai mult



MFM_ch10_seminar

A4 [Propus]: Viteza de învățare a formatorului de piață — cerință

- **Exemplu de urmat**: A3
- **Întrebările**
  - cîte cumpărări consecutive sînt necesare pînă cînd bid-ul depășește 108?
  - cum depinde acest număr de μ?
- **Date și ipoteze**
  - A3 cu $\theta_{\mathbf{0}} = 1/2$, $V_L = 90$, $V_H = 110$; valoarea adevărată V_H; un **scenariu condiționat**, în care fiecare tranzacție este o cumpărare (nu se cere timpul așteptat de învățare)
  - Cronologia: după k cumpărări probabilitatea este θ_k; bid-ul cotat atunci se condiționează de o vînzare ipotetică următoare
- **Cerințe**
  1. Arătați că fiecare cumpărare înmulțește cota $\theta/(1-\theta)$ cu $(1+\mu)/(1-\mu)$, iar fiecare vînzare o împarte la același factor.
  2. Arătați că bid-ul după k cumpărări este $V_L + \Delta V\, \theta_{k-\mathbf{1}}$.
  3. Arătați că bid > 108 cere o cotă peste 9, deduceți $k_{\min} = \lfloor \ln 9 / \ln \frac{\mathbf{1}+\mu}{\mathbf{1}-\mu} \rfloor + 2$ ($\lfloor\cdot\rfloor$: partea întreagă) și evaluați-l pentru $\mu = 0{,}1; 0{,}3; 0{,}5$, verificînd cazul de la limită.
  4. Comparați scenariul condiționat cu traiectorii simulate în care investitorii neinformați pot și să vîndă.
- **Raportați**: cele trei valori ale lui $k_{\min}$ și graficul bid-ului în funcție de k
- **Interpretare**: cum depinde viteza descoperirii prețului de ponderea investitorilor informați?

◂ Înapoi ▸ Lista exercițiilor

A5 [Rezolvat]: Deducerea echilibrului Kyle — cerință

- **Întrebările**
  - cît de agresiv tranzacționează un investitor din interior?
  - cu cît mișcă formatorul de piață prețul pe acțiune?
- **Date și ipoteze**
  - Cronologia: (1) investitorul vede $v \sim N(p_0, \sigma_v^2)$ și trimite x; (2) investitorii zgomot (*noise traders*) trimit $u \sim N(0, \sigma_u^2)$, independent; (3) formatorul de piață vede doar $y = x + u$ și stabilește $p = \mathbb{E}[v \mid y]$ Kyle (1985)
  - Unități: v, p în USD; x, u, y în acțiuni; β în acțiuni per USD; λ în USD per acțiune. Cifre: $\sigma_v = 2$, $\sigma_u = 10\,000$ și $20\,000$
- **Cerințe**
  1. Răspunsul optim al investitorului: dat $p = p_0 + \lambda y$, maximizați profitul așteptat și scrieți β din $x = \beta(v - p_0)$.
  2. Prețul formatorului de piață: dat $x = \beta(v - p_0)$, deduceți λ din teorema proiecției.
  3. Rezolvați cele două ecuații; calculați profitul așteptat, $\mathrm{Var}(x)$ și $\mathrm{Var}(v \mid y)$.
  4. Completați un tabel cu două coloane ($\sigma_u = 10\,000$ și $20\,000$): β, λ, profitul, variația prețului pentru $y = 15\,000$.
- **Raportați**: formulele echilibrului și tabelul
- **Interpretare**: de ce îl ajută mai mult zgomot pe investitorul din interior?

◂ Înapoi ▸ Lista exercițiilor

A5 [Rezolvat]: Echilibrul și cifrele

- Pasul 1. $\mathbb{E}[(v - p_0 - \lambda x - \lambda u)x \mid v] = (v - p_0)x - \lambda x^2$; condiția de ordinul întîi: $x = (v - p_0)/(2\lambda)$, $\beta = 1/(2\lambda)$
- Pasul 2. (v, y) are distribuție Normală bivariată: $\lambda = \mathrm{Cov}(v, y)/\mathrm{Var}(y) = \beta\sigma_v^2/(\beta^2\sigma_v^2 + \sigma_u^2)$
- Pasul 3. Cu $2\lambda\beta = 1$: $\beta^2\sigma_v^2 = \sigma_u^2$, deci $\beta = \sigma_u/\sigma_v$, $\lambda = \sigma_v/(2\sigma_u)$
- Pasul 4. Profitul $\beta\sigma_v^2 - \lambda\beta^2\sigma_v^2 = \sigma_v\sigma_u/2$; $\mathrm{Var}(x) = \sigma_u^2$; $\mathrm{Var}(v \mid y) = \sigma_v^2/2$ în ambele economii
- Interpretare: λ măsoară conținutul informațional per acțiune; mai mult zgomot face fluxul mai puțin informativ, investitorul tranzacționează mai mult și cîștigă mai mult, iar jumătate din informație intră totuși în preț

	$\sigma_u = 10^4$	$\sigma_u = 2 \cdot 10^4$
β	5 000	10 000
λ	0,0001	0,00005
profit (USD)	10 000	20 000
Δp ($y = 15\,000$)	1,50	0,75

A6 [Propus]: Lambda lui Kyle ca pantă de regresie — cerință

- **Exemplu de urmat**: A5
- **Întrebările**
  - ce estimează un raport de tip Amihud într-o economie Kyle?
  - contează alegerea volumului de la numitor?
- **Date și ipoteze**
  - A5 cu $p_0 = 50$ USD, $\sigma_v = 2$, $\sigma_u = 10\,000$; simulare (seed 42): 200 000 de licitații independente, $v = p_0 + \sigma_v z_1$, $u = \sigma_u z_2$, $x = \beta(v - p_0)$, $y = x + u$, $p = p_0 + \lambda y$, cu z_1, z_2 extrageri independente din distribuția Normală standard
  - $|x| + |u|$ este o aproximare a volumului brut de ordine (cele două ordine numărate separat), nu volumul bursei
- **Cerințe**
  1. Arătați că pantele OLS în populație ale lui $p - p_0$ și ale lui $v - p_0$ pe y sînt ambele λ și dați $\mathrm{Var}(y)$.
  2. Arătați că $|r|/(p_0|y|)$, cu $r = (p - p_0)/p_0$, este λ/p_0^2 și convertiți-l în pb la 1 milion USD (factorul $10^4 \times 10^6$).
  3. Înlocuiți $|y|$ cu $|x| + |u|$, deduceți sensul deplasării și măsurați-o prin simulare; repetați cu $\sigma_u = 20\,000$.
- **Raportați**: pantele, raportul pentru ambele numitoare și ambele niveluri de zgomot
- **Interpretare**
  - ce măsoară raportul Amihud în modelul Kyle?
  - de ce este deplasat cu volumul brut?

Înapoi Lista exercițiilor

A7 [Rezolvat]: Execuția Almgren–Chriss — cerință

- **Întrebările**
  - cum ar trebui să împartă un vînzător cu aversiune la risc 1 milion de acțiuni pe cinci zile?
  - cît costă reducerea riscului?
- **Date și ipoteze**
  - Parametrii Almgren & Chriss (2001): $X = 10^6$ acțiuni, $T = 5$ zile, $N = 5$, $\tau = 1$; $\sigma = 0{,}95$ USD pe $\text{zi}^{1/2}$; impactul temporar $\eta = 2{,}5 \cdot 10^{-6}$, permanent $\gamma = 2{,}5 \cdot 10^{-7}$ (USD pe acțiune per acțiune); aversiunea la risc $\lambda = 2 \cdot 10^{-6}$ per USD
  - x_k = deținerea după tranzacția k ($x_0 = X$, $x_5 = 0$), $n_k = x_{k-1} - x_k$; $\mathbb{E}[C] = \frac{1}{2}\gamma X^2 + \frac{\tilde\eta}{\tau}\sum_{k=1}^{5} n_k^2$, $\operatorname{Var}[C] = \sigma^2 \tau \sum_{k=1}^{5} x_k^2$, $\tilde\eta = \eta - \gamma\tau/2$
- **Cerințe**
  1. Calculați $\tilde\eta$, $\tilde\kappa^2 = \lambda\sigma^2/\tilde\eta$ și κ din $\cosh(\kappa\tau) = 1 + \tilde\kappa^2\tau^2/2$.
  2. Completați un tabel cu k, $x_k = X\sinh(\kappa(5-k))/\sinh(5\kappa)$ și n_k.
  3. Calculați $\mathbb{E}[C]$ și $\operatorname{sd}[C]$ și comparați cu vînzarea uniformă ($\lambda = 0$).
  4. Extindere: minimizați $\int_0^T (\eta\dot{x}^2 + \lambda\sigma^2 x^2)\,dt$ cu $x(0) = X$, $x(T) = 0$ prin Euler–Lagrange și comparați κ_c cu κ.
- **Raportați**: tabelul programului, comparația cost–risc și graficul deținerilor
- **Interpretare**: ce obține investitorul în schimbul costului așteptat suplimentar?

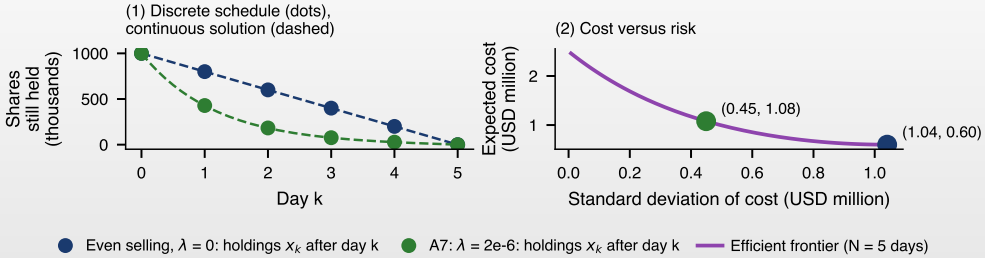
înapoi Lista exercițiilor

A7 [Rezolvat]: Programul, costul și riscul

- Pasul 1. $\tilde{\eta} = 2{,}5 \cdot 10^{-6} - 1{,}25 \cdot 10^{-7} = 2{,}375 \cdot 10^{-6}$; $\tilde{\kappa}^2 = 0{,}76$; $\cosh \kappa = 1{,}38$, $\kappa = 0{,}8463$, $\sinh(5\kappa) = 34{,}40$
- Pasul 2. Tabelul: 57% se vinde în prima zi
- Pasul 3. $\mathbb{E}[C] = 1{,}078$, $\mathrm{sd}[C] = 0{,}449$ milioane USD; uniform: 0,600 și 1,041
- Interpretare: vînzarea concentrată la început costă suplimentar 0,48 milioane USD impact așteptat pentru a reduce abaterea standard a costului la mai puțin de jumătate

k	x_k	n_k
0	1 000 000	–
1	428 599	571 401
2	182 933	245 666
3	76 296	106 637
4	27 643	48 652
5	0	27 643

A7 [Rezolvat]: Euler–Lagrange și frontiera



- $F = \eta \dot{x}^2 + \lambda \sigma^2 x^2$: $\frac{d}{dt}\partial F/\partial \dot{x} = \partial F/\partial x \Rightarrow \ddot{x} = \kappa_c^2 x$, $\kappa_c^2 = \lambda\sigma^2/\eta = 0{,}722$; cu $x(0) = X$, $x(T) = 0$: $x(t) = X \sinh(\kappa_c(T-t))/\sinh(\kappa_c T)$, $x(1) = 427\,151$
- $\kappa_c = 0{,}8497$ față de $\kappa = 0{,}8463$: apropiate aici fiindcă discretizarea și $\tilde{\eta}$ se compensează parțial ($\kappa\tau \approx 0{,}85$ nu este mic); punctele: programul discret, linia punctată: cel continuu

MFM_ch10_seminar

A8 [Propus]: Un vînzător cu aversiune mai mare la risc — cerință

- **Exemplu de urmat**: A7
- **Întrebarea**: cum se schimbă programul cînd aversiunea la risc este de zece ori mai mare?
- **Date și ipoteze**
  - Parametrii din A7: $X = 10^{\mathbf{6}}$, $T = N = 5$, $\sigma = 0{,}95$, $\eta = 2{,}5 \cdot 10^{-\mathbf{6}}$, $\gamma = 2{,}5 \cdot 10^{-\mathbf{7}}$; se schimbă doar λ, la $2 \cdot 10^{-\mathbf{5}}$
- **Cerințe**
  1. Calculați $\tilde{\kappa}^{\mathbf{2}}$, κ, κ_c și deținerile x_k.
  2. Calculați $\mathbb{E}[C]$ și $\mathrm{sd}[C]$ și rapoartele dintre noile valori și cele din A7, cu interpretarea lor.
  3. Suprapuneți deținerile din A7 și A8 pe același grafic.
  4. Opțional: explicați de ce κ și κ_c diferă acum mult mai mult.
- **Raportați**: un tabel de comparație și graficul
- **Interpretare**: cînd ar alege un investitor acest program?

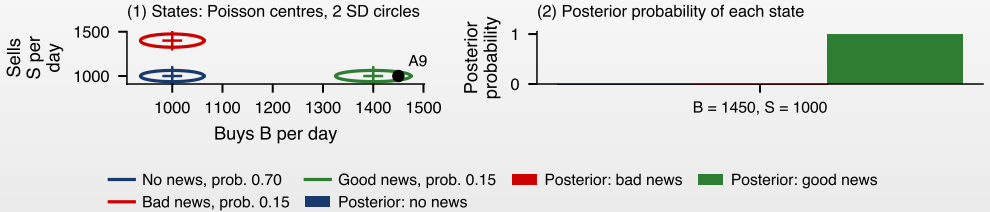
Înapoi Lista exercițiilor

A9 [Rezolvat]: Verosimilitatea PIN, pas cu pas — cerință

- **Întrebările**
  - cum evaluăm verosimilitatea unei zile de cumpărări și vînzări fără depășire numerică?
  - ce informație oferă ea despre acea zi?
- **Date și ipoteze**
  - Easley–Kiefer–O'Hara–Paperman (EKOP) Easley et al. (1996), trei stări: fără știri, prob. $1-\alpha$, $(B,S) \sim \text{Poisson}(\varepsilon_b)$, $\text{Poisson}(\varepsilon_s)$; știre proastă, $\alpha\delta$, vînzările cu intensitatea $\mu+\varepsilon_s$; știre bună, $\alpha(1-\delta)$, cumpărările cu intensitatea $\mu+\varepsilon_b$
  - $\alpha = 0{,}3$, $\delta = 0{,}5$, $\mu = 400$, $\varepsilon_b = \varepsilon_s = 1\,000$; o zi cu $B = 1\,450$, $S = 1\,000$
- **Cerințe**
  1. Scrieți verosimilitatea ca amestec de trei componente și calculați $\text{PIN} = \alpha\mu/(\alpha\mu+\varepsilon_b+\varepsilon_s)$.
  2. Explicați de ce evaluarea directă a termenilor eșuează în dublă precizie.
  3. Scrieți fiecare componentă logaritmică ponderată, combinați-le ca $m + \ln \sum e^{\ell_j - m}$ cu $m = \max_j \ell_j$ și normalizați aceleași componente în probabilități a posteriori.
- **Raportați**: PIN, $\ln L$, cele trei probabilități a posteriori
- **Interpretare**: este probabilitatea a posteriori a unei știri bune același lucru cu PIN?

◂ Înapoi ▸ Lista exercițiilor

A9 [Rezolvat]: Componente logaritmice stabile și probabilități a posteriori



- PIN $= 120/2\,120 = 0{,}0566$; direct, $e^{-\mathbf{1\,000}} \to 0$ și $1\,000^{\mathbf{1450}} \to \infty$: $0 \cdot \infty =$ NaN Lin & Ke (2011). Factorul $K = e^{-\varepsilon_b - \varepsilon_s}(\mu + \varepsilon_b)^B(\mu + \varepsilon_s)^S/(B!S!)$, $x_b = \varepsilon_b/(\mu + \varepsilon_b)$, x_s la fel; $\ln B!$ prin log-gamma
- Componentele: $\ell_{\mathbf{1}} = \ln(1-\alpha) + B \ln x_b + S \ln x_s$, $\ell_{\mathbf{2}} = \ln \alpha\delta - \mu + B \ln x_b$, $\ell_{\mathbf{3}} = \ln \alpha(1-\delta) - \mu + S \ln x_s$; $\ln L = \ln K + m + \ln \sum e^{\ell_j - m} = -11{,}71$; a posteriori $e^{\ell_j - m}/\sum e^{\ell_i - m}$: știre bună ≈ 1
- Interpretare: 450 de cumpărări în plus față de o abatere standard Poisson de circa 32 clasifică ziua aproape sigur; PIN $=$ 0,0566 este ponderea necondiționată așteptată a tranzacțiilor informate, nu această probabilitate a posteriori

MFM_ch10_seminar

A10 [Propus]: Citirea zilelor cu modelul PIN — cerință

- **Exemplu de urmat**: A9
- **Întrebările**
  - cu ce grad de certitudine pot clasifica o zi parametrii PIN cunoscuți?
  - ce se schimbă cînd parametrii sînt necunoscuți?
- **Date și ipoteze**
  - Parametrii din A9; două zile observate (a) $B = 1\,000$, $S = 1\,420$; (b) $B = S = 1\,000$; (c) o zi ipotetică ambiguă, $B = 1\,195$, $S = 1\,000$
- **Cerințe**
  1. Pentru fiecare zi, calculați log-verosimilitatea stabilă și cele trei probabilități a posteriori.
  2. Calculați PIN cînd μ se dublează la 800.
  3. Conceptual: explicați de ce estimarea lui $(\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ din multe zile rămîne dificilă pentru acțiuni foarte lichide.
- **Raportați**: un tabel al zilelor cu probabilitățile a posteriori și noul PIN
- **Interpretare**: poate fi interpretat un dezechilibru al ordinelor întotdeauna ca semn de informație privată?

Înapoi Lista exercițiilor

B5 [Propus]: Ilichiditatea Amihud pe trei piețe — cerință

- **Exemplu de urmat**: B2, B3
- **Întrebările**
  - de cîte ori este BVB mai puțin lichidă decît SUA după raportul Amihud?
  - cît de precis este multiplul?
  - cît de stabil este?
- **Date și ipoteze**
  - Cele 25 de active din B3 (BVB include ETF-ul pe BET); apartenența la grupuri fixă
  - Agregarea: mediem rapoartele zilnice pe fiecare activ, 19 sep. 2024 – 18 sep. 2026, apoi luăm mediana grupului
  - Anii: 19 sep. 2024 – 18 sep. 2025 și 19 sep. 2025 – 18 sep. 2026
- **Cerințe**
  1. Calculați cele trei mediane pe grupuri și raportul mediana BVB / mediana SUA.
  2. Construiți un interval de 95% reeșantionînd activele cu întoarcere, independent în fiecare grup (1 000 de reeșantionări), recalculînd de fiecare dată raportul medianelor.
  3. Calculați corelația Spearman a clasamentelor activelor în cei doi ani.
- **Raportați**: medianele, raportul cu intervalul lui, corelația de rang și grafice pe scală logaritmică
- **Interpretare**
  - de ce depinde intervalul?
  - este ilichiditatea o caracteristică stabilă a activului?

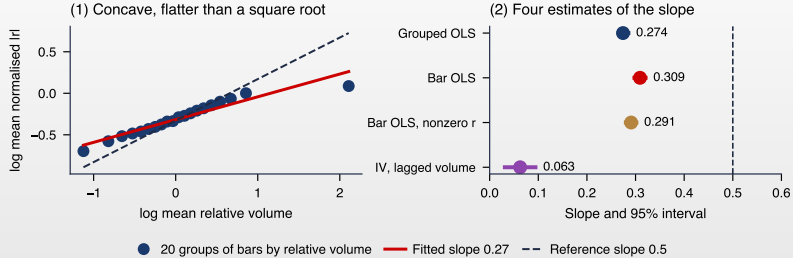
Înapoi Lista exercițiilor

B6 [Rezolvat]: Volum și mișcări de preț în barele SPY — cerință

- **Întrebările**
  - cum crește randamentul absolut cu volumul relativ în barele de 5 minute?
  - poate fi interpretată panta ca lege a impactului?
- **Date și ipoteze**
  - Barele SPY de 5 minute cu volum pozitiv: 115 052 de bare în 1 477 de zile (115 206 bare minus 154 fără volum)
  - Volumul relativ v = volumul barei / volumul median la aceeași oră; $z = |r|/\sigma_{\mathbf{zi}}$, $\sigma_{\mathbf{zi}}$ abaterea standard a randamentelor de 5 minute din acea zi
- **Cerințe**
  1. Împărțiți barele în 20 de grupe egale după v (limitele fixate pe întregul eșantion), regresați $\ln \bar{z}$ pe $\ln \bar{v}$ ($\bar{z}$, $\bar{v}$: mediile pe grupe) și dați un interval bootstrap pe zile de 95% (300 de reeșantionări, aceleași limite).
  2. La nivel de bară, estimați $\ln(|r| + 10^{-\mathbf{6}}) - \ln \sigma_{\mathbf{zi}} = a + b \ln v + e$ cu erori grupate pe zile, testați $b = 0{,}5$ și repetați fără barele cu $r = 0$.
  3. Extindere: instrumentați $\ln v$ cu logaritmul volumului relativ al aceleiași bare din ziua de tranzacționare precedentă; raportați prima etapă și panta estimată prin metoda celor mai mici pătrate în două etape (2SLS, *Two-Stage Least Squares*), cu SE grupată.
- **Raportați**: relația pe grupe, cele patru estimări ale pantei cu intervale și prima etapă
- **Interpretare**
  - este relația concavă?
  - este legea rădăcinii pătrate Tóth et al. (2011)?
  - este plauzibilă restricția de excludere?

◂ Înapoi ▸ Lista exercițiilor

B6 [Rezolvat]: Relația estimată și pantele



- Panta pe grupe 0,274 [0,262; 0,289]; la nivel de bară 0,309 (SE 0,008), $t = (0{,}5 - 0{,}309)/0{,}008 \approx 24$: 0,5 respins; 0,8% din bare au $r = 0$, fără ele 0,291 (SE 0,008)
- Panta în prima etapă 0,339 ($t = 32{,}8$): instrument relevant; panta 2SLS 0,063 (SE 0,018); relevanța nu dovedește excluderea: volumul de ieri la acea oră reflectă activitatea persistentă și știrile programate (corelația cu lag a lui $\ln|r|$ normalizat 0,043)
- Interpretare: concavă, dar mai plată decît o rădăcină pătrată; o bară amestecă cumpărători și vînzători, iar legea privește metaordinul net al unui singur investitor, deci panta 0,5 nu o testează direct

MFM_ch10_seminar

B7 [Propus]: Ceasul Bitcoin — cerință

- **Exemplu de urmat**: B1
- **Întrebarea**: urmează o piață deschisă 24/7 programul piețelor tradiționale?
- **Date și ipoteze**
  - Bare Bitcoin de 5 minute, UTC, 749 de zile (535 lucrătoare, 214 de weekend); randamentele calculate peste bare lipsă sînt eliminate; acoperirea este incompletă: doar 106 zile au toate cele 288 de bare (minimum 26); toate zilele sînt păstrate
  - Ponderarea: întîi media $|r|$ a fiecărei zile într-o fereastră (orele $[h_1, h_2)$), apoi media pe zile; rapoartele sînt rapoarte ale acestor medii
- **Cerințe**
  1. În zilele lucrătoare, calculați media $|r|$ în 13:00–17:00 UTC și în 03:00–11:00 UTC și raportul lor, cu un interval bootstrap pe zile de 95%.
  2. Calculați raportul zile lucrătoare / weekend al mediei $|r|$, cu un interval bootstrap pe zile de 95% (1 000 de reeșantionări, seed 42; extrageți zile întregi din eșantionul comun, deci numărul zilelor lucrătoare și de weekend poate varia).
  3. Desenați cele două profile orare cu benzi bootstrap pe zile.
- **Raportați**: cele două rapoarte cu intervale și profilele
- **Interpretare**: ce informații despre spread, adîncime și volum ar fi necesare pentru o decizie de execuție, dincolo de volatilitate?

◂ Înapoi ▸ Lista exercițiilor

B8 [Rezolvat]: Descoperirea prețului: ETF-ul pe BET și BET-TR — cerință (1/2)

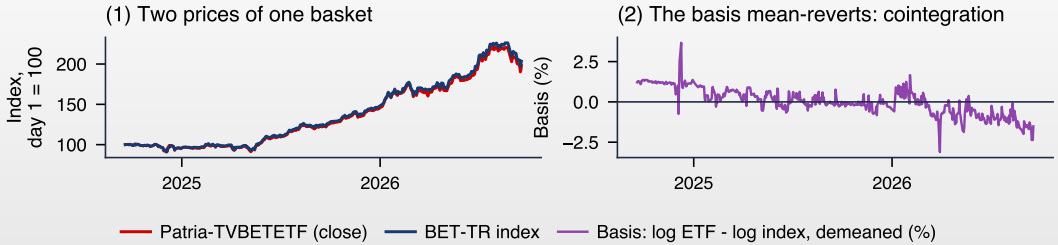
- **Întrebările**
  - domină descoperirea prețului ETF-ul pe BET sau indicele BET-TR?
  - cît de siguri putem fi cu închideri zilnice?
- **Date și ipoteze**
  - Logaritmul închiderilor $y_t = (\ln P_t^{\mathrm{ETF}}, \ln P_t^{\mathrm{idx}})'$ Patria-TVBETETF (zilele cu tranzacții) și BET-TR, zile comune 2024-09-19 – 2026-09-18, 493 de observații VECM; ETF-ul reinvestește dividendele; BET-TR nu se tranzacționează
  - VECM: $\Delta y_t = a_0 + \alpha z_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta y_{t-i} + e_t$, $z_{t-1} = \ln P_{t-1}^{\mathrm{ETF}} - \ln P_{t-1}^{\mathrm{idx}}$, $\mathrm{Var}(e_t) = \Omega$; constantă nerestricționată; p laguri ale *diferențelor*, alese după BIC (*Bayesian Information Criterion*, criteriul informațional bayesian) între 1 și 10 Johansen (1991)

înapoi Lista exercițiilor

B8 [Rezolvat]: Descoperirea prețului: ETF-ul pe BET și BET-TR — cerință (2/2)

- **Cerințe**
  1. Desenați prețurile rebazate și baza; aplicați testul trace Johansen (rangul unu nu stabilește singur $\beta = (1, -1)'$, impus aici).
  2. Estimați VECM și testați $\alpha_{\mathrm{ETF}} = 0$ și $\alpha_{\mathrm{idx}} = 0$ la 5%.
  3. Cu $\alpha_\perp = (-\alpha_{\mathrm{idx}}, \alpha_{\mathrm{ETF}})'$ calculați ponderea componentei $\alpha_\perp / \sum \alpha_\perp$ Gonzalo & Granger (1995) și ponderea informațională $([\alpha_\perp' F]_j)^2 / \alpha_\perp' \Omega \alpha_\perp$ pentru ambele ordonări Cholesky, cu F factorul inferior triunghiular, $FF' = \Omega$, și $[\cdot]_j$ elementul j Hasbrouck (1995).
  4. Bootstrap: reeșantionați blocuri suprapuse de 20 de rînduri ale regresiei $(\Delta y_t, 1, z_{t-\mathbf{1}}, \Delta y_{t-\mathbf{1}}, \ldots)$, de 1 000 de ori; apoi reestimați α pe fiecare jumătate.
- **Raportați**: testul, α cu intervale asimptotice și bootstrap, limitele de ordonare cu intervalele capetelor, cele două jumătăți
- **Interpretare**
  - care preț domină descoperirea prețului?
  - de ce sînt limitele largi?
  - de ce poate ponderea componentei să iasă din $[0, 1]$?

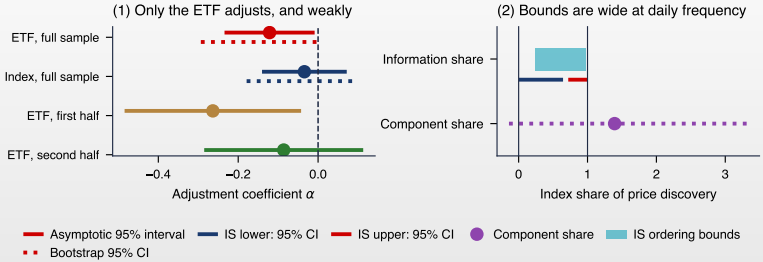
B8 [Rezolvat]: Prețuri, test și ajustare



- BIC: 1 lag; trace 49,6 pentru $r = 0$ (valoarea critică 5% 15,49), 0,10 pentru $r \leq 1$ (3,84): un vector de cointegrare
- $\alpha_{\text{ETF}} = -0{,}12$ ($t = -2{,}11$, respins la 5%), $\alpha_{\text{idx}} = -0{,}03$ ($t = -0{,}63$); bootstrap pe blocuri: ETF $[-0{,}29; 0{,}01]$, indice $[-0{,}18; 0{,}09]$: testul asimptotic respinge, intervalul bootstrap atinge 0, deci dovada este slabă

MFM_ch10_seminar

B8 [Rezolvat]: Ponderi, limite și jumătăți



- Ponderea informațională a indicelui în [0,25; 0,98] (limite de ordonare, nu CI); CI bootstrap pentru limita inferioară [0,00; 0,64], pentru cea superioară [0,72; 1,00]; ponderea componentei 1,39 [−0,14; 3,32]
- Jumătățile: ETF $\alpha = -0{,}26$ ($t = -2{,}34$), apoi -0,09 ($t = -0{,}85$); indicele nu este niciodată semnificativ ($t = -0{,}77$, 0,47)
- Interpretare: ETF-ul pare să se ajusteze spre indice; corelația reziduurilor 0,93 face ordonarea decisivă, de aici limitele largi; ambii α negativi împing ponderea componentei în afara [0, 1]

MFM_ch10_seminar

C2: panelul, eșantionul și datele de intrare

- Date pentru C2 (cerința urmează): observații zilnice între 19 septembrie 2016 și 18 septembrie 2026; în toate etapele intră doar lunile complete: indicele de ilichiditate și modelul autoregresiv AR(1), octombrie 2016–august 2026; regresiile randamentelor, noiembrie 2016–august 2026. Această fereastră lungă se folosește doar aici.
- Acțiuni SUA: AAPL, MSFT, NVDA, JPM, GME, MSTR. Celelalte criptoactive: ETH, SOL, XRP, DOGE, ADA, LTC.
- Panelul este neechilibrat: un activ intră în media transversală lunară doar dacă are cel puțin zece zile valide. Observațiile dinaintea listării nu se completează.
- Încărcați randamentele și valoarea tranzacționată cu `returns(key, start)` și `dollar_volume(key, start)`, cu `start="2016-09-19"`; aliniați datele înainte să calculați rapoartele.

Incertitudinea după corectarea persistenței

- În C2, $\hat{\phi}_c$ este persistența AR(1) corectată (formula: Etapa 3 a cerinței); v_t^c este reziduul calculat cu acest coeficient și cu interceptul reestimat.
- Notăm cu $\hat{d}$ coeficientul lui v_t^c din regresia augmentată a randamentului. s_b este eroarea standard HAC a coeficientului ilichidității decalate, condiționat de corecția estimată.
- Pentru modelul AR(1), s_ϕ este eroarea standard a lui $\hat{\phi}$, iar T numără tranzițiile utilizabile. Aproximarea prin metoda delta: $\widehat{\mathrm{Var}}(\hat{\phi}_c) = (1 + 3/T + 9/T^2)^2 s_\phi^2, \qquad s_{b,c} = \sqrt{s_b^2 + \hat{d}^2 \widehat{\mathrm{Var}}(\hat{\phi}_c)}.$
- Implementarea limitează persistența corectată la 0,9999. Derivata afișată este aproximarea pentru interiorul domeniului, fără plafonare, deci nu este validă cînd plafonul este atins. Amihud & Hurvich (2004)
- Inferența comună: bootstrap pe blocuri suprapuse de 12 luni din tripletele (L_t, L_{t-1}, r_t); fiecare replicare reestimează AR(1), corecția plafonată și regresia finală. Raportați intervalul percentilic de 95%, ponderea replicărilor cu plafonul atins și coeficientul fără plafon.

C2 [Propus]: Prima de ilichiditate pe BVB — cerință (1/2)

- **Exemplu de urmat**: B4, B5
- **Întrebările**: Amihud (2002) prezice că ilichiditatea se reflectă în randamente Amihud (2002)
  - crește ilichiditatea așteptată randamentele viitoare pe BVB?
  - scade o creștere neașteptată a ilichidității randamentele curente?
- **Date și ipoteze**
  - Piețe: BVB (randament BET-TR; ilichiditatea din cele zece acțiuni blue-chip), SUA (SPY; șase acțiuni americane), cripto (Bitcoin; alte șase monede); randamente logaritmice lunare în %, 2016–2026
  - Indicele: media rapoartelor Amihud pe activ și lună (≥ 10 zile); media transversală a logaritmului, L_t; șocul: reziduul modelului autoregresiv AR(1) $L_t = \phi_{\mathbf{0}} + \phi L_{t-\mathbf{1}} + v_t$, cu $\hat{\phi}$ persistența estimată și T numărul de luni

C2 [Propus]: Prima de ilichiditate pe BVB — cerință (2/2)

- **Cerințe**
  1. Etapa 1: construiți indicele și șocul pentru fiecare piață.
  2. Etapa 2: estimați $r_t = a + \mathbf{b_1} L_{t-\mathbf{1}} + \mathbf{b_2}\,\text{șoc}_t + e_t$ cu SE HAC (6 laguri); r_t: randamentul lunar al pieței; Amihud prezice $\mathbf{b_1} > 0$ (investitorii cer o primă pentru ilichiditatea așteptată) și $\mathbf{b_2} < 0$ (un șoc de ilichiditate scade prețurile).
  3. Etapa 3: corectați $\mathbf{b_1}$ pentru deplasarea Stambaugh Stambaugh (1999) (deplasarea lui $\hat{\mathbf{b}}_{\mathbf{1}}$ în eșantioane mici, cînd L_t este persistent și șocurile lui sînt corelate cu randamentele) cu estimatorul cu deplasare redusă: $\hat{\phi}_c = \hat{\phi} + (1 + 3\hat{\phi})/T + 3(1 + 3\hat{\phi})/\mathbf{T^2}$, reziduurile v_t^c din $\hat{\phi}_c$, regresăm r_t pe $L_{t-\mathbf{1}}$ și v_t^c, cu SE care include incertitudinea lui $\hat{\phi}_c$ Amihud & Hurvich (2004).
  4. Etapa 4 (sortări): în fiecare ianuarie ordonați acțiunile blue-chip cu peste 150 de zile de tranzacționare după Amihud din anul anterior; componența jumătăților nelichidă și lichidă se stabilește anual, iar ponderile egale se refac lunar (media randamentelor lunare simple totale ale membrilor); testați diferența medie cu SE HAC.
- **Raportați**: un tabel pe piețe, graficele de referință și o discuție despre putere (puține acțiuni, zece ani) Pele, Mazurencu-Marinescu & Nijkamp (2013)
- **Interpretare**: care dintre cele două predicții ale lui Amihud este confirmată pe fiecare piață?

C3 [Propus]: Costul competiției pentru viteză în SPY — cerință (1/2)

- **Exemplu de urmat**: B1, B4
- **Întrebările**: dacă taxa arbitrajului de latență estimată la Londra s-ar aplica la SPY
  - cît de mare ar fi premiul?
  - cît s-ar schimba în zilele de stres?
- **Date și ipoteze**
  - Studiul de caz Aquilina, Budish & O'Neill (2022): taxa 0,419 pb (tabelul X); tabelul XV: col. 2 $\Pi_t = 0{,}4213 \cdot 10^{-4} V_t$; col. 6 $\Pi_t = 10^{-4}(0{,}3354\, V_t + 0{,}0066\, \sigma_t \bar{V})$; scenariile din tabelul XIV 0,20–0,74 pb; zilele complete SPY 2021–2025
  - Unități: Π_t, V_t în USD; V_t = suma închidere × volum pe 78 de bare; $\sigma_t = 100\sqrt{252 \sum r^2}$ în %; $\bar{V}$ media anuală SPY, nu media eșantionului londonez
  - Rezultatele sînt implicații ale coeficienților transferați, nu profituri măsurate ale curselor din SUA

Înapoi Lista exercițiilor

C3 [Propus]: Costul competiției pentru viteză în SPY — cerință (2/2)

- **Cerințe**
  1. Datele zilnice: calculați V_t și σ_t; raportați mediile anuale și corelația anuală dintre V_t și σ_t, cu intervale de 95% din bootstrap pe zile i.i.d. și pe blocuri suprapuse de 20 de zile; comparați cu corelația zilnică volum–volatilitate de 0,82 din eșantionul londonez (secțiunea VI.A, discuția tabelului XV) Aquilina, Budish & O'Neill (2022).
  2. Extrapolarea anuală: calculați premiul fiecărui an $\sum_t \Pi_t$ după col. 2, col. 6 și cele două scenarii.
  3. Arătați că după col. 6 taxa anuală este $0{,}3354 + 0{,}0066\,\bar{\sigma}$ pb și găsiți pragul lui $\bar{\sigma}$ peste care col. 6 depășește col. 2.
  4. Zilele de stres: desenați taxa la valoarea tranzacționată medie, $0{,}3354 + 0{,}0066\,\sigma_t$, și raportul efectiv $\Pi_t / V_t = 0{,}3354 + 0{,}0066\,\sigma_t\,\bar{V}/V_t$; rezolvați complet ziua de 7 aprilie 2025.
- **Raportați**: tabelul anual, deducerea și pragul, graficul zilnic și calculul complet pentru 7 aprilie 2025
- **Interpretare**
  - ce ipoteze transferă la SPY de azi o taxă măsurată la Londra în 2015?
  - pot datele disponibile stabili semnul efectului lor?

Acronime folosite în acest material (1/2)

- **AAPL**: Apple Inc. (Nasdaq ticker) — *Apple Inc., simbolul bursier Nasdaq*
- **ACF**: AutoCorrelation Function — *funcția de autocorelație*
- **AI**: Artificial Intelligence — *inteligență artificială*
- **AR**: AutoRegressive (model) — *model autoregresiv*
- **BET**: Bucharest Exchange Trading (index) — *indicele de referință al Bursei de Valori București*
- **BET-TR**: BET Total Return (index) — *indicele BET cu dividendele reinvestite*
- **BIC**: Bayesian Information Criterion — *criteriul informațional bayesian*
- **BNR**: Banca Națională a României
- **BVB**: Bursa de Valori București
- **CBOE**: Chicago Board Options Exchange — *bursa de opțiuni din Chicago*
- **CI**: Confidence Interval — *interval de încredere*
- **DVOL**: Dollar Volume — valoarea tranzacționată, în USD
- **EKOP**: Easley–Kiefer–O'Hara–Paperman (model of informed trading) — *modelul Easley–Kiefer–O'Hara–Paperman al tranzacționării informate*
- **ETF**: Exchange-Traded Fund — *fond tranzacționat la bursă*
- **GME**: GameStop Corp. (NYSE ticker) — *GameStop Corp. (simbolul bursier NYSE)*
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors) — *erori standard robuste la heteroscedasticitate și autocorelație*
- **JPM**: JPMorgan Chase & Co. (NYSE ticker) — *JPMorgan Chase & Co. (simbolul bursier NYSE)*

Acronime folosite în acest material (2/2)

- **MA**: Moving Average — *medie mobilă*
- **MSFT**: Microsoft Corporation (Nasdaq ticker) — *Microsoft Corporation (simbolul bursier Nasdaq)*
- **MSTR**: Strategy Inc., formerly MicroStrategy (Nasdaq ticker) — *Strategy Inc., fosta MicroStrategy (simbolul bursier Nasdaq)*
- **NVDA**: NVIDIA Corporation (Nasdaq ticker) — *NVIDIA Corporation, simbolul bursier Nasdaq*
- **NVIDIA**: NVIDIA Corporation (company name, US chip designer) — *NVIDIA Corporation (nume de companie, proiectant american de cipuri)*
- **OHLC**: Open, High, Low, Close — *deschidere, maxim, minim, închidere*
- **OLS**: Ordinary Least Squares — *metoda celor mai mici pătrate*
- **OMV**: Österreichische Mineralölverwaltung — *Administrația austriacă a uleiurilor minerale (grupul OMV)*
- **PIN**: Probability of INformed trading — *probabilitatea tranzacționării informate*
- **SD**: Standard Deviation — *abatere standard*
- **SE**: Standard Error — *eroare standard*
- **SUA**: Statele Unite ale Americii
- **US**: United States — *Statele Unite ale Americii*
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné) — *timpul universal coordonat*
- **VECM**: Vector Error-Correction Model — *model vectorial cu corecția erorilor*
- **VIX**: CBOE Volatility IndeX — *indicele de volatilitate CBOE*
- **XRP**: XRP (the native token of the XRP Ledger) — *tokenul nativ al registrului XRP Ledger*

Bibliografie (1/2)

- Abdi, F., Ranaldo, A. (2017). A simple estimation of bid-ask spreads from daily close, high, and low prices. *Review of Financial Studies*, 30(12), 4437–4480.
- Almgren, R., Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3(2), 5–39.
- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31–56.
- Amihud, Y., Hurvich, C.M. (2004). Predictive regressions: a reduced-bias estimation method. *Journal of Financial and Quantitative Analysis*, 39(4), 813–841.
- Aquilina, M., Budish, E., O'Neill, P. (2022). Quantifying the high-frequency trading "arms race". *Quarterly Journal of Economics*, 137(1), 493–564.
- Brauneis, A., Mestel, R., Riordan, R., Theissen, E. (2021). How to measure the liquidity of cryptocurrency markets? *Journal of Banking & Finance*, 124, 106041.
- Corwin, S.A., Schultz, P. (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance*, 67(2), 719–760.
- Easley, D., Kiefer, N.M., O'Hara, M., Paperman, J.B. (1996). Liquidity, information, and infrequently traded stocks. *Journal of Finance*, 51(4), 1405–1436.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7(1), 1–26.

Bibliografie (2/2)

- Glosten, L.R., Milgrom, P.R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71–100.
- Gonzalo, J., Granger, C.W.J. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, 13(1), 27–35.
- Harris, L. (1990). Statistical properties of the Roll serial covariance bid/ask spread estimator. *Journal of Finance*, 45(2), 579–590.
- Hasbrouck, J. (1995). One security, many markets: determining the contributions to price discovery. *Journal of Finance*, 50(4), 1175–1199.
- Hasbrouck, J. (2007). Empirical Market Microstructure. Oxford University Press.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551–1580.
- Kyle, A.S. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6), 1315–1335.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lin, H.-W.W., Ke, W.-C. (2011). A computing bias in estimating the probability of informed trading. *Journal of Financial Markets*, 14(4), 625–640.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Pele, D.T., Mazurencu-Marinescu, M., Nijkamp, P. (2013). Herding behaviour, bubbles and log periodic power laws in illiquid stock markets: a case study on the Bucharest Stock Exchange. Social Science Research Network (SSRN) Working Paper 2305945.
- Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance*, 39(4), 1127–1139.
- Stambaugh, R.F. (1999). Predictive regressions. *Journal of Financial Economics*, 54(3), 375–421.
- Tóth, B., Lempérière, Y., Deremble, C., de Lataillade, J., Kockelkoren, J., Bouchaud, J.-P. (2011). Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2), 021006.