









Statistica piețelor financiare

Capitolul 8: Estimatori de volatilitate



Daniel Traian PELE

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

Acronime folosite în acest material (1/2)

- **ACF**: Autocorrelation Function (funcția de autocorelație)
- **AI**: Artificial Intelligence (inteligență artificială)
- **AR**: AutoRegressive (model); Abnormal Return (model autoregresiv; randament anormal, în studiile de eveniment)
- **ARCH**: AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă)
- **ARCH-LM**: ARCH Lagrange Multiplier (test) — testul multiplicatorului Lagrange pentru efecte ARCH
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București
- **BVB**: Bursa de Valori București
- **CBOE**: Chicago Board Options Exchange (bursa de opțiuni din Chicago)
- **DM**: Diebold–Mariano (test) — testul Diebold–Mariano de comparare a prognozelor
- **EODHD**: EOD Historical Data (market-data provider) — furnizor de date de piață
- **ES**: Expected Shortfall (pierderea așteptată în coadă)
- **EWMA**: Exponentially Weighted Moving Average (medie mobilă ponderată exponențial)
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă generalizată)

Acronime folosite în acest material (2/2)

- **GK**: Garman–Klass (volatility estimator) — estimatorul de volatilitate Garman–Klass
- **HAR**: Heterogeneous AutoRegressive (model) — model autoregresiv heterogen
- **IGARCH**: Integrated GARCH (GARCH integrat)
- **LM**: Lagrange Multiplier (multiplicatorul Lagrange)
- **MSE**: Mean Squared Error (eroarea pătratică medie)
- **MZ**: Mincer–Zarnowitz (regression) — regresia Mincer–Zarnowitz
- **OHLC**: Open, High, Low, Close (deschidere, maxim, minim, închidere)
- **OMV**: Österreichische Mineralölverwaltung (Administrația austriacă a uleiurilor minerale, grupul OMV)
- **QLIKE**: Quasi-LIKElihood (loss function) — funcția de pierdere de tip cvasi-verosimilitate
- **RV**: Realised Variance (varianța realizată)
- **SE**: Standard Error (eroare standard)
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook) — colecția de Quantlet-uri a manualului Statistics of Financial Markets
- **VIX**: CBOE Volatility IndeX (indicele de volatilitate CBOE)
- **YZ**: Yang–Zhang (volatility estimator) — estimatorul de volatilitate Yang–Zhang

Întrebarea de azi și traseul

- **Întrebarea**: cît de mare este riscul unui activ azi, dacă nu îl putem observa niciodată direct?
  - Capitolul 7: randamentele sînt aproape imprevizibile, dar mărimea lor nu este; aici măsurăm această mărime
  - Capitolul 9 transformă măsurătorile într-un model (ARCH și GARCH)
- **Traseul** capitolului
  - ce este volatilitatea și de ce este latentă; volatilitatea istorică și alegerea ferestrei
  - EWMA (RiskMetrics); estimatorii de amplitudine, care folosesc prețurile de deschidere, maxim, minim și închidere
  - volatilitatea realizată; volatility clustering și testele lui
  - comportamentul pe termen lung, indicele VIX, efectul de levier; compararea prognozelor de volatilitate

Rezultatele învățării

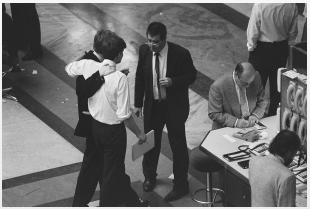
- Definiți volatilitatea, anualizați-o cu frecvența reală de tranzacționare și explicați de ce este o mărime latentă
- Calculați volatilitatea istorică și volatilitatea EWMA și alegeți o fereastră sau un factor de atenuare
- Calculați estimatorii Parkinson, Garman–Klass, Rogers–Satchell și Yang–Zhang și enunțați ipotezele lor
- Detectați volatility clustering cu ACF-ul randamentelor la pătrat, testul Ljung–Box și testul ARCH-LM
- Comparați volatilitatea implicită cu cea realizată și evaluați prognozele de volatilitate cu MSE și QLIKE

Bibliografie și instrumente

- Manual: Franke, Härdle & Hafner (2019), cap. 13 (introducerea: volatilitatea și heteroscedasticitatea condiționată)
  - exerciții rezolvate: Borak, Härdle & López-Cabrera (2013), cap. 13
- Volatilitatea și faptele ei stilizate: Tsay (2010), cap. 3.1; o sinteză despre prognoza volatilității: Poon & Granger (2003)
- Quantlet-urile Python ale capitolului: Quantlets/Ch_08
  - portate din Quantlet-urile SFE SFEkurgarch (boltirea unui proces GARCH) și SFEvolnonparest (volatilitate neparametrică)
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: Applied Time Series Analysis with Python

Volatility clustering și utilizările volatilității

- „Variațiile mari tind să fie urmate de variații mari, de orice semn, iar variațiile mici, de variații mici” Mandelbrot (1963)
  - după crahul din 19 octombrie 1987, piețele au rămas agitate săptămîni la rînd
- Volatilitatea intervine în aproape toate mărimile din finanțe
  - măsurile de risc: VaR și ES (Capitolul 10), raportul Sharpe (Capitolul 1)
  - prețurile opțiunilor, marjele cerute de burse, mărimea pozițiilor
- Acest capitol: **măsurarea** ei din prețuri și testarea volatility clustering



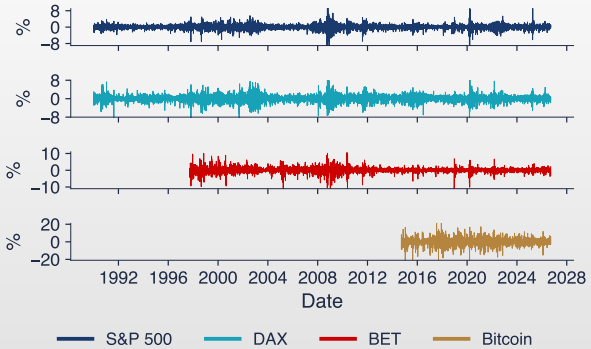
Bursa din Amsterdam, 21 octombrie 1987, la două zile după crah

Foto: Bart Molendijk / Anefo (1987); CC0; Wikimedia Commons

Volatilitatea: definiție și unități de măsură

- Randamentul logaritmic zilnic în %: $r_t = 100\,(\ln P_t - \ln P_{t-1})$ (Capitolul 1)
  - **volatilitatea** = abaterea standard a randamentelor, $\sigma = \sqrt{\mathrm{Var}(r_t)}$
  - are unitatea de măsură a randamentului (% pe zi); varianța σ^2 se măsoară în $\%^2$
- **Anualizarea**: cu A observații pe an și randamente necorelate, $\sigma_{\text{an}} = \sigma_{\text{zi}}\sqrt{A}$
  - A = frecvența reală: circa 252 pentru S&P 500, 250 pentru BET, 365 pentru Bitcoin (tranzacționat în fiecare zi)
  - regula $\sqrt{A}$ este regula rădăcinii pătrate a timpului din Capitolul 7: cere randamente necorelate
- Volatilitatea măsoară dispersia în ambele sensuri: un cîștig mare contează la fel de mult ca o pierdere mare
  - există și alte măsuri ale incertitudinii, de exemplu entropia informațională a randamentelor Pele, Lazar & Dufour (2017)

Volatility clustering pe patru piețe



- Randamente logaritmice zilnice în %: S&P 500 și DAX din 1990, BET din 1997, Bitcoin din 2014; ultima zi: 18 septembrie 2026
- Anii liniștiți alternează cu luni agitate în toate cele patru serii

SFM_ch8_returns_volatility

Interpretarea volatilității pe patru piețe

	Volatilitate anuală	Cea mai mare variație zilnică	Data	Excesul de boltire
S&P 500	18,0%	−12,8%	16 martie 2020	10,9
DAX	21,8%	−13,1%	12 martie 2020	6,0
BET	23,4%	−13,1%	7 ianuarie 2009	9,9
Bitcoin	66,6%	−46,5%	12 martie 2020	12,0

- Cele mai mari variații reprezintă 8–13 abateri standard: valori practic imposibile în cazul distribuției Normale (Capitolul 2)
- Excesul de boltire este mult peste 0 (valoarea pentru distribuția Normală): cozi groase; secțiunea 6 arată că volatility clustering este una dintre cauze

SFM_ch8_returns_volatility

Caracterul latent al volatilității

- Modelul: randamentul este o medie constantă plus un șoc a cărui mărime se schimbă zilnic: $r_t = \mu + \sigma_t z_t$
  - μ: randamentul mediu; z_t: șocuri i.i.d., de medie 0 și variață 1
  - σ_t este **volatilitatea condiționată**: $\sigma_t^2 = \mathrm{Var}(r_t \mid \mathcal{F}_{t-1})$, cu $\mathcal{F}_{t-1}$ informația de la închiderea zilei $t-1$
  - **volatilitatea necondiționată** $\sigma = \sqrt{\mathrm{Var}(r_t)}$ este nivelul ei mediu pe termen lung
- **Latentă**: nu observăm niciodată σ_t, ci doar un randament pe zi
  - r_t^2 este o măsură nedeplasată, dar foarte zgomotoasă, a lui σ_t^2 (cu $\mu = 0$): $E[r_t^2 \mid \mathcal{F}_{t-1}] = \sigma_t^2$
  - o zi liniștită poate apărea într-o perioadă agitată, și invers
- Trei moduri de a obține mai multă informație despre σ_t
  - media pe multe zile: volatilitatea istorică, EWMA
  - întreaga traiectorie a zilei: prețurile maxim și minim
  - prețurile din timpul zilei: volatilitatea realizată

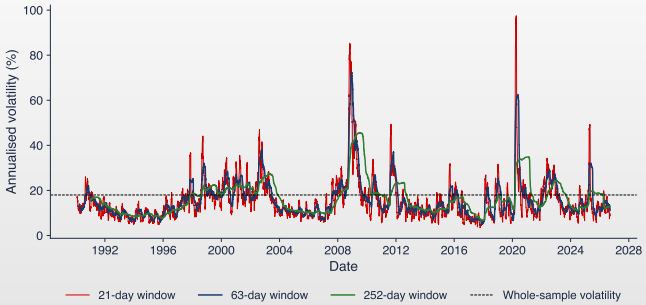
Exemplu rezolvat: anualizarea volatilității

- S&P 500 din 1990: abaterea standard zilnică $\hat{\sigma} = 1{,}1337\%$, $A = 251{,}7$ observații pe an (frecvența reală)
  - $\hat{\sigma}_{\text{an}} = 1{,}1337 \times \sqrt{251{,}7} = 1{,}1337 \times 15{,}866 = 18{,}0\%$
- Bitcoin din 2014: $\hat{\sigma} = 3{,}4852\%$ pe zi, 7 zile pe săptămînă
  - corect: $3{,}4852 \times \sqrt{365{,}3} = 3{,}4852 \times 19{,}114 = 66{,}6\%$
  - greșit, cu convenția de la acțiuni: $3{,}4852 \times \sqrt{252} = 3{,}4852 \times 15{,}875 = 55{,}3\%$
- Pentru a compara active: anualizați fiecare serie cu propriul calendar, apoi comparați

Volatilitatea istorică (pe fereastră mobilă)

- **Volatilitatea istorică** pe o fereastră cu ultimele n zile:
  - $\hat{\sigma}_t = \sqrt{\frac{1}{n-1}\sum_{j=0}^{n-1}(r_{t-j} - \bar{r}_t)^2}$, cu $\bar{r}_t$ media acelorași n randamente
  - j: cîte zile în urmă se află randamentul ($j = 0$ este ziua t); $\hat{\sigma}_t$: estimarea la închiderea zilei t
  - **fereastră mobilă**: în fiecare zi intră cel mai nou randament și iese cel mai vechi
- Varianta cu media zero: $\hat{\sigma}_t^2 = \frac{1}{n}\sum_{j=0}^{n-1} r_{t-j}^2$
  - pentru datele zilnice media este foarte mică în raport cu σ (0,03% față de 1,13% la S&P 500); dacă o fixăm la 0, eliminăm o estimare zgomotoasă
- Ferestre uzuale: $n = 21$ (o lună), 63 (un trimestru), 252 (un an) de zile de tranzacționare
  - fiecare $\hat{\sigma}_t$ se anualizează cu $\sqrt{A}$ și se reprezintă în ultima zi a ferestrei

S&P 500: volatilitatea pe ferestre mobile de 21, 63 și 252 de zile



- Volatilitatea S&P 500 pe ferestre mobile, anualizată; linia punctată: volatilitatea pe întregul eșantion, 18,0%
- **Întrebare pentru sală**: în ce fereastră ați avea încredere în ziua de după un crah?

SFM_ch8_returns_volatility

Interpretarea compromisului dintre ferestre

- Fereastra scurtă ($n = 21$): reacționează repede, dar este zgomotoasă
  - maximul 97,5% pe 6 aprilie 2020; variația zilnică tipică a lui $\ln \hat{\sigma}_t$ este 6,7%
- Fereastra lungă ($n = 252$): netedă, dar lentă
  - maximul doar 45,6%, atins pe 15 iulie 2009, la cîteva luni după vîrful crizei din 2008; variația zilnică 0,8%
- **Răspuns**: o fereastră scurtă sau EWMA (secțiunea următoare), deoarece media pe o fereastră lungă include încă zilele liniștite dinaintea crahului
  - pentru planificarea capitalului se preferă o fereastră lungă: se schimbă lent
- La sfîrșitul eșantionului: 9,6% (21 de zile), 11,1% (63 de zile), 12,9% (252 de zile)

SFM_ch8_returns_volatility

Precizia unei volatilități de selecție

- Cu n randamente Normale i.i.d., eroarea standard relativă a lui $\hat{\sigma}$ este $\mathrm{SE}(\hat{\sigma})/\sigma \approx 1/\sqrt{2n}$
  - $\mathrm{SE}(\hat{\sigma})$: abaterea standard a estimării de la un eșantion la altul; împărțită la σ, este o eroare relativă
  - $n = 21$: 15,4%; $n = 63$: 8,9%; $n = 252$: 4,5%
- Cu cozi groase (coeficient de boltire κ): $\mathrm{SE}(\hat{\sigma})/\sigma \approx \sqrt{(\kappa - 1)/(4n)}$ (metoda delta)
  - S&P 500: $\kappa = 13{,}9$, deci $n = 21$: 39%; $n = 252$: 11%
  - pentru $\kappa = 3$ (distribuția Normală) cele două formule coincid
- **Exemplu rezolvat**: ultima volatilitate pe 21 de zile a S&P 500 este 9,6%
  - interval aproximativ de 95%, randamente Normale: [6,7%; 12,5%]; cu boltirea observată: [2,2%; 17,0%]

1994: RiskMetrics și estimarea zilnică a volatilității

- J.P. Morgan a lansat RiskMetrics în octombrie 1994 și a publicat gratuit metodele și datele J.P. Morgan/Reuters (1996)
  - scopul: un etalon comun pentru riscul de piață, valabil pentru toate băncile
  - volatilitățile și corelațiile a peste 480 de serii, actualizate zilnic
- Metoda: o medie mobilă ponderată exponențial (EWMA) a randamentelor la pătrat
  - un singur parametru, **factorul de atenuare** λ: 0,94 pentru prognozele zilnice, 0,97 pentru cele lunare
  - pentru fiecare serie, λ minimizează rădăcina erorii pătratice medii a prognozelor de varianță; 0,94 este media ponderată pe toate seriile



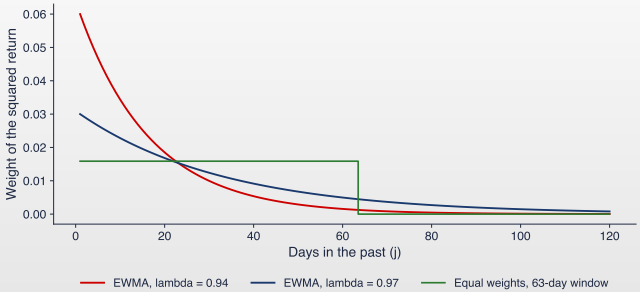
Clădirea J.P. Morgan de la 23 Wall Street, New York

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

Estimatorul EWMA

- Recurența **EWMA** (media zero): $\sigma_t^2 = \lambda\,\sigma_{t-1}^2 + (1-\lambda)\,r_{t-1}^2,\ \ 0 < \lambda < 1$
  - varianța de azi = o medie ponderată a variației de ieri și a randamentului de ieri la pătrat
  - λ: ponderea trecutului; un λ apropiat de 1 dă o estimare netedă, care reacționează lent
  - σ_t^2 folosește randamentele pînă în ziua $t-1$: este prognoza pentru ziua t
  - valoarea inițială: σ_1^2 = media primelor randamente la pătrat; efectul ei dispare după cîteva luni
- Prin dezvoltarea recurenței: $\sigma_t^2 = (1-\lambda)\sum_{j\geq 1}\lambda^{j-1} r_{t-j}^2$
  - ponderile $(1-\lambda)\lambda^{j-1}$ scad geometric și au suma 1
  - **timpul de înjumătățire**: lagul la care ponderea se înjumătățește, $\ln 0{,}5 / \ln\lambda$
- O fereastră mobilă dă ponderea $1/n$ fiecăreia dintre ultimele n zile și 0 tuturor zilelor mai vechi

Ponderile EWMA față de ponderile unei ferestre mobile



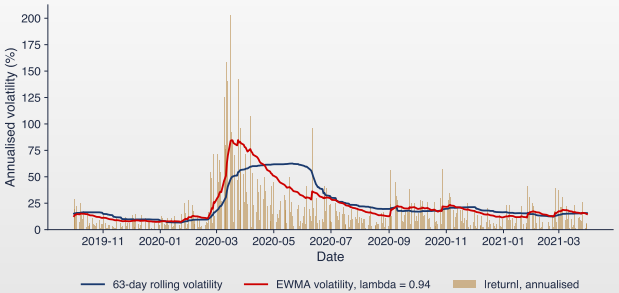
- $\lambda = 0{,}94$: timp de înjumătățire 11,2 zile, lag mediu $1/(1-\lambda) = 16{,}7$ zile, 99% din pondere în ultimele 75 de zile
- $\lambda = 0{,}97$: timp de înjumătățire 22,8 zile, 99% din pondere în ultimele 152 de zile

SFM_ch8_ewma

Exemplu rezolvat: o actualizare EWMA

- Varianța EWMA de ieri: $\sigma^2_{t-1} = 1{,}00$ ($\%^2$), adică o volatilitate zilnică de 1%; randamentul de ieri: $r_{t-1} = -3\%$
  - pasul 1: $\lambda\sigma^2_{t-1} = 0{,}94 \times 1{,}00 = 0{,}94$
  - pasul 2: $(1-\lambda)r^2_{t-1} = 0{,}06 \times 9 = 0{,}54$
  - pasul 3: $\sigma^2_t = 0{,}94 + 0{,}54 = 1{,}48$, deci $\sigma_t = 1{,}217\%$ pe zi, $1{,}217 \times \sqrt{252} = 19{,}3\%$ pe an
- Dacă ziua de azi este liniștită ($r_t = 0$): $\sigma^2_{t+1} = 0{,}94 \times 1{,}48 = 1{,}39$
  - șocul se atenuează cu 6% pe zi; nu dispare niciodată brusc
- O fereastră de 63 de zile ar adăuga azi $9/63 = 0{,}14$ la varianță și ar elimina dintr-odată această contribuție după 63 de zile

Ghost effect: martie–iunie 2020



- S&P 500: volatilitatea pe fereastra mobilă de 63 de zile și volatilitatea EWMA ($\lambda = 0{,}94$), anualizate; barele: $|r_t|$ pe aceeași scală anuală
- **Întrebare pentru sală**: de ce scade brusc linia albastră la mijlocul lui iunie 2020, într-o zi liniștită?

SFM_ch8_ewma

Interpretarea: ghost effect și ponderarea EWMA

- **Răspuns**: pe 15 iunie 2020, randamentul din 16 martie 2020 ($-12{,}8\%$) a ieșit din fereastra de 63 de zile
  - volatilitatea pe fereastra mobilă a scăzut de la 50,4% la 43,0% într-o singură zi, fără nicio știre
  - acest artefact este **ghost effect**: un șoc vechi iese brusc din fereastră
- EWMA reacționează mai repede și uită treptat
  - vîrful EWMA: 84,8% pe 24 martie 2020; vîrful pe 63 de zile: 62,6% abia pe 20 mai 2020
  - în ziua ghost effect, EWMA s-a modificat cu doar $-0{,}9$ puncte procentuale
- Costul reacției rapide: EWMA este mai zgomotos decît o fereastră lungă și nu are un nivel pe termen lung spre care să revină

SFM_ch8_ewma

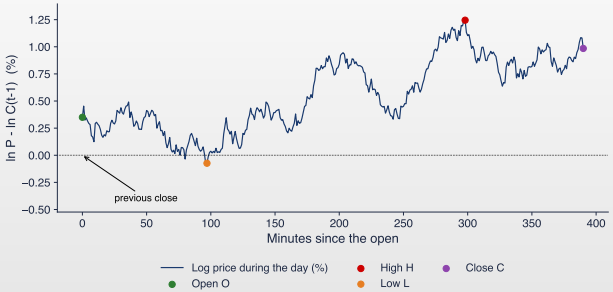
EWMA și GARCH: trimitere la Capitolul 9

- **ARCH** Engle (1982) și **GARCH** Bollerslev (1986): $\sigma_t^2 = \omega + \alpha\, r_{t-1}^2 + \beta\, \sigma_{t-1}^2$
  - $\omega > 0$: o constantă; α: reacția la randamentul de ieri la pătrat; β: ponderea variației de ieri
  - EWMA este cazul particular $\omega = 0$, $\alpha = 1 - \lambda$, $\beta = \lambda$, deci $\alpha + \beta = 1$ (GARCH „integrat", **IGARCH**)
- Consecințe pentru prognozele făcute în ziua t pentru ziua $t + h$
  - EWMA: constantă, $E_t[\sigma_{t+h}^2] = \sigma_{t+1}^2$ pentru orice h (E_t: media condiționată de informația din ziua t): nivelul de azi persistă la nesfîrșit
  - GARCH cu $\alpha + \beta < 1$: atrasă spre varianța pe termen lung $\omega/(1 - \alpha - \beta)$
- Capitolul 9 estimează ω, α, β prin verosimilitate maximă; aici λ este fixat sau ales după un criteriu de prognoză (secțiunea 8)

Prețurile de deschidere, maxim, minim și închidere

- Datele **OHLC**: pentru fiecare zi t, deschiderea O_t, maximul H_t, minimul L_t și închiderea C_t
  - randamentul închidere–închidere folosește 2 prețuri pe zi; H_t și L_t conțin informație despre întreaga traiectorie a zilei
- Notații (diferențe logaritmice în %), fiecare față de deschiderea zilei t:
  - $u_t = \ln(H_t/O_t)$, $d_t = \ln(L_t/O_t)$, $c_t = \ln(C_t/O_t)$; deci $d_t \leq 0 \leq u_t$ și $d_t \leq c_t \leq u_t$
  - randamentul **overnight** $o_t = \ln(O_t/C_{t-1})$; randamentul zilnic este $r_t = o_t + c_t$
  - **amplitudinea** (range) $u_t - d_t = \ln(H_t/L_t)$
- Date: S&P 500 din 2008, DAX din 2006, Bitcoin din 2014, acțiuni BVB din 2010; fișierul BET conține doar închideri

O zi de tranzacționare în datele OHLC



- Preț logaritmic simulat (mișcare browniană, 390 de pași de un minut) după un salt overnight de 0,35%: închiderea păstrează doar punctul final; maximul și minimul rezumă traiectoria
- Două zile cu aceeași închidere pot avea amplitudini foarte diferite, deci riscuri foarte diferite

SFM_ch8_range_estimators

Estimatorul Parkinson (1980)

- Pentru o mișcare browniană fără tendință (drift), cu varianța σ^2 pe zi: $E[(u_t - d_t)^2] = 4 \ln 2 \, \sigma^2$ Parkinson (1980)
  - în cuvinte: pătratul amplitudinii logaritmice, rescalat astfel încît, în medie, să fie egal cu varianța
  - o zi: $\hat{\sigma}^2_{P,t} = \dfrac{(u_t - d_t)^2}{4 \ln 2}$; pe n zile: media acestor valori
  - $4 \ln 2 \approx 2{,}77$: amplitudinea este, în medie, mai mare decît σ
- Ipoteze
  - tranzacționare continuă în timpul zilei, deci maximul și minimul observate sînt extremele reale
  - fără tendință și fără salt overnight: amplitudinea reflectă doar ședința de tranzacționare
- Mult mai precis decît r_t^2: o simulare din această secțiune măsoară cu cît
  - logaritmul amplitudinii se folosește și la estimarea modelelor de volatilitate stochastică Alizadeh, Brandt & Diebold (2002)

Garman–Klass (1980) și Rogers–Satchell (1991)

- **Garman–Klass** Garman & Klass (1980): adaugă randamentul deschidere–închidere
  - $\hat{\sigma}^2_{GK,t} = 0{,}5\,(u_t - d_t)^2 - (2\ln 2 - 1)\,c_t^2$
  - ponderile fac estimatorul nedeplasat și aproape de varianță minimă printre astfel de forme pătratice, în absența tendinței
- **Rogers–Satchell** Rogers & Satchell (1991): nedeplasat pentru orice tendință μ
  - $\hat{\sigma}^2_{RS,t} = u_t(u_t - c_t) + d_t(d_t - c_t)$
  - util pentru active cu o tendință puternică pe fereastră (de exemplu, Bitcoin într-o perioadă de creștere puternică)
- Ambele ignoră însă randamentul overnight o_t: măsoară doar varianța ședinței de tranzacționare

Yang–Zhang (2000): includerea randamentului overnight

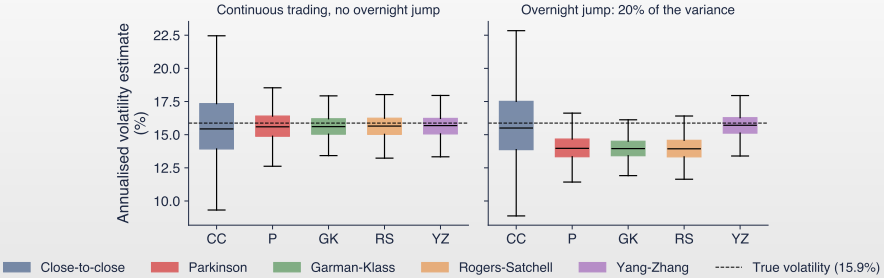
- Pe o fereastră de n zile Yang & Zhang (2000):
  - $\hat{\sigma}^2_{YZ} = \hat{\sigma}^2_O + k\,\hat{\sigma}^2_C + (1-k)\,\hat{\sigma}^2_{RS}, \quad k = \dfrac{0{,}34}{1{,}34 + (n+1)/(n-1)}$
  - $\hat{\sigma}^2_O$, $\hat{\sigma}^2_C$: variațele de selecție ale randamentelor overnight o_t și ale randamentelor deschidere–închidere c_t; $\hat{\sigma}^2_{RS}$: media termenului Rogers–Satchell
- Proprietăți
  - nedeplasat cu tendință și cu salturi la deschidere; k minimizează varianța estimatorului
  - $n = 21$: $k = 0{,}34/(1{,}34 + 22/20) = 0{,}34/2{,}44 = 0{,}14$
- Cere o fereastră de $n \geq 2$ zile (variațe de selecție); Parkinson, Garman–Klass și Rogers–Satchell funcționează și pentru o singură zi

Exemplu rezolvat: S&P 500 pe 16 martie 2020

- Închiderea precedentă 2 711; deschiderea 2 509, maximul 2 563, minimul 2 381, închiderea 2 386
  - $o = -7{,}760$, $u = 2{,}145$, $d = -5{,}223$, $c = -5{,}005$ (în %); închidere–închidere $r = o + c = -12{,}765$
- Estimări ale variației pentru o zi ($\%^2$)
  - închidere–închidere: $r^2 = 163{,}0$
  - Parkinson: $(7{,}368)^2/2{,}773 = 19{,}6$; Garman–Klass: 17,5; Rogers–Satchell: 16,5
  - partea overnight $o^2 = 60{,}2$; overnight plus Rogers–Satchell: 76,7
- Cea mai mare parte a scăderii s-a produs overnight, înainte de deschidere: estimatorii care ignoră o_t nu surprind cea mai mare parte a riscului din această zi

SFM_ch8_range_estimators

Eficiența estimatorilor prin simulare



- 1000 de ferestre de 21 de zile; volatilitatea reală 1% pe zi (15,9% pe an); mișcare browniană intrazilnică pe 4680 de pași (unul la 5 secunde)
- **Eficiența** unui estimator: $\mathrm{Var}(\hat{\sigma}^2_{CC})/\mathrm{Var}(\hat{\sigma}^2)$, CC fiind estimatorul închidere–închidere; panoul din dreapta: 20% din varianța zilnică apare overnight

SFM_ch8_range_efficiency

Interpretarea simulării: precizie și deplasare

	CC	Parkinson	GK	RS	YZ
Eficiența, tranzacționare continuă	1,0	5,4	9,3	7,6	8,5
Deplasarea (%), tranzacționare continuă	0	−2	−3	−3	−2
Deplasarea (%), salt overnight	1	−22	−22	−22	−2
Raportul MSE, salt overnight	1,0	1,8	1,8	1,8	9,0
Deplasarea (%), 26 de prețuri pe zi	0	−23	−32	−32	−28

- Cu tranzacționare continuă, estimatorii de amplitudine sînt de 5–9 ori mai eficienți: o lună de amplitudini conține tot atîta informație cît 5–9 luni de închideri
- Cu salt overnight, Parkinson, GK și RS omit varianța nopții (deplasare de circa −22%); doar YZ rămîne nedeplasat; raportul **MSE** (mean squared error, eroarea pătratică medie) = $\mathrm{MSE}(CC)/\mathrm{MSE}$
- Cu puține tranzacții amplitudinea observată este prea mică: toți estimatorii de amplitudine subestimează (piețe puțin lichide) Molnár (2012)

SFM_ch8_range_efficiency

Estimatorii volatilității pentru cinci active

	Din	CC	P	GK	RS	YZ	Ponderea nopții	$\mathrm{corr}(o, c)$
S&P 500	2008	19,9	15,4	14,6	14,4	16,3	13%	0,22
DAX	2006	20,9	16,7	16,6	16,6	20,3	31%	0,02
Bitcoin	2014	66,6	63,5	62,4	62,5	63,2	0%	0,02
TLV	2010	26,6	21,5	20,6	21,0	25,6	26%	−0,08
SNP	2010	24,5	21,4	20,8	21,4	25,8	27%	−0,20

- Volatilitatea anualizată (%) pe întregul eșantion OHLC; P = Parkinson; ponderea nopții $= \hat{\sigma}_O^2/(\hat{\sigma}_O^2 + \hat{\sigma}_C^2)$
- Acțiuni: Banca Transilvania (TLV) și OMV Petrom (SNP); deschiderea, maximul și minimul ajustate cu același factor ca închiderea; zilele fără tranzacții sînt eliminate

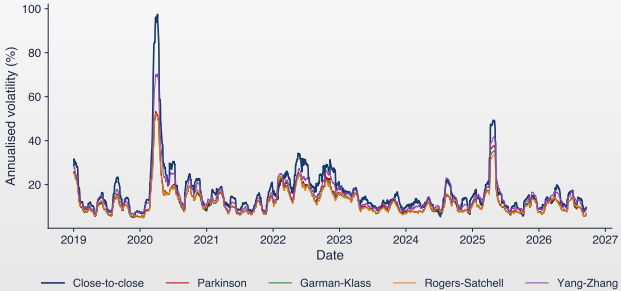
SFM_ch8_range_estimators

Interpretarea diferențelor dintre estimatori

- Bitcoin se tranzacționează 24 de ore din 24: ponderea nopții este aproape 0, iar toți cei cinci estimatori sînt apropiați
  - diferența rămasă: traiectoria zilei este observată prin tranzacții discrete, nu continuu
- DAX și acțiunile BVB: între un sfert și o treime din varianță apare overnight
  - Parkinson, GK și RS sînt mult sub CC; YZ, care include varianța nopții, este apropiat de CC
- S&P 500: chiar și YZ (16,3%) rămîne sub CC (19,9%)
  - randamentele overnight și cele deschidere–închidere sînt corelate pozitiv (0,22); o explicație probabilă: deschiderea indicelui folosește prețuri stale (neactualizate) ale acțiunilor care încă nu s-au tranzacționat
  - $\mathrm{Var}(r) = \mathrm{Var}(o) + \mathrm{Var}(c) + 2\,\mathrm{Cov}(o, c)$: niciun estimator de amplitudine nu vede termenul de covarianță
- Lecția: verificați ipotezele pe date înainte de a avea încredere într-un cîștig de eficiență

SFM_ch8_range_estimators

S&P 500: estimări pe 21 de zile din 2019



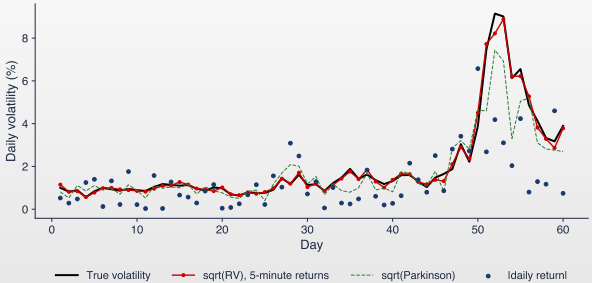
- Volatilitatea pe 21 de zile, anualizată, calculată cu cei cinci estimatori; medii din 2019: CC 16,3%, Parkinson 12,5%, YZ 14,7%
- Liniile estimatorilor de amplitudine sînt mai netede: variația zilnică tipică a estimării logaritmice este 4,4% pentru Parkinson și 4,4% pentru YZ, față de 7,0% pentru CC

SFM_ch8_range_estimators

Varianța realizată din randamentele intrazilnice

- Împărțim ziua t în M intervale (de exemplu $M = 78$ de randamente de cinci minute $r_{t,i}$ într-o ședință de 6,5 ore)
  - $r_{t,i}$: randamentul logaritmic al intervalului i din ziua t, $i = 1, \ldots, M$
  - **varianța realizată** $\mathrm{RV}_t = \sum_{i=1}^{M} r_{t,i}^2$; **volatilitatea realizată** $= \sqrt{\mathrm{RV}_t}$
  - cînd $M \to \infty$, RV_t converge către **varianța integrată** a zilei t, adică varianța reală a zilei Andersen, Bollerslev, Diebold & Labys (2003), Barndorff-Nielsen & Shephard (2002)
- Volatilitatea devine aproape observabilă: o măsură aproape fără eroare pentru fiecare zi
  - se folosește la evaluarea prognozelor (secțiunea 8) și la prognoza directă a volatilității, de exemplu cu modelul HAR Corsi (2009)
  - prețuri Bitcoin de înaltă frecvență și riscul zilnic: Pele & Mazurencu-Marinescu-Pele (2019)
- Datele cursului sînt zilnice; următoarele două slide-uri folosesc prețuri simulate la un minut pentru a arăta comportamentul RV

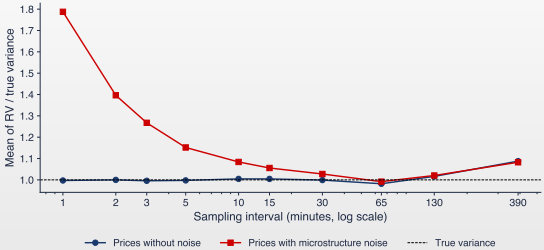
Exemplu simulat: RV comparată cu volatilitatea reală



- 60 de zile simulate, fiecare cu 390 de randamente de un minut; logaritmul volatilității zilnice urmează un proces AR(1) (persistența 0,9)
- Eroarea relativă absolută medie: RV la 5 minute 6%, Parkinson 23%, $|r_t|$ 58%; corelația cu valoarea reală: 0,995; 0,94 și 0,59

SFM_ch8_realised_volatility

Zgomotul de microstructură și signature plot



- Prețurile observate = prețurile reale + **zgomot de microstructură** (oscilația între bid și ask, pasul de cotare), aici cu abaterea standard 0,025%; fiecare randament la pătrat crește atunci, în medie, cu dublul variației zgomotului
- Cu zgomot, RV la un minut supraestimează varianța de 1,79 ori; la 5 minute: 1,15; **signature plot** arată RV în funcție de intervalul de eșantionare
- În practică, RV la 5 minute este greu de depășit Liu, Patton & Sheppard (2015)

SFM_ch8_realised_volatility

1982: Engle și modelul ARCH

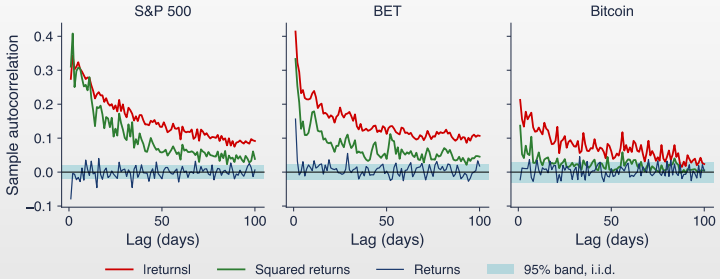
- **Volatility clustering**: perioadele cu randamente mari în valoare absolută alternează cu perioade cu randamente mici Mandelbrot (1963), Cont (2001)
  - randamentele sînt aproape necorelate (Capitolul 7), dar $|r_t|$ și r_t^2 sînt puternic autocorelate
- Robert F. Engle Engle (1982): varianța de azi depinde de șocurile la pătrat de ieri: **ARCH**
  - primul model în care σ_t variază în timp într-un mod previzibil
  - Premiul Nobel pentru economie (2003), împărțit cu Clive Granger Nobel Prize (2003); prelegerea Nobel a lui Engle: Engle (2004)
- Această secțiune: testele pentru volatility clustering; Capitolul 9: modelele



Robert F. Engle (n. 1942), Santiago, 2017

Foto: Econterms (2017); CC BY-SA 4.0; Wikimedia Commons

ACF-ul randamentelor, al pătratelor și al valorilor absolute



- ACF de selecție (funcția de autocorelație, Capitolul 7) pentru r_t, r_t^2 și $|r_t|$, lagurile 1–100, cu banda de 95% $\pm 1{,}96/\sqrt{T}$ pentru date i.i.d.
- S&P 500: toate cele 100 de autocorelații ale lui r_t^2 și $|r_t|$ sînt deasupra benzii; la lagul 100, $|r_t|$ are încă 0,09

SFM_ch8_clustering

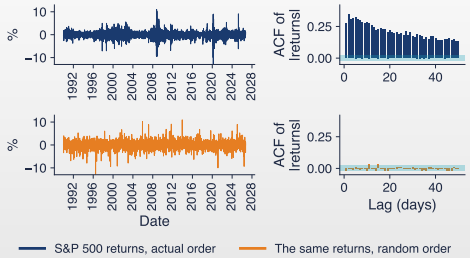
Interpretarea volatility clustering pe trei piețe

	$\hat{\rho}_1(\lvert r\rvert)$	$\hat{\rho}_{10}(\lvert r\rvert)$	$\hat{\rho}_{100}(\lvert r\rvert)$	$\hat{\rho}_1(r^2)$	$\hat{\rho}_{10}(r^2)$	**Laguri în afara benzii,** $\lvert r\rvert$
S&P 500	0,27	0,28	0,09	0,31	0,24	100/100
BET	0,41	0,23	0,11	0,33	0,18	100/100
Bitcoin	0,21	0,12	0,02	0,14	0,04	92/100

- Scădere lentă: un șoc de volatilitate se vede încă după luni de zile; Capitolul 11 numește acest fenomen memorie lungă
- **Efectul Taylor**: $|r_t|$ este mai autocorelat decît r_t^2 (96% dintre laguri la S&P 500, toate lagurile la BET și Bitcoin) Ding, Granger & Engle (1993)
- Bitcoin: volatility clustering mai slab în r_t^2 (26 din 100 de laguri în afara benzii): cîteva zile cu variații extreme domină pătratele

SFM_ch8_clustering

Un test de permutare pentru volatility clustering



- Aceleași randamente S&P 500 într-o ordine aleatoare: aceeași distribuție și aceeași boltire (excesul de boltire 10,9), dar fără volatility clustering
- $\hat{\rho}_{\mathbf{1}}(|r|)$: 0,27 în ordinea reală, $-0{,}02$ după amestecare; Ljung–Box $Q(10)$ pentru $r_t^{\mathbf{2}}$: 8 015 față de 6,3
- Volatility clustering este o proprietate a **ordinii** randamentelor, nu a distribuției lor

SFM_ch8_clustering

Testarea volatility clustering: Ljung–Box pe randamentele la pătrat

- H_0: fără volatility clustering, $\mathrm{Var}(r_t \mid \mathcal{F}_{t-1})$ constantă, deci r_t^2 este necorelat
  - **Testul McLeod–Li** McLeod & Li (1983): statistica Ljung–Box Ljung & Box (1978) aplicată lui r_t^2
  - $Q(m) = T(T+2)\sum_{k=1}^{m} \hat{\rho}_k(r^2)^2/(T-k) \sim \chi^2(m)$ în ipoteza H_0: o sumă ponderată a pătratelor primelor m autocorelații
  - T: numărul de randamente; m: numărul de laguri testate; $\hat{\rho}_k(r^2)$: autocorelația de selecție a lui r_t^2 la lagul k
  - $Q(m) \geq 0$; valorile mari indică pătrate autocorelate; respingem dacă $Q(m)$ depășește cuantila de 95% $\chi^2_{0,95}(m)$
- **Exemplu rezolvat**: S&P 500, $T = 9\,240$, $m = 3$
  - $\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3$ pentru r_t^2: 0,311; 0,408; 0,250
  - termenii $T(T+2)\hat{\rho}_k^2/(T-k)$: 892; 1 540; 579; $Q(3) = 3\,012 \gg \chi^2_{0,95}(3) = 7{,}81$: respingem H_0
- Același test aplicat lui r_t (Capitolul 7) pune o altă întrebare: este **media** previzibilă?

SFM_ch8_clustering

Testul ARCH-LM (Engle, 1982)

- Regresia auxiliară a randamentelor centrate la pătrat e_t^2, $e_t = r_t - \bar{r}$, pe q dintre propriile lor laguri Engle (1982):
  - $e_t^2 = c_0 + c_1 e_{t-1}^2 + \cdots + c_q e_{t-q}^2 + u_t$; H_0: $c_1 = \cdots = c_q = 0$
  - $\bar{r}$: media de selecție; $c_0, \ldots, c_q$: coeficienții regresiei; u_t: eroarea
  - statistica **LM** (Lagrange multiplier, multiplicatorul Lagrange): $\mathrm{LM} = nR^2 \sim \chi^2(q)$, n = numărul de observații din regresie
- S&P 500, $q = 5$: $R^2 = 0{,}241$, $n = 9\,235$
  - $R^2 \in [0, 1]$: proporția din varianța lui e_t^2 explicată de propriul trecut
  - $\mathrm{LM} = 9\,235 \times 0{,}241 = 2\,227 \gg \chi^2_{0,95}(5) = 11{,}07$: efecte ARCH puternice
- Interpretarea lui R^2: circa un sfert din variația de la o zi la alta a lui e_t^2 poate fi anticipată din ultimele cinci zile
- Ambele teste sînt instrumente de diagnostic: arată că volatility clustering există, nu cum îl modelăm (Capitolul 9)

SFM_ch8_clustering

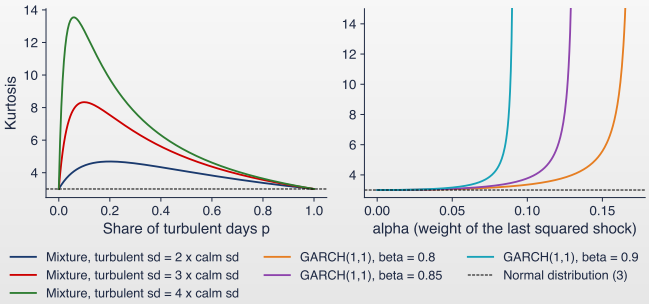
Testele de volatility clustering pe șase serii

	T	$\hat{\rho}_1(r)$	$Q(10)$, r	$\hat{\rho}_1(r^2)$	$Q(10)$, r^2	$Q(10)$, $\lvert r\rvert$	ARCH-LM(5)	Exc. boltire
S&P 500	9 240	−0,079	90,8	0,311	8 015	8 323	2 227	10,9
DAX	9 288	−0,009	21,8	0,155	3 747	5 794	1 209	6,0
BET	7 243	0,156	198,3	0,333	2 500	5 184	1 057	9,9
Bitcoin	4 384	−0,022	20,2	0,137	213	1 086	115	12,0
TLV	3 788	0,010	29,1	0,122	481	1 179	253	13,5
SNP	3 815	−0,015	20,7	0,195	505	944	234	8,9

- Valori critice de 5%: $\chi^2_{0,95}(10) = 18{,}31$ pentru $Q(10)$, $\chi^2_{0,95}(5) = 11{,}07$ pentru ARCH-LM(5); toate statisticile pentru r^2 și $|r|$ resping, cu $p < 0{,}001$
- $Q(10)$ pentru r respinge și el la 5% în toate cele șase serii ($p \leq 0{,}028$), dar este mult mai mic decît pentru r^2 (raportul dintre cele două este între 10,5 și 171,7): previzibilitatea mediei este slabă (Capitolul 7), cea a mărimii este puternică peste tot
- Bitcoin, Banca Transilvania (TLV) și OMV Petrom (SNP) au cele mai mici valori ARCH-LM, deși au cozi groase: $\mathrm{LM} = nR^2$ crește cu lungimea eșantionului, iar aceste serii sînt mai scurte; și R^2 este mai mic (0,026; 0,067; 0,061 față de 0,241 la S&P 500): cîteva zile extreme domină regresia și reduc R^2

SFM_ch8_clustering

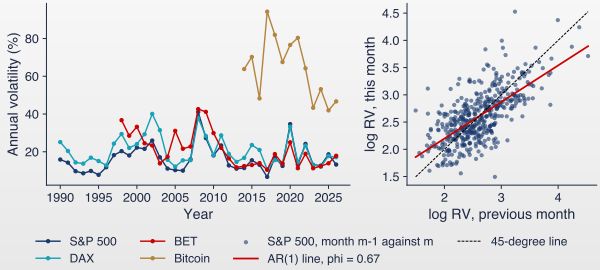
Volatility clustering și cozile groase



- Stînga: zile liniștite cu abaterea standard 1, zile agitate cu abaterea standard 2 (ponderea $p = 0{,}2$): $E[\sigma^2] = 1{,}6$, $E[\sigma^4] = 4{,}0$, coeficientul de boltire $3E[\sigma^4]/E[\sigma^2]^2 = 4{,}69$, deși randamentul fiecărei zile urmează distribuția Normală
- Dreapta: coeficientul de boltire al unui GARCH(1,1), $3 + 6\alpha^2/(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)$, ca în SFEkurgarch; $\alpha = 0{,}10$, $\beta = 0{,}85$: 3,77 (Capitolul 9)

SFM_ch8_kurgarch

Volatilitatea în trei decenii



- Stînga: volatilitatea fiecărui an calendaristic (anualizată, %); dreapta: logaritmul volatilității realizate lunare a S&P 500 (din randamente zilnice) față de luna precedentă
- S&P 500: de la 7% (2017) la 41% (2008); BET: de la 10% (2017) la 43% (2008); Bitcoin: de la 42% (2025) la 94% (2017)

SFM_ch8_long_term_vix

Interpretarea persistenței și a revenirii la medie

- Volatilitatea se schimbă în timp, dar nu se îndepărtează la nesfîrșit: revine la un nivel normal Schwert (1989)
  - AR(1) pentru logaritmul volatilității realizate lunare a S&P 500 (441 de luni): $\hat{\phi} = 0{,}67$ (SE 0,04)
  - $\hat{\phi}$: ponderea estimată a logaritmului volatilității din luna precedentă; o valoare apropiată de 1 înseamnă o revenire lentă la medie
  - timpul de înjumătățire al unui șoc de volatilitate: $\ln 0{,}5 / \ln \hat{\phi} = 1{,}7$ luni; nivelul pe termen lung circa 13,5%
- Volatilitatea lunară are asimetrie la dreapta (coeficientul de asimetrie 3,3), logaritmul ei mult mai puțin (0,6) Andersen, Bollerslev, Diebold & Ebens (2001)
  - media 15,3% este peste mediana 12,9%: cîteva luni de criză ridică media
  - de aceea modelele volatilității sînt adesea scrise pentru $\ln \sigma_t$
- Revenirea la medie este exact ce îi lipsește modelului EWMA și ce are GARCH cu $\alpha + \beta < 1$ (Capitolul 9)

SFM_ch8_long_term_vix

VIX: volatilitatea implicită

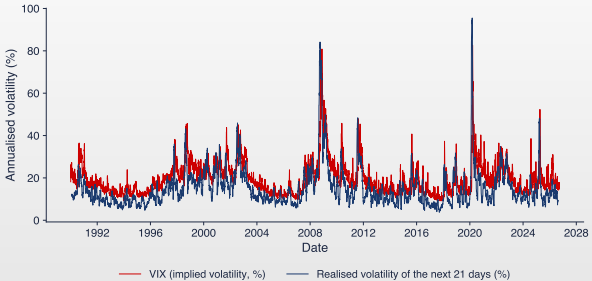
- Opțiunile se tranzacționează la bursă din 1973, cînd membrii bursei Chicago Board of Trade au înființat **CBOE** (Chicago Board Options Exchange)
  - prețul unei opțiuni depinde de volatilitatea pe care piața o așteaptă pînă la scadență
  - **volatilitatea implicită**: volatilitatea care, introdusă într-o formulă de evaluare a opțiunilor, dă prețul observat
- **VIX**, introdus în 1993 Whaley (1993): volatilitatea așteptată a S&P 500 pentru următoarele 30 de zile calendaristice, din prețurile opțiunilor pe S&P 500 Cboe (2026)
  - exprimat în procente anualizate: VIX $= 20$ înseamnă circa $20/\sqrt{12} \approx 5{,}8\%$ pe o lună
  - numit „indicele fricii": crește brusc cînd acțiunile scad Whaley (2009)



Agenți de bursă la Chicago Board of Trade, 31 mai 1973

Foto: U.S. Department of Agriculture (1973); domeniu public; Wikimedia Commons

VIX față de volatilitatea care a urmat



- VIX și volatilitatea realizată a S&P 500 în următoarele 21 de zile de tranzacționare, $\sqrt{(A/21)\sum_{j=1}^{21} r_{t+j}^2}$; maximul VIX: 82,7 pe 16 martie 2020
- **Întrebare pentru sală**: linia roșie este de obicei deasupra sau sub cea albastră?

SFM_ch8_long_term_vix

Interpretarea relației dintre volatilitatea implicită și cea realizată

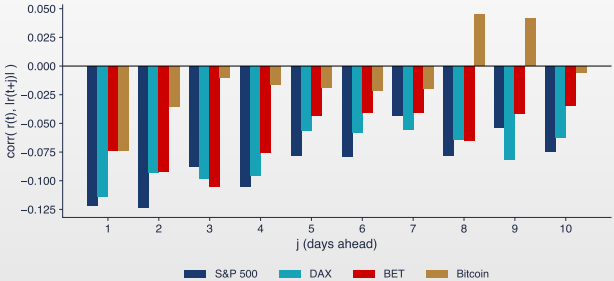
- **Răspuns**: deasupra; VIX a depășit volatilitatea realizată din următoarele 21 de zile în 86% din cele 9 199 de zile
  - media VIX 19,4% față de media volatilității realizate 15,3%: o diferență de 4,1 puncte
  - **prima de risc a variației**: investitorii plătesc pentru a fi asigurați împotriva salturilor de volatilitate Carr & Wu (2009), Bollerslev, Tauchen & Zhou (2009)
- Ca prognoză, VIX este informativ, dar deplasat
  - regresia volatilității realizate viitoare pe VIX: realizată $= -1{,}72 + 0{,}88\,\text{VIX}$, $R^2 = 0{,}52$
  - volatilitatea realizată din ultimele 21 de zile explică mai puțin: $R^2 = 0{,}44$
- Variațiile zilnice ale logaritmului VIX și randamentele S&P 500: corelație $-0{,}71$: frica crește cînd prețurile scad

SFM_ch8_long_term_vix

Efectul de levier

- **Efectul de levier**: randamentele negative sînt urmate de o volatilitate mai mare decît randamentele pozitive de aceeași mărime Christie (1982)
  - explicația prin levier: scăderea prețului acțiunii crește raportul datorii/capital propriu, deci acțiunea devine mai riscantă
  - explicația prin feedback de volatilitate: o volatilitate așteptată mai mare crește randamentul cerut, deci prețul scade azi French, Schwert & Stambaugh (1987), Bekaert & Wu (2000)
- O măsură simplă: $\mathrm{corr}(r_t, |r_{t+j}|)$ pentru $j = 1, 2, \ldots$
  - j: numărul de zile în avans; corr: corelația de selecție pe toate zilele t
  - negativă: o scădere azi este urmată de o volatilitate mai mare mîine; zero într-un model simetric
- Capitolul 2 a arătat asimetria distribuției; aici asimetria apare în **dinamica** volatilității

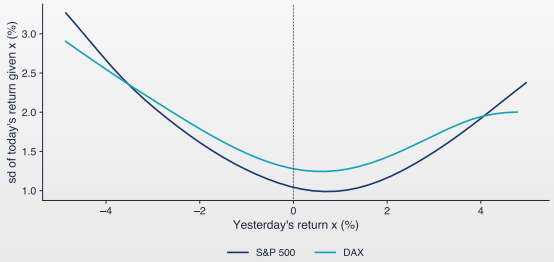
Randamentele negative și volatilitatea viitoare



- $\mathrm{corr}(r_t, |r_{t+j}|)$, $j = 1, \ldots, 10$; S&P 500: $-0{,}122$ la $j = 1$; în sens invers, $\mathrm{corr}(|r_{t-1}|, r_t) = 0{,}025$
- Media lui r^2_{t+1} după o zi de scădere împărțită la cea după o zi de creștere: S&P 500 1,51, DAX 1,46, BET 1,41, Bitcoin 1,06

SFM_ch8_leverage

Volatilitatea neparametrică în funcție de randamentul de ieri



- $\hat{\sigma}(x) = \sqrt{\hat{m}_2(x) - \hat{m}_1(x)^2}$, $\hat{m}_1$, $\hat{m}_2$: regresii liniare locale ale lui r_t și r_t^2 pe $r_{t-1} = x$ (nucleu cvartic, lățimea de bandă 0,04 pe randamente zecimale, ca în SFEvolnonparest)
- S&P 500: $\hat{\sigma}(-2\%) = 1{,}61\%$ față de $\hat{\sigma}(+2\%) = 1{,}16\%$; minimul este la $x \approx 0{,}7\%$, nu la 0: un impact asimetric al știrilor

SFM_ch8_leverage

Problema variabilei proxy

- O prognoză h_t a lui σ_t^2 se face la închiderea zilei $t-1$; valoarea reală σ_t^2 nu este observată niciodată
  - comparăm h_t cu o variabilă **proxy** $\hat{\sigma}_t^2$: o mărime măsurabilă cu $E[\hat{\sigma}_t^2 \mid \mathcal{F}_{t-1}] = \sigma_t^2$
- Variabile proxy, de la zgomotoase la precise: r_t^2, un estimator de amplitudine, varianța realizată
  - cu r_t^2 chiar și o prognoză perfectă pare slabă: Andersen & Bollerslev (1998) au arătat că un R^2 mic nu înseamnă un model prost
- **Regresia Mincer–Zarnowitz** Mincer & Zarnowitz (1969): $\hat{\sigma}_t^2 = a + b\, h_t + u_t$
  - a: interceptul; b: panta; u_t: eroarea; o prognoză bună are $a \approx 0$, $b \approx 1$
  - R^2 măsoară ce parte din variația variabilei proxy este explicată de prognoză

Funcții de pierdere: MSE și QLIKE

- Pierderea medie pe zilele de evaluare; o valoare mai mică este mai bună Patton (2011)
  - L: pierderea dintr-o zi; h_t: prognoza variației; $\hat{\sigma}_t^2$: variabila proxy
  - MSE: $L = (\hat{\sigma}_t^2 - h_t)^2$
  - **QLIKE** (quasi-likelihood, cvasi-verosimilitate): $L = \hat{\sigma}_t^2 / h_t + \ln h_t$, adică, pînă la constante, opusul log-verosimilității Normale
- Ambele sînt **robuste**: cu o variabilă proxy zgomotoasă, dar nedeplasată, ordonează prognozele la fel ca varianța reală
  - pierderile calculate pe $\sqrt{h_t}$ sau pe $\ln h_t$ nu sînt robuste și pot prefera prognoza greșită
- **Exemplu rezolvat**: varianța reală și variabila proxy egale cu 1; două prognoze, $h = 0{,}5$ și $h = 2$
  - MSE: $(1 - 0{,}5)^2 = 0{,}25$ față de $(1 - 2)^2 = 1$: MSE penalizează mai mult supraestimarea
  - QLIKE: $1/0{,}5 + (-0{,}693) = 1{,}307$ față de $1/2 + 0{,}693 = 1{,}193$: QLIKE penalizează mai mult subestimarea, eroarea costisitoare în managementul riscului

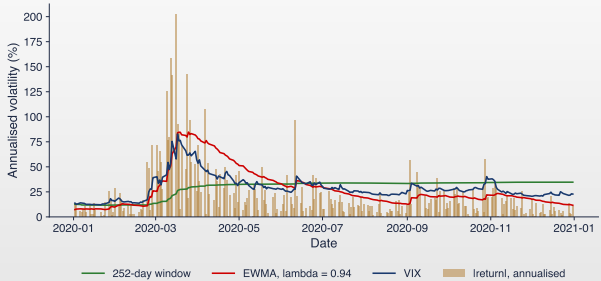
S&P 500: compararea prognozelor

	21 de zile	63 de zile	252 de zile	EWMA 0,94	EWMA 0,97	VIX
MSE	18,99	20,77	21,43	17,77	19,02	16,70
QLIKE	0,876	0,914	1,099	0,808	0,844	0,799
MZ R^2, proxy r^2	0,14	0,05	0,01	0,17	0,11	0,25
MZ R^2, proxy de amplitudine	0,24	0,09	0,02	0,29	0,19	0,42

- Prognoze ale varianței pentru ziua următoare, pentru 4 202 de zile, începînd din 2010; proxy r_t^2; prognoza VIX: VIX^2/A; proxy de amplitudine: o_t^2 plus termenul Garman–Klass
- Testul **DM** (Diebold–Mariano) al egalității QLIKE Diebold & Mariano (1995): EWMA 0,94 față de 21 de zile $t = -4{,}50$ ($p < 0{,}001$); VIX față de EWMA 0,94 $t = -0{,}32$ ($p = 0{,}751$)
- EWMA este semnificativ mai bun decît fereastra de 21 de zile; VIX și EWMA nu diferă semnificativ; fereastra lungă este cea mai slabă; o variabilă proxy precisă crește fiecare R^2, dar nu schimbă ordinea

SFM_ch8_forecast_evaluation

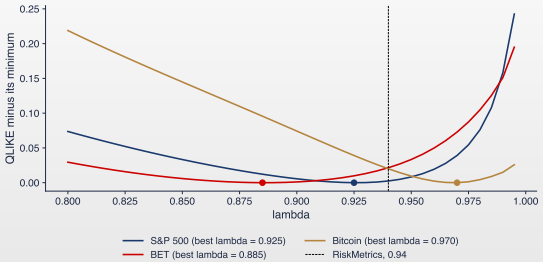
Prognozele din 2020



- Prognozele volatilității S&P 500 pentru ziua următoare în 2020 (anualizate): fereastra de 252 de zile, EWMA ($\lambda = 0{,}94$) și VIX; barele: $|r_t|$
- VIX a început să crească cu o zi de tranzacționare înaintea primei scăderi mari din februarie 2020; EWMA a crescut imediat după ea; fereastra de 252 de zile s-a mișcat lent și a rămas aproape de vîrf pînă la sfîrșitul anului

SFM_ch8_forecast_evaluation

Alegerea lui λ după QLIKE



- QLIKE mediu al prognozelor EWMA pentru ziua următoare (proxy r_t^2, din 2010; Bitcoin din 2016) minus minimul lui, pentru λ între 0,80 și 0,995
- Cel mai bun λ: S&P 500 0,925, BET 0,885, Bitcoin 0,970; valoarea RiskMetrics 0,94 are un QLIKE mai mare cu doar 0,002 la S&P 500
- Un singur λ pentru toate activele este un compromis comod, nu un optim

SFM_ch8_forecast_evaluation

O întrebare de cercetare deschisă

- **Ajută estimatorii de amplitudine la Bursa de Valori București, unde multe acțiuni se tranzacționează rar?**
  - teoria: amplitudinile sînt de 5–9 ori mai eficiente; dar tranzacționarea rară micșorează amplitudinea observată (secțiunea 4)
  - zile cu tranzacții, dar cu închiderea neschimbată, din 2010: TLV 9%, SNP 8%, BRD 10%, TGN 10%
  - raportul volatilităților Parkinson / închidere–închidere: TLV 0,81, SNP 0,87, BRD 0,90, TGN 0,87 (S&P 500: 0,78)
  - QLIKE al prognozelor pentru ziua următoare, cu fereastra de 21 de zile, CC față de YZ: TLV 2,022 față de 2,039; TGN 1,750 față de 1,636
- Întrebarea rămîne deschisă: răspunsul diferă de la o acțiune la alta; contează lichiditatea, licitația de deschidere sau mărimea mișcării overnight?
- Instrumentele AI pot accelera un astfel de studiu, dar rezultatele lui trebuie verificate Wang et al. (2023)

SFM_ch8_bvb_range

Utilizări posibile ale instrumentelor AI

- **Literatura**: o listă a studiilor despre volatilitatea de amplitudine și cea realizată pe piețe emergente și puțin lichide, cu datele și estimatorii folosiți
- **Cod**: o primă versiune a celor cinci estimatori, a prognozelor pe ferestre mobile și a comparațiilor QLIKE pentru toate acțiunile din BET
- **Robustețe**: grupe de lichiditate după volum, alte ferestre, cele două jumătăți 2016–2020 și 2021–2026
- Exemplu de prompt
  - `Write a Python function that computes the Parkinson, Garman-Klass, Rogers-Satchell and Yang-Zhang variance on rolling 21-day windows from daily open, high, low and close prices, and compares their next-day forecasts with QLIKE against squared close-to-close returns.`

Verificări necesare

- Constantele: $4 \ln 2$ la Parkinson, $2 \ln 2 - 1$ la Garman–Klass, k la Yang–Zhang; o primă versiune cu $\ln 2$ sau cu $k = 0{,}34$ este greșită
- Prețurile acțiunilor: deschiderea, maximul și minimul trebuie ajustate pentru dividende și splituri cu același factor ca închiderea
- Anualizarea cu frecvența reală: 365 pentru cripto, circa 250 pentru acțiunile BVB
- Fără informație din viitor: prognoza pentru ziua t folosește doar datele pînă în ziua $t - 1$
- Referințele: fiecare lucrare citată trebuie să existe; verificați DOI-ul

Idee de proiect

- **Întrebarea**: ce estimator zilnic de volatilitate ar trebui să folosească un manager de risc din România?
  - date: acțiuni BVB din 2010 (TLV, SNP, BRD, TGN și altele cu date OHLC), S&P 500 și DAX ca termeni de comparație (EODHD)
- Pași
  - cei cinci estimatori pe ferestre de 21 de zile; variantele EWMA ale fiecăruia; prognoze pentru ziua următoare
  - QLIKE și MSE cu proxy-ul r_t^2; teste Diebold–Mariano; rezultate pe grupe de lichiditate
  - legați cîștigul YZ față de CC de ponderea zilelor fără volum și de ponderea nopții
- Livrabile: un tabel, un grafic și un paragraf despre ce pot și ce nu pot arăta datele
- Declarați orice utilizare a instrumentelor AI și enumerați erorile acestora pe care le-ați corectat

Idei de reținut

- Volatilitatea este abaterea standard a randamentelor; este latentă și variază în timp; se anualizează cu frecvența reală
- La ferestrele istorice, viteza de reacție se plătește cu zgomot; EWMA ($\lambda = 0{,}94$) reacționează repede și evită ghost effect
- Estimatorii de amplitudine folosesc traiectoria zilei: mult mai eficienți, dar deplasați de salturile overnight și de tranzacționarea rară; Yang–Zhang adaugă varianța nopții
- Volatility clustering: $|r_t|$ și r_t^2 sînt autocorelate; se detectează cu Ljung–Box pe r_t^2 și cu ARCH-LM
- Volatilitatea revine la medie; VIX depășește de obicei volatilitatea realizată; scăderile cresc volatilitatea viitoare
- Comparați prognozele cu pierderi robuste (QLIKE) și o variabilă proxy; testați diferențele

Formule de reținut

Mărimea	**Formula**
Volatilitatea anualizată	$\sigma_{\text{an}} = \sigma_{\text{zi}}\sqrt{A}$
EWMA	$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2$, timp de înjumătățire $\ln 0{,}5/\ln\lambda$
Parkinson	$(u_t - d_t)^2/(4\ln 2)$
Garman–Klass	$0{,}5(u_t - d_t)^2 - (2\ln 2 - 1)c_t^2$
Rogers–Satchell	$u_t(u_t - c_t) + d_t(d_t - c_t)$
Yang–Zhang	$\hat\sigma_O^2 + k\hat\sigma_C^2 + (1-k)\hat\sigma_{RS}^2$, $k = 0{,}34/(1{,}34 + (n+1)/(n-1))$
Varianța realizată	$\mathrm{RV}_t = \sum_i r_{t,i}^2$
McLeod–Li	$Q(m) = T(T+2)\sum_{k=1}^m \hat\rho_k(r^2)^2/(T-k) \sim \chi^2(m)$
ARCH-LM	$\mathrm{LM} = nR^2 \sim \chi^2(q)$
MSE, QLIKE	$(\hat\sigma_t^2 - h_t)^2$, $\hat\sigma_t^2/h_t + \ln h_t$

Autoevaluare

- **Întrebare**: o acțiune are $H = 102$, $L = 98$ într-o zi; care este volatilitatea zilnică Parkinson?
  - **Răspuns**: $\ln(102/98) = 4{,}00\%$; $\sqrt{4{,}00^2/2{,}77} = 2{,}40\%$
- **Întrebare**: testul Ljung–Box pe randamente nu respinge, iar pe randamentele la pătrat respinge puternic; ce concluzie trageți?
  - **Răspuns**: media este imprevizibilă, varianța este previzibilă: volatility clustering fără autocorelația randamentelor
- **Întrebare**: de ce nu poate EWMA cu $\lambda = 0{,}94$ să anticipeze revenirea volatilității la nivelul normal?
  - **Răspuns**: $\alpha + \beta = 1$ și $\omega = 0$: prognozele lui sînt constante la orice orizont
- Urmează: Capitolul 9, modelele ARCH și GARCH

Bibliografie (1/3)

- Alizadeh, S., Brandt, M.W., Diebold, F.X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3), 1047–1091.
- Andersen, T.G., Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39(4), 885–905.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1), 43–76.
- Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2), 579–625.
- Barndorff-Nielsen, O.E., Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, 64(2), 253–280.
- Bekaert, G., Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, 13(1), 1–42.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Bollerslev, T., Tauchen, G., Zhou, H. (2009). Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11), 4463–4492.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Carr, P., Wu, L. (2009). Variance risk premiums. *Review of Financial Studies*, 22(3), 1311–1341.
- Cboe Global Markets (2026). Cboe Volatility Index (VIX). Product page and methodology.
- Christie, A.A. (1982). The stochastic behavior of common stock variances: value, leverage and interest rate effects. *Journal of Financial Economics*, 10(4), 407–432.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174–196.

Bibliografie (2/3)

- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Ding, Z., Granger, C.W.J., Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1), 83–106.
- Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Engle, R.F. (2004). Risk and volatility: econometric models and financial practice. *American Economic Review*, 94(3), 405–420.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.
- French, K.R., Schwert, G.W., Stambaugh, R.F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1), 3–29.
- Garman, M.B., Klass, M.J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1), 67–78.
- Hansen, P.R., Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), 873–889.
- J.P. Morgan/Reuters (1996). *RiskMetrics – Technical Document*, 4th ed. New York: Morgan Guaranty Trust Company.
- Liu, L.Y., Patton, A.J., Sheppard, K. (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1), 293–311.
- Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4), 394–419.
- McLeod, A.I., Li, W.K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, 4(4), 269–273.
- Mincer, J.A., Zarnowitz, V. (1969). The evaluation of economic forecasts. In J.A. Mincer (ed.), *Economic Forecasts and Expectations*, 3–46. NBER.

Bibliografie (3/3)

- Molnár, P. (2012). Properties of range-based volatility estimators. *International Review of Financial Analysis*, 23, 20–29.
- Nobel Prize Outreach (2003). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2003: Robert F. Engle III and Clive W.J. Granger. nobelprize.org.
- Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1), 61–65.
- Patton, A.J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246–256.
- Pele, D.T., Lazar, E., Dufour, A. (2017). Information entropy and measures of market risk. *Entropy*, 19(5), 226.
- Pele, D.T., Mazurencu-Marinescu-Pele, M. (2019). Using high-frequency entropy to forecast Bitcoin's daily Value at Risk. *Entropy*, 21(2), 102.
- Poon, S.-H., Granger, C.W.J. (2003). Forecasting volatility in financial markets: a review. *Journal of Economic Literature*, 41(2), 478–539.
- Rogers, L.C.G., Satchell, S.E. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1(4), 504–512.
- Schwert, G.W. (1989). Why does stock market volatility change over time? *Journal of Finance*, 44(5), 1115–1153.
- Tsay, R.S. (2010). *Analysis of Financial Time Series*, 3rd ed. Wiley.
- Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60.
- Whaley, R.E. (1993). Derivatives on market volatility. *Journal of Derivatives*, 1(1), 71–84.
- Whaley, R.E. (2009). Understanding the VIX. *Journal of Portfolio Management*, 35(3), 98–105.
- Yang, D., Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73(3), 477–492.