








Modelling Financial Markets

Seminar 2: Market Efficiency — Hands-on



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

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

From returns to conclusions

- 1. A1: compute a variance ratio and interpret dependence.
- 2. B1: compare the estimate with its classical and robust uncertainty.
- 3. C1: audit whether the code supports the economic conclusion.
- For each result, distinguish the estimated effect, its statistical uncertainty and its economic meaning.
- After A1, B1 and C1: applications to new data, then further methods.

Assessment criteria

- correct derivations
- inference robust to heteroskedasticity and dependence, with the decision at a stated level
- out-of-sample evaluation that uses only the information available at the forecast date
- the AI (Artificial Intelligence) critique documented in `AI_ERRORS.md`

A1: from returns to dependence

- First learn the return notation and the random-walk assumptions on the following pages.
- Your output is a reproducible worksheet, followed by one sentence on the sign of dependence.
- A sample identity may differ from a population identity because the estimators use end terms and divisors that differ.

Notation Used in Every Exercise: Returns and Sample Moments

- **Daily log return** of a series with closing prices P_t

$$r_t = \ln P_t - \ln P_{t-1}, \qquad t = 1, \ldots, T$$

  - T: the number of returns
  - q: a horizon in days; $r_t + r_{t-1} + \cdots + r_{t-q+1}$ is the log return over q days

- Sample mean, deviations and variance

$$\bar{r} = \hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \qquad e_t = r_t - \bar{r}, \qquad \hat{\sigma}^2 = \frac{1}{T-1}\sum_{t=1}^{T} e_t^2$$

  - $\hat{}$ (hat): an estimate computed from the sample; without a hat: the population value
  - $\mathbb{E}[\cdot]$, $\mathrm{Var}(\cdot)$, $\mathrm{Cov}(\cdot,\cdot)$: expectation, variance and covariance of the population

- **i.i.d.**: independent and identically distributed
  - the benchmark of the classical tests
  - daily returns are not i.i.d.: calm years alternate with bursts such as 2008–2009 and 2020 (volatility clustering, Chapter 1; chart: Overview of the Data)

Notation Used in Every Exercise: Autocorrelation

- **Autocorrelation of order** k: the correlation between today's return and the return k days earlier

$$\rho_k = \frac{\mathrm{Cov}(r_t, r_{t-k})}{\mathrm{Var}(r_t)}, \qquad \hat{\rho}_k = \frac{\sum_{t=k+1}^{T} e_t e_{t-k}}{\sum_{t=1}^{T} e_t^2}$$

  - k: the lag in days; $\rho_k \in [-1, 1]$; $\rho_k = 0$ for every $k \geq 1$: no linear predictability from the past

- Standard error of $\hat{\rho}_k$ when $\rho_k = 0$
  - i.i.d. returns: $\mathrm{SE}(\hat{\rho}_k) = 1/\sqrt{T}$, hence the i.i.d. null band $\pm 1.96/\sqrt{T}$
  - robust (volatility clustering allowed, Chapter 1): $\mathrm{SE}(\hat{\rho}_k) = \sqrt{\hat{\tau}_k}$

- Robust variance of $\hat{\rho}_k$ and the robust t statistic

$$\hat{\tau}_k = \frac{\sum_{t=k+1}^{T} e_t^2 e_{t-k}^2}{\left(\sum_{t=1}^{T} e_t^2\right)^2}, \qquad \hat{\delta}_k = T\hat{\tau}_k, \qquad t_k = \frac{\hat{\rho}_k}{\sqrt{\hat{\tau}_k}}$$

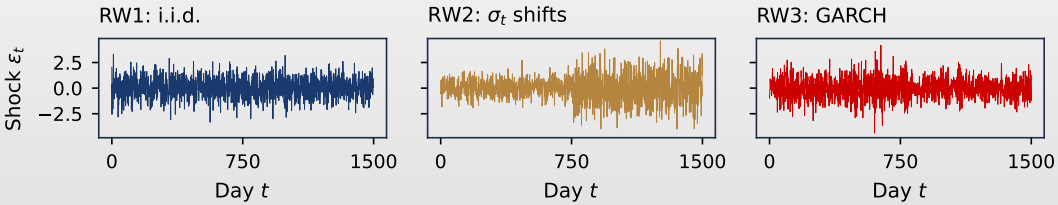
  - $\hat{\delta}_k$: how many times the robust variance exceeds the i.i.d. one
  - $\hat{\delta}_k \approx 1$ for i.i.d. returns; $\hat{\delta}_k > 1$ when large moves follow large moves

Weak-Form Efficiency

- Details: Chapter 2 and Franke, Härdle & Hafner (2019)
- **Efficient market** Fama (1970): prices reflect the available information
  - **weak form**: the information is the past prices and returns
  - **semi-strong form**: all public information (announcements, news; B9)
  - **strong form**: also private information
- Testable consequence of the weak form
  - past returns do not allow risk-adjusted profits after costs
  - in general, the restriction concerns the part of the return above the required return: that excess cannot be forecast from past returns
  - with a constant required return, efficiency implies that $\mathbb{E}[r_t \mid r_{t-1}, r_{t-2}, \dots]$ does not depend on the past: a constant conditional mean
- **Joint-hypothesis problem** Fama (1991): every test of efficiency also tests a model of expected returns
  - a rejection means inefficiency *or* a wrong model of expected returns
  - example (B9): a negative abnormal return after a tax may be an overreaction or a higher risk premium that the market model misses
- Sign of the dependence
  - $\rho_k > 0$: persistence (momentum); $\rho_k < 0$: reversal (mean reversion)

Three Random Walks (1/2)

- **RW1**: Random Walk 1 — independent, identically distributed increments.
- **RW2**: Random Walk 2 — independent increments, possibly with different distributions.
- **RW3**: Random Walk 3 — uncorrelated increments, possibly dependent.
- **GARCH**: Generalized AutoRegressive Conditional Heteroskedasticity.



Three Random Walks (2/2)

- **Random walk** of the log price Campbell, Lo & MacKinlay (1997)

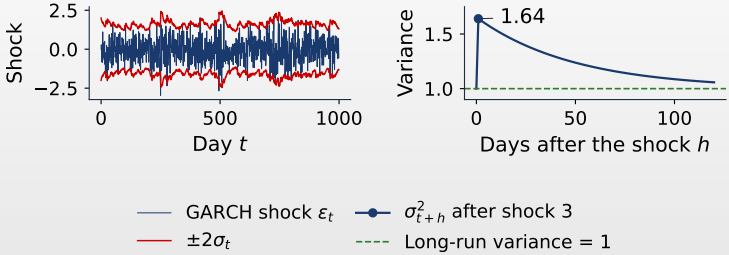
$$\ln P_t = \mu + \ln P_{t-\mathbf{1}} + \varepsilon_t, \qquad \text{so } r_t = \mu + \varepsilon_t$$

  - μ: the drift, the expected daily log return; ε_t: the shock of day t, with mean 0
  - RW1: ε_t i.i.d. (left panel)
  - RW2: ε_t independent, the variance may change (middle panel)
  - RW3: ε_t only uncorrelated, $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$, while ε_t^2 may be autocorrelated (right panel)
  - example of RW3: the GARCH(1,1) model Bollerslev (1986), with $\varepsilon_t = \sigma_t z_t$ (next slide)
- All three panels have $\hat{\rho}_{\mathbf{1}} \approx 0$; the right one has volatility clustering, the relevant null for daily returns is therefore RW3

GARCH(1,1), an Example of RW3

- $\sigma_t^2 = \omega + a\,\varepsilon_{t-1}^2 + b\,\sigma_{t-1}^2$ is the **conditional variance**: the variance of the next shock given the information available today
  - $\omega > 0$, $a, b \geq 0$, $a + b < 1$: the variance stays positive and reverts to its long-run level $\omega/(1 - a - b)$
- Illustrative example: $\omega = 0.02$, $a = 0.08$, $b = 0.90$
  - long-run variance: $0.02/(1 - 0.98) = 1$
  - after a shock $\varepsilon_{t-1} = 3$ on an ordinary day ($\sigma_{t-1}^2 = 1$): $\sigma_t^2 = 0.02 + 0.08 \times 9 + 0.90 \times 1 = 1.64$
- More on GARCH: Chapter 5

GARCH(1,1): a Graphical Illustration



- Left: a simulated GARCH(1,1) with $\omega = 0.02$, $a = 0.08$, $b = 0.90$; horizontal axis the day, vertical axis the shock ε_t
- The band $\pm 2\sigma_t$ widens after large shocks and narrows in calm periods: volatility clustering
- Right: after a shock $\varepsilon_t = 3$ the expected variance jumps from 1 to 1.64, then returns to 1 at the rate $a + b = 0.98$ a day; horizontal axis the days after the shock

RW1, RW2 and RW3 Compared (1/2)

	What the shocks ε_t must satisfy	What is relaxed	Mini-example (RW chart)
RW1	independent and identically distributed (i.i.d.)	–	i.i.d. Normal shocks with constant variance (left panel); a discrete example: $+1\%$ or -1% with probability $1/2$
RW2	independent; the distribution may change over time	identical distribution	$\pm 1\%$ a day in a calm year, $\pm 3\%$ in a turbulent one, fixed in advance: still independent, but the variance changes (middle panel)
RW3	only uncorrelated: $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$	independence	symmetric GARCH: after a $\pm 3\%$ day the variance of the next shock rises, so a large move becomes more likely, but its sign stays unpredictable (right panel)

RW1, RW2 and RW3 Compared (2/2)

- Each step relaxes one requirement: RW1 $\Rightarrow$ RW2 $\Rightarrow$ RW3 (when the variances exist)
- RW2 with a variance fixed in advance (calm year, turbulent year) is independent but not stationary: calm and turbulent periods alone do not prove dependence
- For the tests (Testing VR(q) $= 1$): $z(q)$ is built for RW1, $z^*(q)$ for RW3; RW3 still allows nonlinear predictability of the mean

The Variance Ratio

- Uncorrelated returns with constant variance σ^2: the variance of a q-day sum is $q\sigma^2$
  - variances add up only when the covariances are zero
- **Variance ratio** (VR) Lo & MacKinlay (1988): the variance of the q-day return relative to q one-day variances

$$\mathrm{VR}(q) = \frac{\mathrm{Var}(r_t + r_{t-1} + \cdots + r_{t-q+1})}{q\,\mathrm{Var}(r_t)}$$

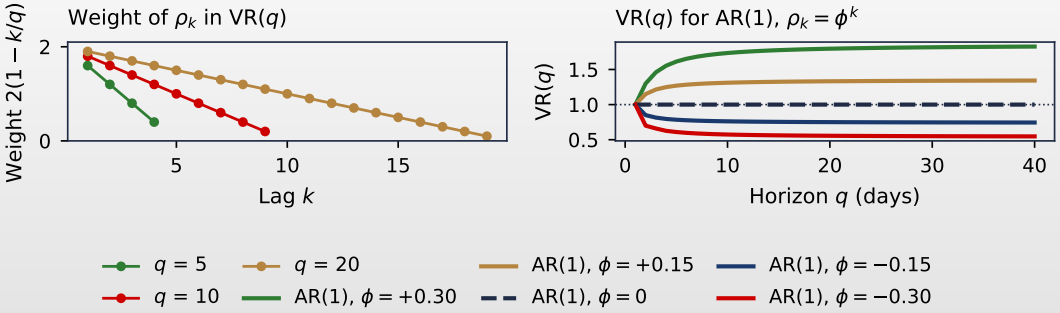
  - population identity: $\mathrm{VR}(q) = 1$ for uncorrelated returns with constant variance; $\mathrm{VR}(1) = 1$ always
  - with time-varying variance (RW2, RW3) this identity can fail; under the conditions of Lo & MacKinlay (1988) the estimator $\widehat{\mathrm{VR}}(q)$ still tends to 1 and the robust $z^*(q)$ (defined below) is asymptotically $N(0,1)$
  - $\mathrm{VR}(q) > 1$: persistence, q-day moves larger than q independent days
  - $\mathrm{VR}(q) < 1$: reversal, part of each move is undone in the following days
- Case $q = 2$

$$\mathrm{Var}(r_t + r_{t-1}) = 2\sigma^2 + 2\mathrm{Cov}(r_t, r_{t-1}) \;\Rightarrow\; \mathrm{VR}(2) = 1 + \rho_1$$

  - example: $\rho_1 = -0.1$ gives $\mathrm{VR}(2) = 0.9$, two-day sums vary 10% less than two separate days

The Variance Ratio as a Weighted Sum (1/2)

AR: AutoRegressive model; AR(p) has order p (in the event study, AR is the abnormal return).



The Variance Ratio as a Weighted Sum (2/2)

- Expanding the variance of the q-day sum: lag k appears in $2(q-k)$ covariances

$$\mathrm{VR}(q) = 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho_k$$

- Left: the weight $2(1-k/q)$ falls linearly with the lag; right: $\mathrm{VR}(q)$ for AR(1) returns
  - AR(1), the first-order autoregression: $r_t = \phi\, r_{t-1} + u_t$, $|\phi| < 1$, u_t an i.i.d. shock; then $\rho_k = \phi^k$
  - right, theoretical curves: as q grows, $\mathrm{VR}(q)$ moves towards $(1+\phi)/(1-\phi)$, above 1 for $\phi > 0$ and below 1 for $\phi < 0$
  - example: $\rho_1 = -0.1$, all other $\rho_k = 0$: $\mathrm{VR}(5) = 1 + 2(1-1/5)(-0.1) = 0.84$
- **What do you think**: if $\mathrm{VR}(3) = 1$ in the population, are returns uncorrelated?
  - **Answer**: no: $\rho_1 = 0.05$ and $\rho_2 = -0.10$ give $\mathrm{VR}(3) = 1 + \frac{4}{3}(0.05) + \frac{2}{3}(-0.10) = 1$
  - $\mathrm{VR}(q)$ tests one weighted combination of the autocorrelations, not each autocorrelation

A1 [Solved]: The Variance Ratio, Line by Line — Task

- **Questions**
  - what does $\widehat{\mathrm{VR}}(2)$ compute?
  - how does it relate to $\hat{\rho}_1$?
- **Data**: eight daily returns in %, in time order: $1.0; -0.5; 0.8; 0.3; -1.2; 0.6; 0.9; -0.4$; no other data; simple estimator, sums of squares divided by the number of terms
- **Tasks**
  1. Compute the mean $\hat{\mu}$, the deviations $e_t = r_t - \hat{\mu}$ and the one-day variance $\hat{\sigma}_1^2 = \frac{1}{8}\sum_t e_t^2$.
  2. Compute the lag-1 autocorrelation $\hat{\rho}_1 = \sum_{t=2}^{8} e_t e_{t-1} / \sum_{t=1}^{8} e_t^2$.
  3. Form the seven overlapping two-day sums $s_t = r_{t-1} + r_t$ and compute $\hat{\sigma}_2^2 = \frac{1}{7}\sum_t (s_t - 2\hat{\mu})^2$.
  4. Compute $\widehat{\mathrm{VR}}(2) = \hat{\sigma}_2^2/(2\hat{\sigma}_1^2)$ and compare it with $1 + \hat{\rho}_1$.
- **Report**: a worksheet with columns t, r_t, e_t, e_t^2, $e_t e_{t-1}$, s_t, $(s_t - 2\hat{\mu})^2$ and one sentence on the sign of $\widehat{\mathrm{VR}}(2) - 1$
- Notebook: section A1

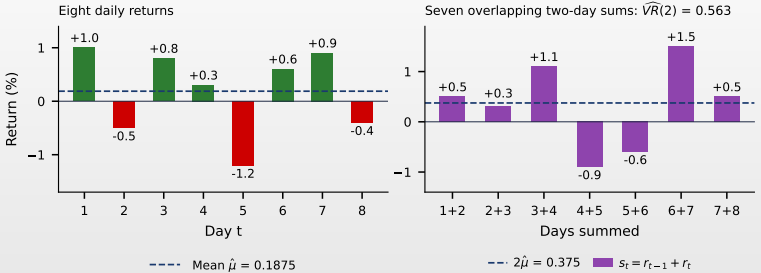
Back Exercise list

A1 [Solved]: Worksheet

t	r_t	e_t	e_t^2	$e_t e_{t-1}$	s_t	$(s_t - 2\hat{\mu})^2$
1	$+1.0$	$+0.8125$	0.6602	–	–	–
2	-0.5	-0.6875	0.4727	-0.5586	$+0.5$	0.0156
3	$+0.8$	$+0.6125$	0.3752	-0.4211	$+0.3$	0.0056
4	$+0.3$	$+0.1125$	0.0127	$+0.0689$	$+1.1$	0.5256
5	-1.2	-1.3875	1.9252	-0.1561	-0.9	1.6256
6	$+0.6$	$+0.4125$	0.1702	-0.5723	-0.6	0.9506
7	$+0.9$	$+0.7125$	0.5077	$+0.2939$	$+1.5$	1.2656
8	-0.4	-0.5875	0.3452	-0.4186	$+0.5$	0.0156
sum			4.4688	-1.7639		4.4044

- $\hat{\mu} = 1.5/8 = 0.1875$; $\hat{\sigma}_1^2 = 4.4688/8 = 0.5586$; $\hat{\rho}_1 = -1.7639/4.4688 = -0.395$
- $\hat{\sigma}_2^2 = 4.4044/7 = 0.6292$; $\widehat{\mathrm{VR}}(2) = 0.6292/(2 \times 0.5586) = 0.563$, against $1 + \hat{\rho}_1 = 0.605$

A1 [Solved]: Results and Interpretation



- $\mathrm{VR}(2) = 1 + \rho_1$ holds for population moments; in a sample the two differ by end terms and divisors (7 sums, 8 squares): 0.563 vs 0.605 is a convention gap, not an error
- Interpretation: consecutive returns reverse, so two-day sums vary less than two single days: mean reversion
  - with $T = 8$ this is arithmetic, not evidence: no test is attempted

MFM_ch2_seminar

Why the Sample VR Identity Has End Terms

- For this illustration use $e_t = r_t - \bar{r}$, $S = \sum e_t^2$, one-day variance S/T, and the mean squared centred two-day sum divided by 2.
- Expanding the squared sums gives $\widehat{VR}(2) = \frac{T}{T-1}\left[1 + \hat{\rho}_1 - \frac{e_1^2 + e_T^2}{2S}\right]$.
- Toy returns $(2, 0, -2)$ give $S = 8$, $\hat{\rho}_1 = 0$ and two-day sums $(2, -2)$: the scaled variance is $4/2 = 2$, so $\widehat{VR}(2) = 2/(8/3) = 0.75$.
- The identity gives $(3/2)[1 - 8/16] = 0.75$, while $1 + \hat{\rho}_1 = 1$. Endpoints and divisors explain the gap; the Lo–MacKinlay estimator corrects the bias through the divisor m.

A1: check before continuing

- ⊡ Can someone reproduce every column from the eight returns, with the same divisors?
- ⊡ Explain why overlapping two-day sums share one daily return.
- ⊡ State the direction of dependence without claiming a profitable trading strategy.

B1: an estimate needs uncertainty

- You now move from eight returns to real market samples.
- Learn the robust standard error and the variance-ratio tests before reading the results.
- Deliver a table that states the market, horizon, statistic, decision and assumptions.

Where the Data Come From

- **S&P 500**: Standard & Poor's 500.
- **BET**: Bucharest Exchange Trading (index).
- **BET-TR**: BET Total Return (index).
- Files: data/market; close for indices, adjusted_close for bank stocks (dividends and splits included).
- S&P 500: GSPC.INDX.csv; BET: BET.csv; BET-TR: BETTR.INDX.csv (first close 23 September 2014).
- Banca Transilvania: TLV.RO.csv; BRD (BRD–Groupe Société Générale, BVB ticker): BRD.RO.csv; Bitcoin: BTC-USD.CC.csv, loaded from 2011 for B6.
- Other series and transformations: the Data section of chapter2_seminar_notebook.ipynb.

Markets and data sources

- **US**: United States.
- **BVB**: Bursa de Valori București — the Bucharest Stock Exchange.
- **WIG20**: Warszawski Indeks Giełdowy 20 — the Warsaw Stock Exchange index of the 20 largest companies.
- **BUX**: Budapesti Értéktőzsde index — the Budapest Stock Exchange index.
- **PX**: Prague Stock Exchange index (index Burzy cenných papírů Praha).
- **BIST**: Borsa İstanbul (BIST 100 index) — the Istanbul Stock Exchange (BIST 100 index).
- **IPC**: Índice de Precios y Cotizaciones — the Mexican Stock Exchange index.
- **FRED**: Federal Reserve Economic Data.

Data Used

Series	Price used	Calendar	Used in
S&P 500	close	US trading days	A3, B1–B3, B4 (supplied p-values), B5, C1
BET	close	BVB trading days	A2, B1, B3, B4 (supplied p-values), B7, C2
BET-TR and 8 emerging indices	close (WIG20, BUX, PX, BIST 100, Bovespa, IPC, Nifty 50, Shanghai)	own calendar each	B7, C2
Euro Stoxx 50	close	own calendar	B4 (p-values supplied)
Bitcoin	close (USD)	every day	B6
Banca Transilvania, BRD	adjusted close (dividends, splits)	common days with BET	B9
S&P 500 monthly	Shiller S&P Composite, S&P 500 closes, SPY dividends; 3-month T-bill (FRED, TB3MS)	monthly averages, month-end labels	B8
49 industry portfolios	Kenneth French Data Library, value-weighted, monthly	month ends	C3

- Daily closing prices of the course data, one series per asset, from 2000 (Bitcoin from 2011 for B6; in the lecture, Bitcoin enters the full-sample tests from 2014) to 18 Sep 2026; daily log returns (Notation: returns) on each series' own calendar
- Cleaning rule for indices: weekdays only; a close equal to the previous close is removed (usually a filled holiday)

Setup: Run and Verify

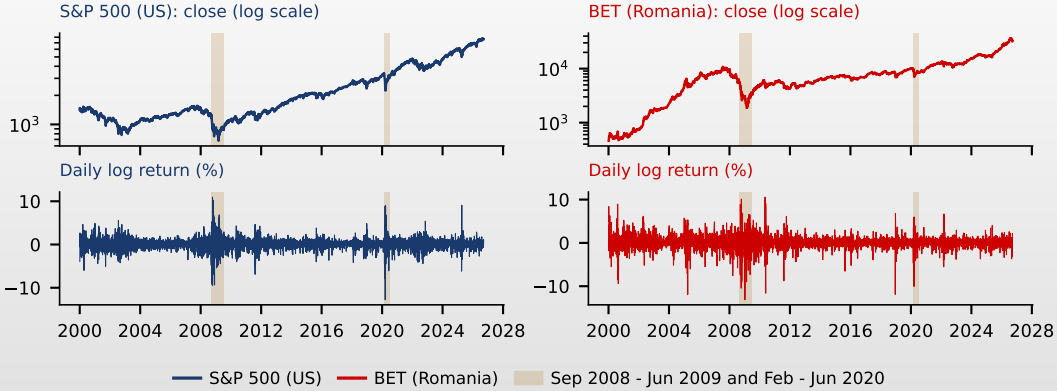
```
# notebook sections "Data" and
# "Efficiency tests" define
#   log_returns(name, start=None)
#   variance_ratio(r, q) -> (VR, z, z*)
#   chow_denning(r), runs_test(r), dfa_hurst(r)
sp = log_returns('sp500')    # S&P 500
bet = log_returns('bet')     # BET
print(len(sp), sp.index[0].date(),
      sp.index[-1].date(), len(bet))
print(sp.head(3).round(5))
```

- Expected: 6 717 S&P 500 returns from 2000-01-04 to 2026-09-18; 6 670 BET returns

date	close	r_t
2000-01-03	1455.22	–
2000-01-04	1399.42	−0.03910
2000-01-05	1402.11	+0.00192
2000-01-06	1403.45	+0.00096

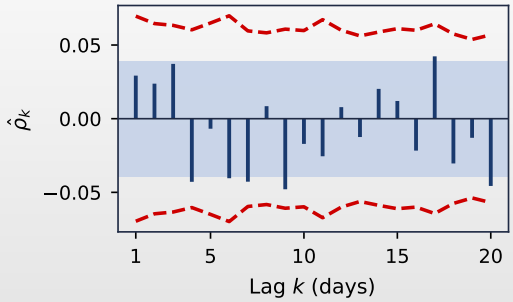
- If your count or first value differs, check the series, the price type (adjusted or unadjusted close) and the cleaning rule

Overview of the Data



MFM_ch2_seminar

Two Null Bands for the Autocorrelation



i.i.d. null band $\pm 1.96/\sqrt{T}$

robust null band $\pm 1.96\sqrt{\hat{\tau}_k}$

$\hat{\rho}_k$ of a GARCH series

- Simulated GARCH returns, $T = 2\,500$: $\rho_k = 0$ at every lag
  - i.i.d. null band $\pm 1.96/\sqrt{T} = \pm 0.039$ (shaded)
  - robust null band $\pm 1.96\sqrt{\hat{\tau}_k}$, on average ± 0.061 (dashed)
- Reading the chart
  - 6 of 20 lags leave the i.i.d. null band: false rejections, since the true ρ_k are all 0
  - no lag leaves the robust null band
  - here $\bar{\delta}_k \approx 2.5$: the robust standard error is $\sqrt{2.5} \approx 1.6$ times larger
- Why
  - for i.i.d. returns $\mathrm{Var}(e_t e_{t-k}) = \sigma^4$ (hence $\mathrm{SE} = 1/\sqrt{T}$); it is larger when e_t^2 and e_{t-k}^2 are positively correlated

The Lo–MacKinlay Estimator

- One-day variance: $\hat{\sigma}_a^2 = \hat{\sigma}^2$, the sample variance of the notation slide on returns, with $\hat{\mu} = \bar{r}$
  - the subscript a marks the one-day estimator
- Variance of the **overlapping** q-day sums, divided by q

$$\hat{\sigma}_c^2(q) = \frac{1}{m} \sum_{t=q}^{T} \Big(\sum_{j=0}^{q-1} r_{t-j} - q\hat{\mu} \Big)^2, \qquad m = q(T-q+1)\Big(1 - \frac{q}{T}\Big)$$

  - the subscript c marks the estimator from overlapping q-day sums
  - overlapping: a new q-day sum starts every day, $T - q + 1$ sums in total
  - $q\hat{\mu}$: the expected q-day return; m: a divisor that removes the small-sample bias
- The estimator

$$\widehat{\mathrm{VR}}(q) = \frac{\hat{\sigma}_c^2(q)}{\hat{\sigma}_a^2}$$

  - in A1 the simple version (divisors T and $T-1$, no bias correction) follows the same idea, step by step

Testing $\mathrm{VR}(q) = 1$

- Under RW1 (i.i.d.): the classical statistic

$$z(q) = \frac{\widehat{\mathrm{VR}}(q) - 1}{\sqrt{2(2q-1)(q-1)/(3qT)}} \approx N(0,1)$$

  - the denominator is the i.i.d. standard error of $\widehat{\mathrm{VR}}(q)$; for $q = 2$ it equals $1/\sqrt{T}$
- Under RW3 (volatility clustering allowed): the robust statistic

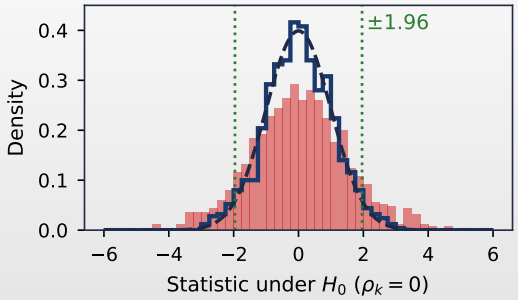
$$z^*(q) = \frac{\sqrt{T}\,(\widehat{\mathrm{VR}}(q) - 1)}{\sqrt{\hat{\theta}(q)}}, \qquad \hat{\theta}(q) = \sum_{k=1}^{q-1} \Big(\frac{2(q-k)}{q}\Big)^2 \hat{\delta}_k$$

  - $\hat{\theta}(q)$: the robust variance of $\sqrt{T}\,\widehat{\mathrm{VR}}(q)$; the weights are those of ρ_k in $\mathrm{VR}(q)$, squared
  - $\hat{\delta}_k = T\hat{\tau}_k$ from the Notation slides; $|z^*(q)| > 1.96$ rejects RW3 at 5%

When $z^*(q)$ Is Valid, and an Example

- When $z^*(q)$ is valid Lo & MacKinlay (1988): RW3 plus three conditions
  - finite fourth moments
  - the dependence between distant days dies out (mixing)
  - $\mathbb{E}[\varepsilon_t^2 \varepsilon_{t-j}\varepsilon_{t-k}] = 0$ for $j \neq k$: then $\hat{\rho}_j$ and $\hat{\rho}_k$ are asymptotically uncorrelated, so their robust variances $\hat{\delta}_k$ add up in $\hat{\theta}(q)$
  - GARCH(1,1) with symmetric shocks satisfies the cross-moment condition; a strong leverage effect (Chapter 1) can violate it
  - finite fourth moments and mixing need extra parameter restrictions: with Normal shocks, finite fourth moments require $(a+b)^2 + 2a^2 < 1$
  - the illustrative GARCH of these slides satisfies them: $(0.08 + 0.90)^2 + 2 \times 0.08^2 = 0.973 < 1$
- Example: $T = 2\,500$, $\widehat{\mathrm{VR}}(2) = 0.95$
  - $z(2) = -0.05 \times 50 = -2.5$: a rejection at 5%
  - with $\hat{\delta}_1 = 4$: $z^*(2) = -2.5/\sqrt{4} = -1.25$, no rejection

Null Distributions of z and z^*



- 1 000 simulated GARCH samples, $T = 2\,500$: no autocorrelation, so H_0 of RW3 is true
  - $z(2)$ is too dispersed (standard deviation 1.55, not 1)
  - $|z(2)| > 1.96$ in 18.4% of the samples instead of 5%
  - $z^*(2)$ is close to $N(0,1)$: 6.2% rejections
- Lesson for B1
  - with volatility clustering a large $|z|$ alone is not evidence: compare $|z^*|$ with 1.96
  - For $q = 2$, $z^* = z/\sqrt{\hat{\delta}_1}$: volatility clustering cuts precision as if the sample were $\hat{\delta}_1$ times smaller

Several Horizons at Once: Chow–Denning

- Testing four horizons separately at 5% rejects *some* horizon more often than 5%
  - the more horizons, the more chances of a false rejection (multiple testing, Multiple Testing: FWER, Bonferroni and Holm)
- **Chow–Denning** statistic Chow & Denning (1993): the largest robust statistic

$$\mathrm{CD} = \max_{q \in \{2,5,10,20\}} |z^*(q)|$$

  - critical value c from $(2\Phi(c) - 1)^4 = 0.95$, Φ the $N(0,1)$ distribution function, i.e. $c = 2.49$ (not 1.96)
  - p-value: $p = 1 - (2\Phi(\mathrm{CD}) - 1)^4$
  - the exponent 4 is the number of horizons: the bound treats them as independent, so the test is slightly conservative
- Reading
  - CD > 2.49: RW3 is rejected jointly at 5%; the horizon with the largest $|z^*|$ shows where the departure from RW3 is strongest

Which variance-ratio horizons do we test?

- Lo & MacKinlay (1988), Table 1a: four horizons, $q = 2, 4, 8, 16$, with weekly observations.
- Here the four test horizons are $q = 2, 5, 10, 20$ days: two days, about one week, two weeks and one trading month for equities.
- The lecture profile extends to $q = 60$ descriptively; it does not change the four horizons tested here.
- Four horizons are a stated design choice, not a guarantee of sufficiency.

B1 [Solved]: Classical vs Robust Variance Ratio — Task

- **Question**: do the S&P 500 and BET reject the random walk once volatility clustering is allowed?
- **Data**: S&P 500 and BET, daily log returns 2000–2026, $T = 6\,717$ and $6\,670$; H_0 of $z(q)$: RW1 (i.i.d.); H_0 of $z^*(q)$: RW3 plus the conditions of Lo & MacKinlay (1988)
- **Tasks**
  1. For the S&P 500 and $q = 2$, compute $\hat{\sigma}_a^2$, $\hat{\sigma}_c^2(2)$, $\widehat{\mathrm{VR}}(2)$, $\hat{\delta}_1$, $z(2)$ and $z^*(2)$ with the Lo–MacKinlay formulas (The Lo–MacKinlay Estimator).
  2. For both markets and $q = 2, 5, 10, 20$, compute $\widehat{\mathrm{VR}}(q)$, $z(q)$, $z^*(q)$ and the two-sided p-values with `variance_ratio(r, q)`.
  3. Compute the Chow–Denning statistic $\max_q |z^*(q)|$ and decide at 5% with the critical value 2.49; report its p-value (Several Horizons at Once: Chow–Denning).
- **Report**: a table with one row per market and q ($\widehat{\mathrm{VR}}$, z, z^*, p), the Chow–Denning (CD) decision, and one sentence on why $|z| > |z^*|$
- Notebook: section B1

B1 [Solved]: S&P 500, $q = 2$, Step by Step

- Returns in %: $T = 6\,717$, $\hat{\mu} = 0.0247$; one-day variance $\hat{\sigma}_a^2 = \frac{1}{T-1}\sum_t e_t^2 = 1.4708$
- $6\,716$ overlapping sums $r_t + r_{t-1} - 2\hat{\mu}$; divisor $m = q(T - q + 1)(1 - q/T) = 13\,428$; $\hat{\sigma}_c^2(2) = 1.3246$
- $\widehat{\mathrm{VR}}(2) = 1.3246/1.4708 = 0.9006$
- Classical: $\mathrm{SE} = \sqrt{2(2q-1)(q-1)/(3qT)} = \sqrt{1/T} = 0.0122$, so $z(2) = (-0.0994)/0.0122 = -8.15$
- Robust: $\hat{\delta}_1 = T\sum_t e_t^2 e_{t-1}^2/(\sum_t e_t^2)^2 = 4.967$ (it would be about 1 under i.i.d. returns); for $q = 2$, $\hat{\theta}(2) = \hat{\delta}_1$
  - $z^*(2) = \sqrt{T}\,(\widehat{\mathrm{VR}}(2) - 1)/\sqrt{\hat{\theta}(2)} = 81.96 \times (-0.0994)/\sqrt{4.967} = -3.66$, $p = 0.0003$

MFM_ch2_seminar

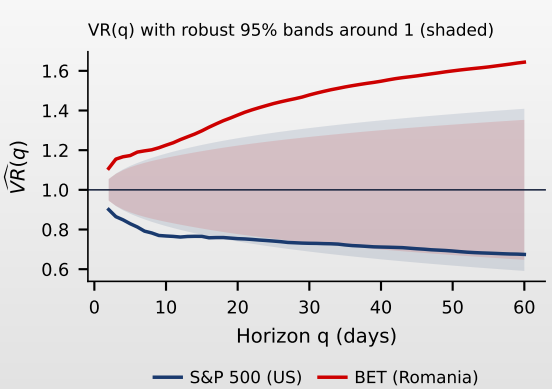
B1 [Solved]: Both Markets, Four Horizons

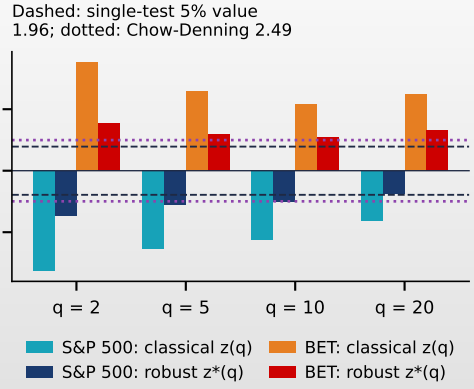
	q	$\widehat{\mathrm{VR}}(q)$	$z(q)$	p (RW1)	$z^*(q)$	p (RW3)
S&P 500	2	0.901	-8.15	< 0.0001	-3.66	0.0003
	5	0.829	-6.39	< 0.0001	-2.78	0.0055
	10	0.767	-5.65	< 0.0001	-2.49	0.013
	20	0.754	-4.06	< 0.0001	-1.85	0.064
BET	2	1.108	8.83	< 0.0001	3.84	0.0001
	5	1.173	6.44	< 0.0001	2.98	0.0029
	10	1.225	5.44	< 0.0001	2.73	0.0063
	20	1.377	6.19	< 0.0001	3.31	0.0009

- Two-sided p-values: $p = 2(1 - \Phi(|z|))$, for $z(q)$ under RW1 and for $z^*(q)$ under RW3
- Chow–Denning: S&P 500 3.66 ($p = 0.0010$), BET 3.84 ($p = 0.0005$), both above 2.49: the joint test rejects RW3 for both markets

MFM_ch2_seminar

B1 [Solved]: Chart





MFM_ch2_seminar

B1 [Solved]: Interpretation

- Interpretation: reversal in the S&P 500 ($\widehat{\mathrm{VR}} < 1$), significant with the robust statistic up to $q = 10$; persistence in BET ($\widehat{\mathrm{VR}} > 1$) at all four horizons
- Why $|z| > |z^*|$: volatility clustering makes $\hat{\delta}_1 = 4.967$ for the S&P 500, so the robust standard error at $q = 2$ is $\sqrt{\hat{\delta}_1} \approx 2.2$ times the i.i.d. one; the robust statistic has about half the absolute magnitude ($|z|/|z^*|$ between 1.87 and 2.30)
- A rejection is not a trading rule: costs, risk and stability decide (B2)

MFM_ch2_seminar

B1: check before continuing

- ⊡ Explain why the point estimate can stay the same while the rejection decision changes.
- ⊡ For each conclusion, name the null hypothesis and the horizon tested.
- ⊡ Non-rejection is not proof of independence or proof of market efficiency.

C1: apply the method on your own

- Prerequisites: the worksheet in A1 and the robust test in B1.
- Check every claim of the AI answer yourself before consulting the solution.
- For every claim, identify the code that computes it and the assumption that supports it.

C1 [Proposed]: Critique an AI Answer — the Answer

AI: Artificial Intelligence.

- **Prompt** sent by a student to an AI assistant: "Test whether the S&P 500 (daily closes, 2000–2026) follows a random walk with the variance ratio. Give Python code and interpret the result"

```
r = np.log(sp).diff().dropna()      # sp: raw S&P 500 closes; audit cleaning against B1
T = len(r)
def vr(q):
    rq = r.rolling(q).sum().dropna()  # overlapping q-day returns
    return rq.var() / (q * r.var())
q = 5
z = (vr(q) - 1) / np.sqrt(2 * (2*q - 1) * (q - 1) / (3 * q * T))
p = 2 * stats.norm.sf(abs(z))
```

- **The AI's interpretation**: (a) "VR(2) $= 0.900$, VR(5) $= 0.829$, $z(5) = -6.41$, $p < 0.001$: the random walk is rejected"; (b) "since VR(2) $= 1 + 2\rho_1$, $\rho_1 \approx -0.050$"; (c) "VR < 1 means momentum: buy the S&P 500 after up days"

Back Exercise list

C1 [Proposed]: Critique an AI Answer — Task

- **Question**: which claims of the AI answer does its code actually support?
- **Data**: the code and interpretation on the previous slide; raw S&P 500 closes for the B1 period; use the index cleaning rule (slide Data Used) for your numerical checks
- **Tasks**
  1. Audit four points in a table (line, claim, verdict (supported / error / convention difference / not computed by the shown code), check): the estimator convention, the null and its standard error, the VR identity, the economic interpretation.
  2. Say which reported numbers the shown lines actually evaluate; defining a function does not execute it.
  3. Give two numerical checks: $\hat{\rho}_1$ computed directly, and $z^*(5)$ from `variance_ratio`.
  4. Write the corrected interpretation in two sentences.
- **Report**: the audit table, the two checks, the corrected interpretation; log each error as in `AI_ERRORS.md`
- **Model**: A1, B1 [Solved]
- Notebook: section B1

The core route is complete

- You can calculate a variance ratio, choose an inference procedure and qualify an economic conclusion.
- Keep the A1 worksheet, the B1 comparison and the C1 audit as evidence of these skills.

Practice and further study

- ⊡ Practice: check a new sample
- ⊡ Practice: calendar effects
- ⊡ Further study: long memory
- ⊡ Further study: other tests and markets
- ⊡ Further study: forecasting
- ⊡ Further study: an event study
- ⊡ Further study: research projects

Check a new sample

- After B1, test whether the full-sample result also holds outside crises.
- Compare the retained sample with the full sample: changes in the VR estimate and in its robust uncertainty answer different questions.
- Back to module selection.

B2 [Proposed]: Is the S&P 500 Reversal a Crisis Effect? — Task

- **Question**: does the S&P 500 reversal of B1 survive without the two crisis periods?
- **Data**
  - S&P 500, daily log returns 2000–2026, $T = 6\,717$; crisis days, fixed in advance: 1 Sep 2008 – 30 Jun 2009 and 15 Feb – 30 Jun 2020
  - Removing a period leaves contiguous segments. A lag pair $(r_t, r_{t-\mathbf{1}})$ with $t = 1$ Jul 2009 and $t - 1 = 29$ Aug 2008 would join two days ten months apart: it is forbidden; so is any 5-day sum that spans a gap
- **Tasks**
  1. On all days compute $\hat{\rho}_{\mathbf{1}}$, the i.i.d. $t = \hat{\rho}_{\mathbf{1}}\sqrt{T}$, the robust $t = \hat{\rho}_{\mathbf{1}}/\sqrt{\hat{\tau}_{\mathbf{1}}}$ and $z^*(5)$.
  2. Remove the crisis days and repeat within the retained segments, using the pooled mean and pooled sum of squared deviations. Keep T as the count of retained returns; in the VR divisor replace the number of overlapping sums by the valid within-segment count. Report the kept returns, the valid lag pairs and the valid 5-day sums.
  3. Decide at 5% with the robust t: its SE (Standard Error) $\sqrt{\hat{\tau}_{\mathbf{1}}}$ is derived under $\rho_{\mathbf{1}} = 0$, so $\pm 1.96\sqrt{\hat{\tau}_{\mathbf{1}}}$ is a test band around zero, not an interval around $\hat{\rho}_{\mathbf{1}}$.
  4. Give 95% intervals for $\rho_{\mathbf{1}}$ in both samples with the HAC (Heteroskedasticity and Autocorrelation Consistent, standard errors) SE of the estimator (Newey–West on its influence function, Chapter 2), valid also when $\rho_{\mathbf{1}} \neq 0$.
- **Report**
  - a two-row table and one sentence
  - how does the estimate change?
  - is removing crisis days legitimate?
- **Model**: B1 [Solved]
- Notebook: section B2

Back Exercise list

Calendar effects

- Use B1 to distinguish an estimate from its uncertainty; then learn regression and multiple-testing adjustments here.
- Compare weekday means and report whether the conclusion survives adjustment for the family of tests.
- Back to module selection.

Calendar Effects as a Regression

OLS: Ordinary Least Squares.

- **Turn of the month** (TOM), the window of Lakonishok & Smidt (1988): the last trading day of a month and the first three of the next

$$r_t = a + b\,\mathrm{TOM}_t + u_t$$

  - the monthly effect was first documented by Ariel (1987), on a longer window (the last day and the first half of the month)
  - $\mathrm{TOM}_t = 1$ on a turn-of-the-month day, 0 otherwise (a dummy variable); u_t: the residual
- With a 0/1 regressor OLS compares two means

$$\hat{a} = \bar{r}_{\mathbf{other}}, \qquad \hat{b} = \bar{r}_{\mathbf{TOM}} - \bar{r}_{\mathbf{other}}$$

  - H_0: $b = 0$, no TOM effect; $t = \hat{b}/\,\mathrm{SE}(\hat{b})$, SE: the standard error
- The other effects of B4
  - **January**: monthly returns on a January dummy, same regression
  - **day of the week**: five dummies; a Wald test of equal means, $W = (\mathrm{R}\hat{\boldsymbol{\beta}})'(\mathrm{R}\hat{\mathrm{V}}\mathrm{R}')^{-1}\mathrm{R}\hat{\boldsymbol{\beta}} \approx \chi^2_4$
  - $\mathrm{R}\hat{\boldsymbol{\beta}}$: the four differences from Monday's mean; $\hat{\mathrm{V}}$: their Newey–West covariance (next slide)
  - χ^2_m: the chi-squared distribution with m degrees of freedom, the distribution of a sum of m squared independent $N(0,1)$ variables; here $m = 4$
  - reject equal means at 5% when $W > 9.49$, the 95% quantile of χ^2_4
  - with a single restriction $W = t^2$, and the threshold becomes $1.96^2 = 3.84$

Newey–West (HAC) Standard Errors

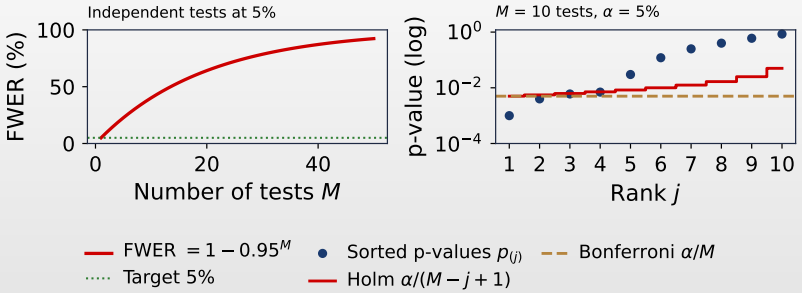
- **HAC** (heteroskedasticity- and autocorrelation-consistent) Newey & West (1987): the variance of $\hat{\boldsymbol{\beta}}$ when the residuals change variance and are correlated over L lags

$$\widehat{\mathrm{Var}}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{X})^{-1}\,\hat{\mathbf{S}}\,(\mathbf{X}'\mathbf{X})^{-1}, \qquad \hat{\mathbf{S}} = \sum_{j=-L}^{L}\left(1 - \frac{|j|}{L+1}\right)\sum_{t} \mathbf{x}_t \hat{u}_t \hat{u}_{t-j} \mathbf{x}_{t-j}'$$

  - $\mathbf{X}$: the $T \times 2$ matrix with rows $\mathbf{x}_t' = (1, \mathrm{TOM}_t)$; $\hat{u}_t$: the OLS residuals
  - weights $1 - |j|/(L+1)$ (Bartlett): 1 at $j = 0$, falling linearly to 0 after lag L
  - $L = 0$: the White (heteroskedasticity-robust) errors; $L = 5$ in B3
  - L is fixed before looking at the results: 5 for daily data (B3), 3 or 12 for monthly data (B4, B8)
- Classical OLS errors
  - $\widehat{\mathrm{Var}}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}_u^2(\mathbf{X}'\mathbf{X})^{-1}$: constant variance, no autocorrelation
  - HAC and OLS errors differ when $\mathbf{x}_t\hat{u}_t$ is heteroskedastic or autocorrelated: HAC is larger with positive autocorrelation and can be smaller with negative autocorrelation
  - the point estimate is the same, only the uncertainty changes

Multiple Testing: FWER, Bonferroni and Holm

FWER: Family-Wise Error Rate.



- **FWER** of M independent tests at 5%: $1 - 0.95^M$ (left), $M = 10$: 40.1%
  - **Bonferroni**: reject test i if $p_i < \alpha/M$; p_i: its p-value, α: the target FWER
  - **Holm** Holm (1979): sort $p_{(1)} \leq \cdots \leq p_{(M)}$, reject while $p_{(j)} < \alpha/(M - j + 1)$, stop at the first failure
  - both keep FWER $\leq \alpha$; Holm rejects at least as much (right: 4 rejections against 2)
  - in finance: hundreds of published anomalies were tested on the same returns Harvey, Liu & Zhu (2016); a step-down procedure that uses the dependence between the tests: Romano & Wolf (2005)

B3 [Solved]: The Turn-of-the-Month Effect with HAC Errors — Task

- **Question**: do returns on turn-of-the-month days exceed those on other days?
- **Data**
  - S&P 500 and BET, daily log returns 2000–2026 in basis points (bp; 1 bp = 0.01%), $r_t = 10\,000 \ln(P_t / P_{t-1})$, each on its own trading calendar; the incomplete last month (September 2026) is dropped
  - TOM_t: the turn-of-the-month dummy of Calendar Effects as a Regression; example: 31 Jul and 3, 4, 5 Aug 2026 are TOM days
- **Tasks**
  1. Build TOM_t from the trading calendar (rank of the day within its month, from the start and from the end).
  2. Estimate $r_t = a + b\,\text{TOM}_t + u_t$ by OLS; $\hat{a}$ is the mean on other days, $\hat{a} + \hat{b}$ the mean on TOM days.
  3. Compute the OLS and the Newey–West HAC standard errors of $\hat{b}$ with 5 lags Newey & West (1987), both t-statistics and the HAC p-value.
  4. Convert $\hat{b}$ into an approximate extra log return per month (4 TOM days), before costs.
- **Report**: a table (share of TOM days, $\hat{a}$, $\hat{b}$, both SEs and t), a 95% interval chart, one sentence on why OLS and HAC differ
- Notebook: section B3

Back Exercise list

B3 [Solved]: Design Matrix

date	r_t (bp)	1	TOM_t
2026-07-27	+1.6	1	0
2026-07-28	+21.0	1	0
2026-07-29	−152.8	1	0
2026-07-30	+164.7	1	0
2026-07-31	+69.8	1	1
2026-08-03	+146.8	1	1
2026-08-04	+177.4	1	1
2026-08-05	−16.8	1	1
2026-08-06	−17.6	1	0
2026-08-07	+61.7	1	0

- S&P 500 rows around the July/August 2026 boundary: the regressors are a column of ones and the TOM dummy
- OLS: $\hat{b} = \bar{r}_{\mathbf{TOM}} - \bar{r}_{\mathbf{other}}$; S&P 500 $5.1 - 1.9 \approx 3.2$ bp; BET $20.4 - 3.1 \approx 17.3$ bp
- HAC: the Newey–West variance of Newey–West (HAC) Standard Errors with $L = 5$, i.e. weights $1 - |j|/6$
- In code: `sm.OLS(r, X).fit(cov_type='HAC', cov_kwds={'maxlags': 5})`

MFM_ch2_calendar_anomalies

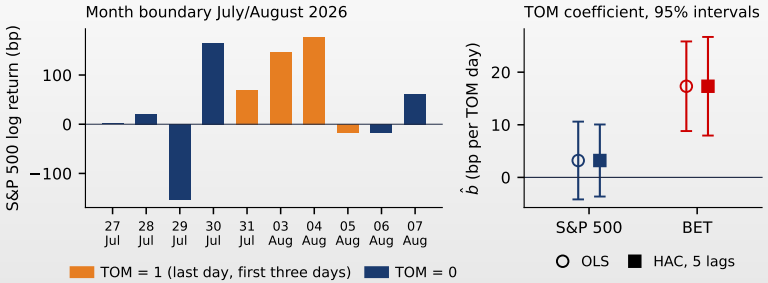
B3 [Solved]: Estimates

	TOM share	$\hat{a}$ (bp)	$\hat{b}$ (bp)	SE_{OLS}	t_{OLS}	SE_{HAC}	t_{HAC} (p)
S&P 500	19.1%	1.9	3.2	3.8	0.85	3.5	0.92 (0.36)
BET	19.2%	3.1	17.3	4.3	3.99	4.8	3.62 (0.0003)

- BET: HAC 95% interval $[8.0, 26.7]$ bp; the HAC SE exceeds the OLS SE because $x_t \hat{u}_t$ is positively autocorrelated: the four TOM days are consecutive, so inside each block it inherits the return autocorrelation ($\hat{\rho}_1 = 0.109$); t falls from 3.99 to 3.62, still a rejection
- S&P 500: $[-3.6, 10.1]$ bp contains 0; here the HAC SE is below the OLS SE, because the autocorrelation is negative ($\hat{\rho}_1 = -0.099$)
- $4\hat{b} \approx 69$ bp (BET): an approximate extra monthly log return relative to the other-day mean, before costs, not a strategy return

MFM_ch2_calendar_anomalies

B3 [Solved]: Results and Interpretation



- Chart, right panel: the BET interval excludes 0 with both standard errors; the S&P 500 interval contains 0 with both
- Interpretation: one effect in one market was picked among many: correct for the effects tried (B4) before looking for a mechanism (salary and pension inflows)

MFM_ch2_calendar_anomalies

B4 [Proposed]: Fifteen Calendar Tests and Holm — Task (1/2)

- **Question**: which of the fifteen calendar effects survive a correction for multiple testing?
- **Data**: the 15 p-values in the table below: 5 markets $\times$ 3 effects (day of week: Wald test of equal means, HAC 5 lags; January: monthly returns, HAC 3 lags; TOM: as in B3), 2000–2026

Market	Effect	p	Market	Effect	p
S&P 500 (US)	day of week (equal means)	0.826	Euro Stoxx 50	turn of the month	0.227
S&P 500 (US)	January	0.426	WIG20 (Poland)	day of week (equal means)	0.227
S&P 500 (US)	turn of the month	0.358	WIG20 (Poland)	January	0.686
BET (Romania)	day of week (equal means)	0.305	WIG20 (Poland)	turn of the month	0.085
BET (Romania)	January	0.137	BUX (Hungary)	day of week (equal means)	0.933
BET (Romania)	turn of the month	0.0003	BUX (Hungary)	January	0.024
Euro Stoxx 50	day of week (equal means)	0.632	BUX (Hungary)	turn of the month	0.0003
Euro Stoxx 50	January	0.774			

Back Exercise list

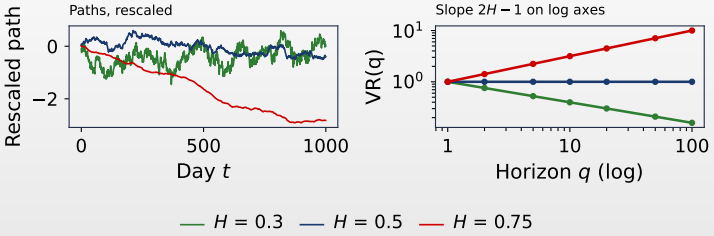
B4 [Proposed]: Fifteen Calendar Tests and Holm — Task (2/2)

- **Tasks**
  1. Count the raw rejections at 5% and compute the FWER of 15 independent tests.
  2. Apply Bonferroni: reject if $p_i < 0.05/15$.
  3. Apply Holm: sort the p-values, compare $p_{(j)}$ with $0.05/(16 - j)$ and stop at the first failure.
- **Report**: the sorted table with ranks, thresholds and three decision columns (raw, Bonferroni, Holm), and one sentence on BUX January
- **Model**: B3 [Solved]
- Notebook: section B4

Long memory

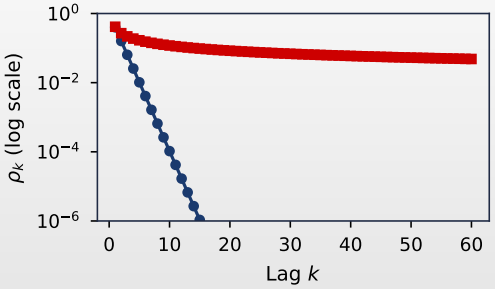
- After A1 and B1, distinguish an implied Hurst exponent from evidence of long memory.
- Compare scaling across horizons with the reference null bands; overlapping windows describe changes but do not provide independent tests.
- Back to module selection.

The Hurst Exponent



- **Hurst exponent** $H \in (0, 1)$, from the hydrology of the Nile Hurst (1951): how fast the spread of a cumulated series grows with the horizon
  - the estimator of Hurst (1951): the rescaled range R/S, the range of the cumulated deviations from the mean divided by the standard deviation, which grows like n^H with the sample length n
  - self-similar increments (fractional Gaussian noise, fGn: the q-day sum has the distribution of one day scaled by q^H): $\mathrm{Var}(r_t + \cdots + r_{t-q+1}) = q^{2H}\mathrm{Var}(r_t)$, so $\mathrm{VR}(q) = q^{2H-1}$
  - $H = 0.5$: no memory; $H > 0.5$: smooth, trending paths; $H < 0.5$: anti-persistent, rough paths (left)
- Implied exponent $H(q)$: the H that solves $\mathrm{VR}(q) = q^{2H-1}$ at a single horizon q; A2 derives its formula

Short and Long Memory



- **Short memory**: the AR(1) of The Variance Ratio as a Weighted Sum

$$r_t = \phi\, r_{t-1} + u_t, \qquad \rho_k = \phi^k$$

  - geometric decay: a straight line on the log scale (blue)
  - VR(q) $\to (1+\phi)/(1-\phi)$, bounded, so $H(q) \to 0.5$ as q grows

- **Long memory**, fGn with $H > 0.5$

$$\rho_k = \tfrac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right)$$

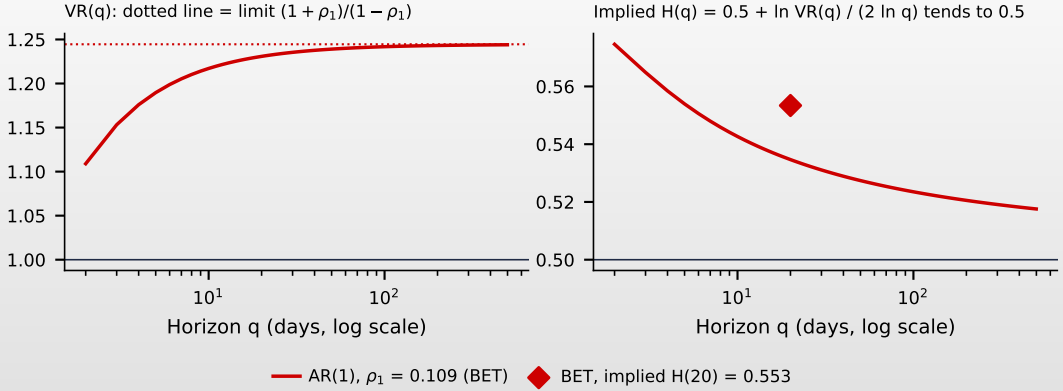
  - hyperbolic decay $\rho_k \approx H(2H-1)k^{2H-2}$ (red): at $k = 20$, 0.084 against 10^{-8} for AR(1)

- One horizon q cannot tell the two apart: compare $H(q)$ over many q (A2, A3)
  - a formal test: the statistic V of Lo (1991), a rescaled range computed with a long-run (HAC) standard deviation

A2 [Solved]: The Implied Hurst Exponent of BET — Task

- **Question**: does BET's $\widehat{\mathrm{VR}}(20) > 1$ imply long memory?
- **Data**: BET, daily log returns 2000–2026 (closing prices): $\widehat{\mathrm{VR}}(20) = 1.377$, $\hat{\rho}_1 = 0.109$
- **Tasks**
  1. Take logarithms of the fractional-Gaussian-noise relation $\mathrm{VR}(q) = q^{2H-1}$ (The Hurst Exponent) and solve for the implied $H(q)$.
  2. Compute the implied $H(20)$ for BET.
  3. Compute VR(20) and $H(20)$ of a short-memory AR(1) with $\rho_k = 0.109^k$, using the weighted sum of The Variance Ratio as a Weighted Sum.
- **Report**: the formula, the two $H(20)$ values with three decimals, and whether one horizon can identify long memory
- Notebook: section A2

A2 [Solved]: Chart



MFM_ch2_seminar

A2 [Solved]: Interpretation

- (1) $\ln \mathrm{VR}(q) = (2H - 1) \ln q$, so $H(q) = 0.5 + \ln \mathrm{VR}(q)/(2 \ln q)$
- (2) $\ln 1.377 = 0.3199$, $2 \ln 20 = 5.991$: $H(20) = 0.5 + 0.3199/5.991 = 0.553$
- (3) AR(1): $\mathrm{VR}(20) = 1.231$, $H(20) = 0.535$, and $H(250) = 0.520$ (the curve of the chart)
  - The curve is theoretical and runs to $q = 500$ only to show the limit $H(q) \to 0.5$
  - Its horizons are not the four empirical tests, which stay at $q = 2, 5, 10, 20$
- Interpretation: short memory alone gives most of BET's $H(20)$, and its $H(q)$ drifts back to 0.5: one horizon cannot identify long memory

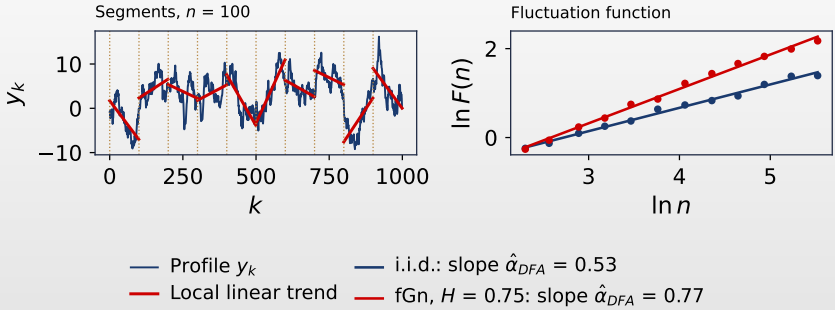
MFM_ch2_seminar

A3 [Proposed]: The Implied Hurst Exponent of the S&P 500 — Task

- **Question**: is the S&P 500 "anti-persistent", or does it only reverse in the short run?
- **Data**: S&P 500, daily log returns 2000–2026 (closing prices): $\widehat{\mathrm{VR}}(20) = 0.754$; a short-memory benchmark, the AR(1) with $\rho_1 = -0.1$, so $\rho_k = (-0.1)^k$
- **Tasks**
  1. Compute the implied $H(20)$ of the S&P 500 with the formula of A2.
  2. Compute VR(20) and $H(20)$ of the AR(1) benchmark.
  3. Compute VR(250), $H(250)$ and the limit $\mathrm{VR}(\infty) = (1+\rho_1)/(1-\rho_1)$ of the AR(1) benchmark.
  4. Plot VR(q) and the implied $H(q)$ of the benchmark for q from 2 to 500, next to the curve of A2.
- **Report**: the six numbers, the chart and one sentence: does $H(20) < 0.5$ require long memory?
- **Model**: A2 [Solved]
- Notebook: section A3

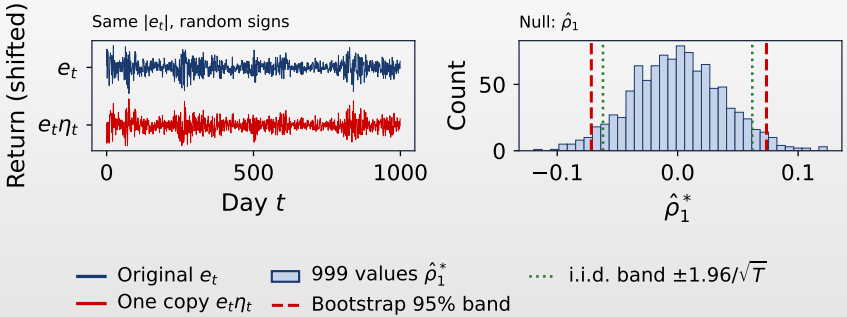
Back Exercise list

Detrended Fluctuation Analysis (DFA)



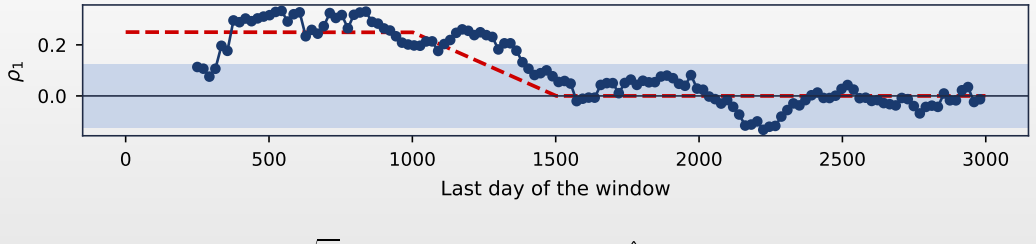
- **DFA** (Detrended Fluctuation Analysis) Peng et al. (1994): estimates H from how the detrended spread grows with n
  - profile $y_k = \sum_{i \le k}(r_i - \bar{r})$, cut into segments of length n, a line fitted in each (left)
  - $F(n)$: root mean square of the residuals around the local lines; $F(n) \propto n^{\alpha_{DFA}}$
  - $\hat{\alpha}_{DFA}$: OLS slope of $\ln F(n)$ on $\ln n$ (right); 0.5 without memory, above 0.5 with long memory

The Sign-Flip (Wild) Bootstrap



- **Wild bootstrap** Kim (2006): new series $e_t^* = e_t\eta_t$, with $\eta_t = \pm 1$ with probability $1/2$, independent
  - $|e_t^*| = |e_t|$: volatility clustering is kept; random signs destroy any autocorrelation, so H_0 holds in every copy
  - repeat B times (here 999), compute the statistic each time; the 2.5% and 97.5% quantiles form the null band
  - for the simulated series of the chart ($T = 1\,000$) the null band for $\hat{\rho}_1$ is $[-0.072; 0.074]$, wider than the i.i.d. null band ± 0.062

Rolling Windows and Adaptive Markets



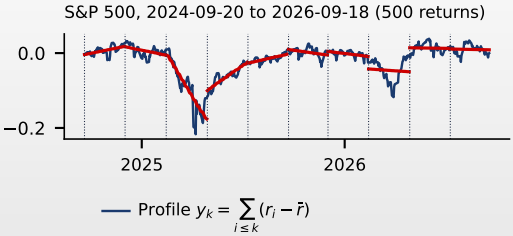
- **Rolling window**: re-estimate on the last W days, moving the window forward by a fixed step
  - simulated: $\rho_1 = 0.25$ for 1 000 days, then fading to 0; $W = 250$, step 21 days
  - overlapping windows share most data: neighbouring estimates are correlated, so a count of rejections is only descriptive
- **Adaptive markets hypothesis** (AMH) Lo (2004): efficiency changes over time with competition and learning
  - a predictable pattern attracts traders until it fades, as in the simulation after day 1 000

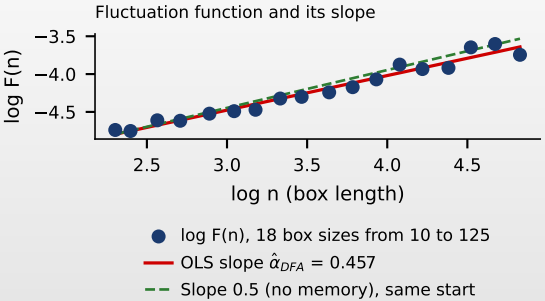
B5 [Solved]: Rolling DFA on the S&P 500 — Task

- **Question**: does the S&P 500 DFA exponent ever leave a null band that allows volatility clustering?
- **Data**
  - S&P 500, daily log returns 2000–2026; windows of 500 returns moved by 21 days (297 windows); function `dfa_hurst(x)`
  - DFA grid: 18 box lengths spaced evenly in $\log n$ from 10 to 125 (a quarter of the window), truncated to integers $(10, 11, 13, \ldots, 107, 124)$, duplicates removed
  - For a box length n the profile is cut into $k = \lfloor 500/n \rfloor$ boxes; the remaining $500 - kn$ points are discarded
  - Random seed of the bootstrap in each window: 42 + the index of the last day of the window
- **Tasks**
  1. For the last window (2024-09-20 to 2026-09-18) build the profile, detrend it in boxes and fit $\log F(n)$ on $\log n$.
  2. Compute α_{DFA} in every window.
  3. In each window build a sign-flip null band (The Sign-Flip (Wild) Bootstrap) from 199 series $e_t \eta_t$; keep the 2.5% and 97.5% quantiles of α_{DFA}.
  4. Compute the share of windows outside their own null band and outside one fixed i.i.d. Student-t_4 null band (300 simulations of length 500): Student t with 4 degrees of freedom has much heavier tails than the Normal, so this band keeps fat tails but no volatility clustering.
- **Report**: the one-window fit, the rolling chart, both shares, and one sentence on what the null band assumes
- Notebook: section B5

Back Exercise list

B5 [Solved]: One Window, Step by Step

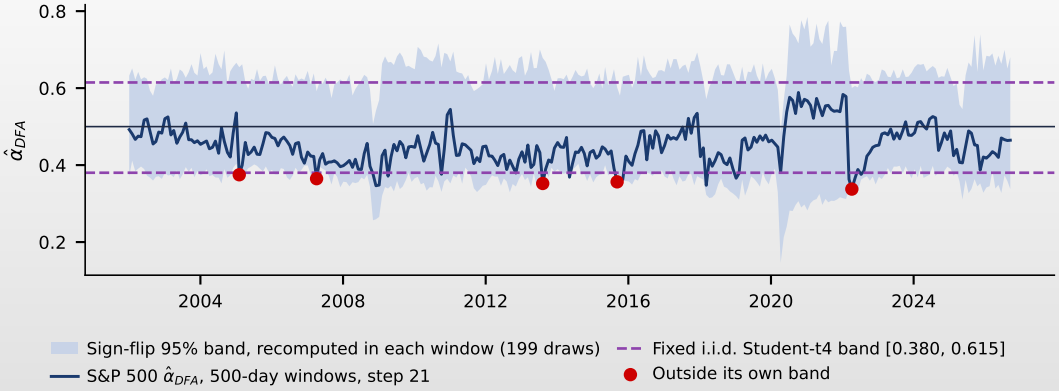




- Left: the profile y_k of the last window, detrended box by box as in Detrended Fluctuation Analysis (DFA)
- 18 box lengths from 10 to 125; the OLS slope of $\log F(n)$ on $\log n$ is $\hat{\alpha}_{DFA} = 0.457$ (right), close to 0.5

MFM_ch2_seminar

B5 [Solved]: All Windows: Chart



MFM_ch2_seminar

B5 [Solved]: All Windows: Interpretation

- Outside their own null band: 1.7% of 297 windows; outside the fixed Student-t_4 null band $[0.380, 0.615]$: 6.7%
- Mean $\hat{\alpha}_{DFA} = 0.451$; minimum 0.338 (window ending 2022-04-08), maximum 0.589 (2020-10-08); the null band is widest (0.503) in the window ending 2021-01-08
- Interpretation: the null band is **pointwise**: 95% coverage for each window taken alone, not for all 297 windows at once; under the null about 5% of the windows are expected to leave it on average, not guaranteed in correlated windows
- It is a benchmark under symmetric signs, independent given $|e_t|$, not a confidence interval
- Overlapping windows are correlated, so the shares describe, they do not test stability; the 1.7% share does not indicate memory beyond what chance would produce; non-rejection does not prove there is no memory

MFM_ch2_seminar

B6 [Proposed]: Adaptive Efficiency of Bitcoin — Task

- **Question**: has Bitcoin moved towards efficiency since its early years, as the AMH Lo (2004) would allow?
- **Data**: Bitcoin (USD) closing prices, daily log returns 7 days a week from 1 Jan 2011 to 18 Sep 2026 (course file BTC-USD.CC.csv, including the early years); 180 windows of 365 days moved by 30 days; H_0: RW3 in every window
- **Tasks**
  1. Compute the robust $z^*(5)$ in each window.
  2. In each window build a sign-flip null band for $z^*(5)$ with 999 replications (random seed = 42 + index of the last day of the window) and keep its 2.5% and 97.5% quantiles.
  3. Separately for windows ending in 2012–2016 and 2017–2026, count those outside their null band and outside ± 1.96 and compare with $0.05\times$ the number of windows.
  4. Compare the 2012–2016 windows with Urquhart (2016) (next slide): which of his tests address RW3, which can reject only because of volatility clustering, and do the conclusions agree?
- **Report**: the chart, the flagged windows with z^* and null band, both counts, and one sentence: what do the counts establish, and what can they not establish?
- **Model**: B5, B1 [Solved]
- Notebook: section B6

Back Exercise list

B6: What the Urquhart Tests Detect

BDS: Brock–Dechert–Scheinkman–LeBaron test.

- Compare with Urquhart (2016): 2010–2016, a battery of classical tests
  - Ljung–Box Ljung & Box (1978): $Q(m) = T(T+2)\sum_{k=1}^{m} \hat{\rho}_k^2/(T-k)$, approximately χ_m^2 under i.i.d. returns, tests the first m autocorrelations jointly
  - runs and the rank test of Bartels (1982), which replaces returns by their ranks and is robust to fat tails
  - automatic VR Choi (1999): the horizon q is chosen from the data
  - BDS Brock, Dechert, Scheinkman & LeBaron (1996): tests independence by how often pairs of m-day histories lie close together; it also rejects under RW3 (volatility clustering), so it is not a test of RW3
  - the rescaled range R/S of The Hurst Exponent

Other tests and markets

- After B1, examine joint horizons, runs and comparison markets.
- Identify which departure each test detects and keep the horizon family separate from the market family.
- Back to module selection.

Before comparing markets

- B7 combines tests across horizons and across markets; these are two different families of hypotheses.

A4 [Proposed]: The Variance Ratio and Its Null Distribution — Task

- **Questions**
  - why is $\sqrt{T}(\widehat{\mathrm{VR}}(q) - 1)$ approximately Normal with variance $2(2q-1)(q-1)/(3q)$?
  - when does this fail?
- **Data**: r_t covariance-stationary: constant mean and variance $\gamma_{\mathbf{0}}$, autocovariances $\gamma_k = \mathrm{Cov}(r_t, r_{t-k})$ that depend only on the lag k, not on t, $\rho_k = \gamma_k/\gamma_{\mathbf{0}}$; under RW1 $\sqrt{T}(\hat{\rho}_{\mathbf{1}}, \ldots, \hat{\rho}_{q-\mathbf{1}})$ tends to $N(0, I_{q-\mathbf{1}})$, $I_{q-\mathbf{1}}$ the identity matrix (independent, unit variances); S&P 500: $T = 6\,717$, $\widehat{\mathrm{VR}}(2) = 0.901$
- **Tasks**
  1. For $q = 3$, write $\mathrm{Var}(r_t + r_{t-\mathbf{1}} + r_{t-\mathbf{2}})$ as the sum of the nine covariances and count how many times $\gamma_{\mathbf{0}}$, $\gamma_{\mathbf{1}}$, $\gamma_{\mathbf{2}}$ appear.
  2. Generalise: show that $\mathrm{VR}(q) = 1 + \sum_{k=\mathbf{1}}^{q-\mathbf{1}} a_k \rho_k$ with $a_k = 2(1 - k/q)$.
  3. Under RW1, with $a = (a_{\mathbf{1}}, \ldots, a_{q-\mathbf{1}})'$, compute $a' I a = \sum_k a_k^{\mathbf{2}}$ and simplify it to $2(2q-1)(q-1)/(3q)$; evaluate it for $q = 2, 5, 20$.
  4. Compute $z(2)$ for the S&P 500 and explain why I must be replaced by $\mathrm{diag}(\delta_k)$, the diagonal matrix of the robust variances δ_k, under RW3.
- **Report**: the counting, the variance in two columns (asymptotic, of $\sqrt{T}(\widehat{\mathrm{VR}} - 1)$; finite-sample, of $\widehat{\mathrm{VR}}$: divide by T) and $z(2)$
- **Model**: A1 [Solved]
- Notebook: section A4

Back Exercise list

The Runs Test

- **Run**: a maximal block of consecutive returns with the same sign Wald & Wolfowitz (1940)
  - n_1: the number of positive days, n_2: of negative days, $n = n_1 + n_2$
  - R: the number of runs, equal to the number of sign changes $+1$
- Under H_0 (independent signs), conditional on n_1 and n_2

$$\mathbb{E}[R] = \frac{2n_1 n_2}{n} + 1, \qquad \mathrm{Var}[R] = \frac{2n_1 n_2(2n_1 n_2 - n)}{n^2(n-1)}, \qquad z = \frac{R - \mathbb{E}[R]}{\sqrt{\mathrm{Var}[R]}} \approx N(0,1)$$

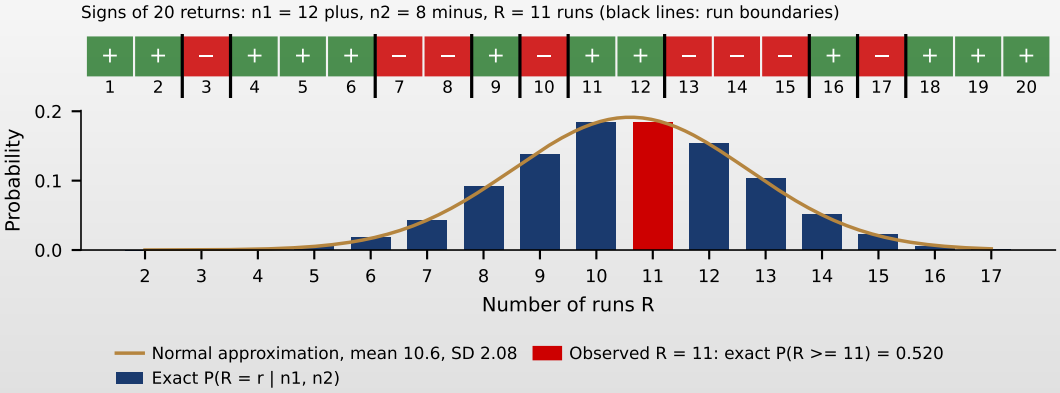
  - too few runs ($z < 0$): signs persist; too many runs ($z > 0$): signs alternate
  - for small n use the exact distribution of R given n_1, n_2 instead of the Normal approximation
- Example: ++--+-: $n_1 = 3$, $n_2 = 3$, $R = 4$
  - $\mathbb{E}[R] = 2 \cdot 9/6 + 1 = 4$: exactly as many runs as expected

A5 [Solved]: The Runs Test, Step by Step — Task

- **Question**: are 20 signs consistent with independence?
- **Data**: signs of 20 consecutive daily returns, no zeros: `++-+++--+-++---+-+++`; under H_0 the signs are independent with a constant probability of a positive day Wald & Wolfowitz (1940)
- **Tasks**
  1. Number the signs 1 to 20, mark every sign change and count n_1, n_2 and the runs R.
  2. Compute $\mathbb{E}[R]$ and $\mathrm{Var}[R]$ with the formulas of The Runs Test.
  3. Compute the approximate z and its two-sided p-value.
  4. Compare with the exact conditional distribution of R given n_1 and n_2.
- **Report**: n_1, n_2, R, $\mathbb{E}[R]$, $\mathrm{Var}[R]$, z, both p-values and what too few or too many runs would mean
- Notebook: section A5

A5 [Solved]: Chart

SD: Standard Deviation.



MFM_ch2_seminar

A5 [Solved]: Interpretation

- (1) Runs: $++ \mid - \mid +++ \mid -- \mid + \mid - \mid ++ \mid --- \mid + \mid - \mid +++$, so $n_1 = 12$, $n_2 = 8$, $R = 11$ (10 sign changes)
- (2) $\mathbb{E}[R] = 2 \cdot 12 \cdot 8/20 + 1 = 10.60$; $\mathrm{Var}[R] = 192 \cdot 172/(400 \cdot 19) = 4.345$
- (3) Approximate $z = 0.40/2.085 = 0.19$, two-sided $p = 0.85$
- (4) Exact: $P(R = 11) = 0.183$, $P(R \geq 11) = 0.520$, $P(R \leq 11) = 0.663$, so the exact two-sided $p = \min\{1,\ 2\min[P(R \geq 11), P(R \leq 11)]\} = 1$
- Interpretation: do not reject independence
  - too few runs would mean persistent signs
  - too many runs would mean alternating signs

MFM_ch2_seminar

B7 [Proposed]: BET-TR and Emerging Markets — Task

- **Question**: are emerging markets less efficient as a group?
- **Data**: BET-TR (first close 23 Sep 2014) and WIG20, BUX, PX, BIST 100, Bovespa, IPC, Nifty 50, Shanghai, closing prices, 2000–2026, each on its own calendar
- **Tasks**
  1. Start with BET and BET-TR over the same period (BET-TR's date bounds), on each series' retained observations: classical $z(2)$, robust $z^*(2)$, Chow–Denning.
  2. Repeat for the eight other indices and report three decisions separately: classical $q = 2$, robust $q = 2$, robust joint over four horizons.
  3. Before any group-wide claim, correct the joint p-values for BET-TR and the eight listed indices (nine tests, Holm).
- **Report**: one table (one row per market), the matched BET comparison and one sentence: how much of the BET versus BET-TR gap changes when the periods match? Different retained observations still prevent attributing the remaining gap solely to dividends.
- **Model**: B1 [Solved]
- Notebook: section B7

Back Exercise list

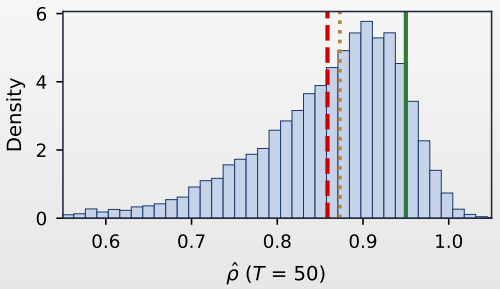
Forecasting

- After the course section on predictive regressions, evaluate forecasts using only information available at the forecast date.
- Compare out-of-sample squared forecast errors with the historical-mean benchmark; an in-sample fit does not establish predictive improvement.
- Back to module selection.

Before estimating the forecast model

- Use B3 as the worked example of regression with HAC inference.
- A6, developed next, supplies the other prerequisite: bias with a persistent predictor.

Predictive Regressions and the Kendall Bias



- **Predictive regression**: next period's return on a predictor x_t known today

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}, \quad x_{t+1} = c + \rho x_t + v_{t+1}$$

  - x_t: e.g. the log dividend–price ratio $dp_t = \ln D_t - \ln P_t$, very persistent (ρ near 1)
  - u, v: shocks; a price rise lowers x, so $\operatorname{corr}(u, v) < 0$

- **Kendall bias** Kendall (1954): OLS underestimates a persistent ρ in small samples

$$\mathbb{E}[\hat{\rho} - \rho] \approx -\frac{1 + 3\rho}{T}$$

  - chart: $\rho = 0.95$, $T = 50$; the mean of $\hat{\rho}$ is 0.859, below the truth
  - when u and v are correlated this bias passes to $\hat{\beta}$ Stambaugh (1999); A6 derives how

A6 [Proposed]: The Stambaugh Bias — Task (1/2)

- **Question**: why does a persistent valuation ratio look like a significant predictor too often?
- **Data**
  - System, notation and Kendall bias $\mathbb{E}[\hat{\rho} - \rho]$: Predictive Regressions and the Kendall Bias, with x_t the log dividend–price ratio
  - Write $u = \psi v + e$, $\psi = \sigma_{uv}/\sigma_v^2$, with $\sigma_{uv} = \mathrm{Cov}(u, v)$, σ_u, σ_v the shock SDs and e the part of u unrelated to v Stambaugh (1999)
  - Illustrative annual inputs (as fractions, not in %): $\rho = 0.98$, $T = 60$ years, $\mathrm{corr}(u, v) = -0.9$, $\sigma_u = 0.17$, $\sigma_v = 0.14$; true $\beta = 0$

Back Exercise list

A6 [Proposed]: The Stambaugh Bias — Task (2/2)

- **Context**: the same T pairs and demeaned regressor in both OLS regressions; jointly Normal i.i.d. shocks make e independent of the full regressor path.
- **Tasks**
  1. With demeaned $\tilde{x}_t = x_t - \bar{x}$, show that $\hat{\beta} - \beta = \psi(\hat{\rho} - \rho) + \sum \tilde{x}_t e_{t+1} / \sum \tilde{x}_t^2$ and derive $\mathbb{E}[\hat{\beta} - \beta]$.
  2. Compute ψ, $\mathbb{E}[\hat{\rho} - \rho]$ and the bias of $\hat{\beta}$.
  3. Compute the asymptotic standard error $\mathrm{SE}(\hat{\beta}) \approx \sigma_u / (\sigma_x \sqrt{T})$, with the stationary $\sigma_x = \sigma_v / \sqrt{1 - \rho^2}$, and the bias in units of this standard error.
  4. Set $\alpha = c = 0$, seed 42, and draw $x_0 \sim N(0, \sigma_x^2)$ independently of the shocks. Simulate 10 000 samples with Normal shocks; report the mean and the standard deviation of $\hat{\beta}$, the average OLS standard error and the share of samples with $|t| > 1.96$.
- **Report**: the derivation, ψ, the bias, the asymptotic SE, the simulated SD, and what a positive bias does to the false-rejection rate
- **Model**: the worked Stambaugh example of Chapter 2
- Notebook: section A6

Out-of-Sample Forecasts

OS: Out-of-Sample.

- Recursive forecast: at each t estimate on data up to t only, forecast $\hat{r}_{t+1}$; the benchmark is the historical mean $\bar{r}_t$
- **Out-of-sample R^2** Welch & Goyal (2008)

$$R^2_{OS} = 1 - \frac{\sum_t (r_{t+1} - \hat{r}_{t+1})^2}{\sum_t (r_{t+1} - \bar{r}_t)^2}$$

  - $R^2_{OS} > 0$: the model beats the historical mean; it can be negative
- **Clark–West** Clark & West (2007): for nested models, where the smaller model is a special case of the larger one (the mean is the model with $\beta = 0$)

$$f_{t+1} = (r_{t+1} - \bar{r}_t)^2 - \left[(r_{t+1} - \hat{r}_{t+1})^2 - (\bar{r}_t - \hat{r}_{t+1})^2\right]$$

  - t statistic of the mean of f_{t+1} (HAC), one-sided: reject equal accuracy when $t > 1.645$, the 95% quantile of $N(0,1)$
  - the test of Diebold & Mariano (1995): the HAC t statistic of the mean of $d_{t+1} = (r_{t+1} - \bar{r}_t)^2 - (r_{t+1} - \hat{r}_{t+1})^2$, the plain difference of squared errors
  - the last term of f_{t+1} removes the noise of estimating β, which Diebold–Mariano penalises
- **Campbell–Thompson** restrictions Campbell & Thompson (2008)
  - set $\hat{\beta} = 0$ when it has the wrong sign and replace a negative forecast by 0

B8 [Proposed]: Predictive Regression and an Out-of-Sample Test — Task (1/2)

ETF: Exchange-Traded Fund.

- **Question**: does the dividend–price ratio predict S&P 500 excess returns, in and out of sample?
- **Data**
  - Monthly table 1934–2026, $T = 1\,111$
  - P: the Shiller S&P Composite monthly average price (to 2026-06), extended with monthly averages of S&P 500 daily closes, scaled by the mean Shiller/extension ratio over the last 12 overlapping months
  - D: trailing 12-month dividends (Shiller, interpolated from quarterly totals), extended with the dividends of the SPY (SPDR S&P 500 ETF Trust) fund, an ETF on the S&P 500
  - SPY dividends: inferred from the gap between its closing price and its dividend-adjusted price, spliced on the common period
  - Excess return: $r_{t+\mathbf{1}} = \ln((P_{t+\mathbf{1}} + D_{t+\mathbf{1}}/12)/P_t) - \ln(1 + \text{TB3MS}_t/1\,200)$, with TB3MS (FRED series: 3-month Treasury bill, secondary market rate) the FRED series of the 3-month Treasury bill rate (secondary market), known at t
  - Predictor: $dp_t = \ln D_t - \ln P_t$, used one month later

Back Exercise list

B8 [Proposed]: Predictive Regression and an Out-of-Sample Test — Task (2/2)

- **Tasks**
  1. Estimate $r_{t+1} = \alpha + \beta\, dp_t + u_{t+1}$ by OLS with Newey–West errors (12 lags); estimate the AR(1) of dp and corr$(\hat{u}, \hat{v})$; correct $\hat{\beta}$ for the Stambaugh bias (A6).
  2. From January 1970, at each month re-estimate on data up to the previous month and forecast; the benchmark is the historical mean of the same data; compute R^2_{OS} Welch & Goyal (2008) (pseudo out-of-sample: recursive, but on data revised or interpolated afterwards, here the dividends).
  3. Compute the Clark–West t (HAC, 12 lags, one-sided) Clark & West (2007) and, for comparison, the Diebold–Mariano t Diebold & Mariano (1995); impose the Campbell–Thompson restrictions (slope sign, non-negative forecast) Campbell & Thompson (2008).

- **Report**: $\hat{\beta}$, t, $\hat{\rho}$, the bias, R^2_{OS}, both t statistics, and why Diebold–Mariano is the wrong test for nested models
- **Model**: B3; the worked Stambaugh example of Chapter 2
- Notebook: section B8

An event study

- After the course section on abnormal returns, distinguish the measured event reaction from its statistical significance.
- Estimate the normal return before the event, cumulate abnormal returns over the event window, and account for prediction uncertainty.
- Back to module selection.

- **AR (event study)**: Abnormal Return.
- **CAR**: Cumulative Abnormal Return.
- **BMP**: Boehmer–Musumeci–Poulsen.

Event Studies

- **Market model** on the estimation window (T_0 days before the event) MacKinlay (1997)

$$r_{i,t} = a_i + b_i\, r_{m,t} + \varepsilon_{i,t}, \qquad AR_{i,t} = r_{i,t} - \hat{a}_i - \hat{b}_i\, r_{m,t}$$

  - AR here is the abnormal return, not the autoregressive model AR(p)
  - $r_{m,t}$: the market (BET) return; $\hat{\sigma}_\varepsilon$: the residual SD
  - day 0: the first trading day on which the market can react to the announcement
  - $[\tau_1, \tau_2]$: the event window, in days relative to day 0, e.g. $[0, +5]$

- Cumulated abnormal return and its variance

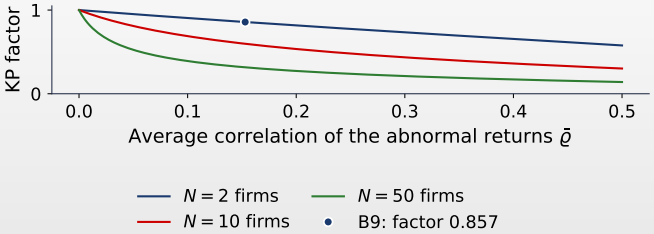
$$CAR[\tau_1, \tau_2] = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}, \qquad \mathrm{Var}(AR_t) = \hat{\sigma}_\varepsilon^2 \left(1 + \mathrm{x}_t' (\mathrm{X}_0' \mathrm{X}_0)^{-1} \mathrm{x}_t\right)$$

  - $\mathrm{x}_t = (1, r_{m,t})'$; X_0: the rows x_t' of the estimation window
  - the bracket (prediction error) exceeds 1 when $r_{m,t}$ is far from its usual values, e.g. a market crash on day 0

Testing the reaction across firms

- Several firms on the same date: their residuals are correlated (average correlation $\bar{\varrho}$)
  - intuition, the raw cross-sectional test for N firms: $t = \overline{CAR}/(s_{CAR}/\sqrt{N})$, with $\overline{CAR}$ and s_{CAR} the mean and SD of the N values; it assumes independent firms
  - Boehmer, Musumeci & Poulsen (1991) (BMP): the same cross-sectional test on *standardised* abnormal returns, each divided by its prediction-error SD
  - Kolari–Pynnönen Kolari & Pynnönen (2010): multiply the BMP statistic by $\sqrt{(1-\bar{\varrho})/(1+(N-1)\bar{\varrho})}$, N: number of firms

Correlated Firms in an Event Study



- Horizontal axis: the average correlation $\bar{\varrho}$ of the abnormal returns; vertical axis: the Kolari–Pynnönen factor that multiplies the BMP statistic
- With $\bar{\varrho} = 0$ the factor is 1: independent firms need no correction
- The factor falls faster with more firms: $N = 50$ and $\bar{\varrho} = 0.1$ give 0.39, so the uncorrected statistic is about 2.6 times too large
- Dot: B9, two banks with $\bar{\varrho} = 0.153$, factor 0.857

B9 [Proposed]: The Bank Tax, TLV and BRD Jointly — Task (1/2)

- **Question**: how much of the bank-tax fall survives proper event-study inference?
- **Data**
  - Banca Transilvania (TLV) and BRD (dividend- and split-adjusted prices) and BET (closing prices), prices aligned on common trading days, then log returns
  - Timeline: tax announced on the evening of 18 Dec 2018; day 0 = 19 Dec; estimation window days $[-250, -11]$ ($T_0 = 240$); event window $[0, +5]$
  - X_0 (240×2) and X_e (6×2): the $[1, r_{m,t}]$ rows of the estimation and event windows; ι: a 6×1 vector of ones; I_6: the 6×6 identity matrix; N, $\bar{\varepsilon}$: as in Event Studies

B9 [Proposed]: The Bank Tax, TLV and BRD Jointly — Task (2/2)

- **Tasks**
  1. For TLV fit the market model on $[-250, -11]$, compute $AR_{\mathbf{0}}$ (the abnormal return on day 0), the naive $t = AR_{\mathbf{0}}/\hat{\sigma}_\varepsilon$ and the t with the prediction-error variance; repeat for BRD.
  2. For the equal-weighted portfolio of the two banks fit the market model to its own returns and compute its $AR_{\mathbf{0}}$ and t; compare with the t that assumes independent banks.
  3. Compute $CAR[0, +5]$ and $CAR[+1, +5]$ with $\mathrm{Var}(CAR) = \iota' \hat{\sigma}_\varepsilon^{\mathbf{2}} [\boldsymbol{I_6} + X_e (X_{\mathbf{0}}' X_{\mathbf{0}})^{-\mathbf{1}} X_e'] \iota$ (adapted to the window used). For a window of L days, use its L rows of X_e, L ones and I_L: $L = 6$ for $[0, +5]$, $L = 5$ for $[+1, +5]$ (drop day 0).
  4. Compute the residual correlation $\bar{\varrho}$ and the Kolari–Pynnönen factor $\sqrt{(1 - \bar{\varrho})/(1 + (N - 1)\bar{\varrho})}$ Kolari & Pynnönen (2010).
- **Report**: one table (TLV, BRD, portfolio), the AR/CAR chart, and whether the reversal contradicts semi-strong efficiency; note that BET contains both banks
- **Model**: Event-study concepts; Chapter 2 bank-tax example; B3 for HAC. Optional: exercise B7 of Seminar 0
- Notebook: section B9

Research projects

- Complete the forecasting module first; use a control group or cross-validation to challenge your conclusion.
- Keep model selection separate from evaluation and state which alternative explanation the comparison can rule out.
- Back to module selection.

Project prerequisites

- For C2: complete B5 on rolling windows and the adaptive-market interpretation.
- For C3: complete B8 on out-of-sample forecast evaluation.
- The following pages introduce block bootstrap for C2 and ridge with cross-validation for C3.

Block Bootstrap and a Control Group

FTSE: Financial Times Stock Exchange (FTSE Russell, index provider).

- **Moving-block bootstrap** Künsch (1989): resample blocks of ℓ consecutive days, not single days
  - block starts drawn uniformly from $1, \ldots, T - \ell + 1$; blocks are glued until the series has length T
  - the dependence inside each block (autocorrelation, volatility clustering) is kept
  - percentile interval: the 2.5% and 97.5% quantiles of the B bootstrap estimates
  - ℓ must be long enough for the dependence that matters (autocorrelation, volatility clustering); C2 uses $\ell = 20$ days, a choice rather than a guarantee, so check the sensitivity to ℓ
- **Before/after contrast** for an event (a reclassification)

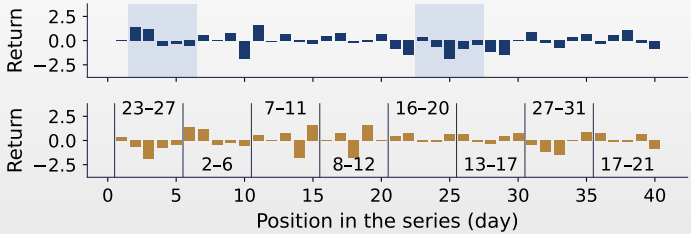
$$\Delta_i = \hat{\rho}_{\mathbf{1},i}^{\,\text{after}} - \hat{\rho}_{\mathbf{1},i}^{\,\text{before}}$$

  - reclassification: FTSE Russell ranks markets as developed, advanced emerging, secondary emerging or frontier (size, liquidity, foreign access, settlement)
  - Romania moved from frontier to secondary emerging on 21 Sep 2020 (C2)
  - i: a market; other events in the same years (a pandemic, new listings) also change Δ_i: confounders
- **Control group**: markets without the event

$$\Delta_{\text{treated}} - \overline{\Delta}_{\text{control}}$$

  - valid under *parallel trends*: without the event the treated market would have changed like the controls

The Moving-Block Bootstrap: a Graphical Illustration



- Top: 40 simulated daily returns; bottom: one bootstrap series of the same length, glued from 8 blocks of $\ell = 5$ consecutive days
- The labels give the source days of each block: the order inside a block is kept, so short-range dependence survives
- Blocks are drawn with replacement: days 8–11 appear twice (blocks 7–11 and 8–12), days 32–40 not at all

C2 [Proposed]: BVB Efficiency after the 2020 FTSE Upgrade — Task (1/2)

- **Question**: did BET's lag-1 autocorrelation move towards zero after the upgrade, relative to WIG20 and BUX?
- **Data**
  - FTSE Russell (2019): announced 26 Sep 2019, effective 21 Sep 2020 (Frontier to Secondary Emerging); background: Pele & Mazurencu-Marinescu (2012)
  - FTSE Russell ranks markets as developed, advanced emerging, secondary emerging or frontier, by size, liquidity, access for foreign investors and settlement
  - The upgrade brings passive funds that track emerging-market indices
  - BET, WIG20, BUX, closing prices; before: 21 Sep 2015 – 20 Sep 2020; after: 21 Sep 2020 – 21 Sep 2025; data after 21 Sep 2025 only in a separately reported extension

Back Exercise list

C2 [Proposed]: BVB Efficiency after the 2020 FTSE Upgrade — Task (2/2)

COVID-19: COronaVIrus Disease 2019.

- **Tasks**
  1. For each market compute $\hat{\rho}_{\mathbf{1}}$ with its HAC SE (Newey–West on the influence function, Chapter 2) and $z^*(5)$ in both windows.
  2. Build a 95% interval for the change in $\hat{\rho}_{\mathbf{1}}$ by a moving-block bootstrap, independent within each period (blocks of 20 days, 999 replications, seed 42) Künsch (1989).
  3. Plot rolling 250-day $\hat{\rho}_{\mathbf{1}}$ for the three markets with the robust 5% test band $0 \pm 1.96\sqrt{\hat{\tau}_{\mathbf{1}}}$ (SE under $\rho_{\mathbf{1}} = 0$; a test band around zero, not an interval around $\hat{\rho}_{\mathbf{1}}$), marking the announcement, the effective date and the end of the formal window.
  4. List the confounders (COVID-19, the 2023 Hidroelectrica listing, liquidity) and say what a control-adjusted contrast would need.
- **Report**: a Quantlet folder and one slide: conclusion, main limitation, link with the AMH Lo (2004); judge movement towards zero, not the sign
- **Model**: B1, B5 [Solved]
- Notebook: section C2

Many Predictors: Ridge and Cross-Validation

- With J predictors close to the number of observations, OLS fits noise: high in-sample R^2, poor forecasts
- **Ridge regression**: penalise large coefficients

$$\hat{\mathrm{h}}_\lambda = \arg\min_{\mathbf{h}} \sum_t (y_t - \mathrm{x}_t'\mathrm{h})^2 + \lambda\|\mathrm{h}\|^2 = (\mathrm{X}'\mathrm{X} + \lambda\mathrm{I})^{-1}\mathrm{X}'\mathrm{y}$$

  - h: the vector of the J coefficients; I: the $J \times J$ identity matrix
  - $\lambda \geq 0$: the penalty; $\lambda = 0$ is OLS, $\lambda \to \infty$ shrinks every coefficient to 0
  - predictors standardised first (mean 0, SD 1), so that the penalty treats them alike
- **Cross-validation**: choose λ by the prediction error on data left out of the estimation
  - leave-one-year-out: estimate without year y, predict year y, repeat for every year, average the errors
- **In-sample against out-of-sample** Martin & Nagel (2022)
  - with learning, in-sample tests reject efficiency although nothing is predictable in real time; only the out-of-sample result is implementable

C3 [Proposed]: Big Data and Efficiency, In and Out of Sample — Task (1/2)

- **Question**: with 238 predictors, is in-sample predictability of industry returns matched out of sample?
- **Data**
  - Case study of Chapter 2, Martin & Nagel (2022): with rational learning, in-sample tests reject efficiency although out-of-sample returns are unpredictable; replicate the design of Sec. 7 of their paper
  - 49 industry portfolios, Kenneth French Data Library, value-weighted monthly returns, July 1926 to the last month; missing values -99.99 and -999 dropped month by month; returns divided by 100
  - Dimensions: in month t, X_t is $N_t \times 238$ (returns and squared returns at lags 2–120), N_t the industries with a complete 120-month history; h_t the 238×1 coefficients
  - $r_{\mathbf{OOS},t+\mathbf{1}} = \frac{\mathbf{1}}{N_{t+\mathbf{1}}} h_t' X_{t+\mathbf{1}}' r_{t+\mathbf{1}}$ is the return of a long–short portfolio whose weights $X_{t+\mathbf{1}} h_t / N_{t+\mathbf{1}}$ are the demeaned industry forecasts, with h_t estimated on data up to t only
  - $r_{\mathbf{IS}}$: the same return with h estimated on a window that contains the evaluated month; the gap between the two measures how much of the predictability was not available in real time

Back Exercise list

C3 [Proposed]: Big Data and Efficiency, In and Out of Sample — Task (2/2)

CV: Cross-Validation.

- **Tasks**
  1. Demean the returns and each predictor across industries every month, scale each predictor to unit cross-sectional SD, give each month equal weight.
  2. Estimate pooled OLS h_t over the 240 months ending in t (data up to t only); compute $r_{\mathbf{OOS},t+\mathbf{1}} = \frac{\mathbf{1}}{N_{t+\mathbf{1}}} h_t' X_{t+\mathbf{1}}' r_{t+\mathbf{1}}$ and $r_{\mathbf{IS}}$ over the same window; state the first forecast month.
  3. Average $r_{\mathbf{OOS}}$ over 10-year windows with descriptive ± 2 SE bands (as in Fig. 8 of the paper).
  4. Fit a ridge from January 1971 with standardised predictors; choose the penalty on a log grid by leave-one-year-out CV (it may use later years: a validation, not a forecast); plot the coefficients by lag.
- **Report**: the chart of step 3, the means of $r_{\mathbf{IS}}$ and $r_{\mathbf{OOS}}$ in the first and in the second part of the period with SE, the coefficient plot, and why 49 portfolios can differ from thousands of stocks
- **Model**: B3, B8
- Notebook: section C3 (lecture Quantlet `MFM_ch2_big_data`)

Exercise Map

Exercise	Question	Type	Model
A1	what does $\widehat{\mathrm{VR}}(2)$ compute, line by line?	Solved	–
A2, A3	does an implied $H(20) \neq 0.5$ prove long memory?	Solved, Proposed	A2
B1	do the S&P 500 and BET reject RW3 with robust errors?	Solved	–
B2	is the S&P 500 reversal a crisis effect?	Proposed	B1
B3	is the turn-of-the-month effect real?	Solved	–
B4	which of the 15 calendar effects survive Holm?	Proposed	B3
C1	what is wrong in an AI variance-ratio answer?	Proposed	A1, B1
After the lecture	B5 [Solved], B6 [Proposed]: rolling DFA and rolling variance ratios with bootstrap null bands	homework	B1, B5
Extensions	A4–A6, B7–B9, C2, C3 (optional, after the seminar)	Solved, Proposed	on each task

- **[Solved]**: full solution in the slides and in the notebook, a model to copy
- **[Proposed]**: you solve it; the solution is discussed in class
- Part A: derivations on paper
- Part B: computation and inference on real data, each exercise ending with an interpretation question
- Part C: open questions and the critique of an AI answer

Further-Study Topics: What Each One Needs

Topic	Exercises	Prerequisites
1. Theory	A4 (null distribution), A5 [Solved] (runs), A6 (Stambaugh bias)	A1, B1
2. Many markets	B7	B1
3. Predictability	B8	A6, B3
4. Event study	B9	Event-study concepts; Chapter 2 bank-tax example; B3 for HAC. Optional: exercise B7 of Seminar 0
5. Project seeds (Part C)	C2, C3	B1, B5; the case study of Chapter 2

- Each exercise has its own notebook section

Acronyms used in this material (1/2)

- **AI**: Artificial Intelligence
- **AMH**: Adaptive Markets Hypothesis
- **AR(p)**: AutoRegressive model of order p.
- **AR**: Abnormal Return, in event studies.
- **BDS**: Brock–Dechert–Scheinkman–LeBaron (test)
- **BET**: Bucharest Exchange Trading (index)
- **BET-TR**: BET Total Return (index)
- **BIST**: Borsa İstanbul (BIST 100 index) — *the Istanbul Stock Exchange (BIST 100 index)*
- **BUX**: Budapesti Értéktőzsde index — *the Budapest Stock Exchange index*
- **BVB**: Bursa de Valori București — *the Bucharest Stock Exchange*
- **CD**: Chow–Denning (multiple variance-ratio test)
- **CI**: Confidence Interval
- **COVID-19**: COronaVIrus Disease 2019
- **CV**: Cross-Validation
- **DFA**: Detrended Fluctuation Analysis
- **ETF**: Exchange-Traded Fund
- **FRED**: Federal Reserve Economic Data
- **FTSE**: Financial Times Stock Exchange (FTSE Russell, index provider)

Acronyms used in this material (2/2)

- **FWER**: Family-Wise Error Rate
- **GARCH**: Generalized AutoRegressive Conditional Heteroskedasticity
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **IPC**: Índice de Precios y Cotizaciones — *the Mexican Stock Exchange index*
- **OLS**: Ordinary Least Squares
- **PX**: Prague Stock Exchange index (index Burzy cenných papírů Praha)
- **RW1**: Random Walk 1 (i.i.d. increments)
- **RW2**: Random Walk 2 (independent increments)
- **RW3**: Random Walk 3 (uncorrelated increments)
- **SD**: Standard Deviation
- **SE**: Standard Error
- **TOM**: Turn of the Month
- **US**: United States.
- **VR**: Variance Ratio
- **WIG20**: Warszawski Indeks Giełdowy 20 — *the Warsaw Stock Exchange index of the 20 largest companies*

References (1/5)

- Ariel, R.A. (1987). A monthly effect in stock returns. *Journal of Financial Economics*, 18(1), 161–174.
- Bartels, R. (1982). The rank version of von Neumann's ratio test for randomness. *Journal of the American Statistical Association*, 77(377), 40–46.
- Boehmer, E., Musumeci, J., Poulsen, A.B. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, 30(2), 253–272.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Brock, W.A., Dechert, W.D., Scheinkman, J.A., LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15(3), 197–235.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997). The Econometrics of Financial Markets. Princeton University Press.
- Campbell, J.Y., Thompson, S.B. (2008). Predicting excess stock returns out of sample: can anything beat the historical average? *Review of Financial Studies*, 21(4), 1509–1531.
- Choi, I. (1999). Testing the random walk hypothesis for real exchange rates. *Journal of Applied Econometrics*, 14(3), 293–308.
- Chow, K.V., Denning, K.C. (1993). A simple multiple variance ratio test. *Journal of Econometrics*, 58(3), 385–401.

References (2/5)

- Clark, T.E., West, K.D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1), 291–311.
- Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Fama, E.F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25(2), 383–417.
- Fama, E.F. (1991). Efficient capital markets: II. *Journal of Finance*, 46(5), 1575–1617.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- French, K.R. (2026). Kenneth R. French Data Library. Tuck School of Business, Dartmouth College.
- FTSE Russell (2019). FTSE Russell announces results of country classification review (press release, 26 September 2019).
- Harvey, C.R., Liu, Y., Zhu, H. (2016). . . . and the cross-section of expected returns. *Review of Financial Studies*, 29(1), 5–68.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.

References (3/5)

- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1), 770–799.
- Kendall, M.G. (1954). Note on bias in the estimation of autocorrelation. *Biometrika*, 41(3–4), 403–404.
- Kim, J.H. (2006). Wild bootstrapping variance ratio tests. *Economics Letters*, 92(1), 38–43.
- Kolari, J.W., Pynnönen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies*, 23(11), 3996–4025.
- Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17(3), 1217–1241.
- Lakonishok, J., Smidt, S. (1988). Are seasonal anomalies real? A ninety-year perspective. *Review of Financial Studies*, 1(4), 403–425.
- Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Lo, A.W. (1991). Long-term memory in stock market prices. *Econometrica*, 59(5), 1279–1313.
- Lo, A.W. (2004). The adaptive markets hypothesis. *Journal of Portfolio Management*, 30(5), 15–29.

References (4/5)

- Lo, A.W., MacKinlay, A.C. (1988). Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies*, 1(1), 41–66.
- MacKinlay, A.C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13–39.
- Martin, I.W.R., Nagel, S. (2022). Market efficiency in the age of big data. *Journal of Financial Economics*, 145(1), 154–177.
- 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. (2012). Modelling stock market crashes: the case of Bucharest Stock Exchange. *Procedia – Social and Behavioral Sciences*, 58, 533–542.
- Peng, C.-K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2), 1685–1689.
- Romano, J.P., Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4), 1237–1282.
- Stambaugh, R.F. (1999). Predictive regressions. *Journal of Financial Economics*, 54(3), 375–421.
- Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, 148, 80–82.

References (5/5)

- Wald, A., Wolfowitz, J. (1940). On a test whether two samples are from the same population. *Annals of Mathematical Statistics*, 11(2), 147–162.
- Welch, I., Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, 21(4), 1455–1508.