








# Modelling Financial Markets

## Seminar 0: Financial Markets in 2026 — 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

## How Do Prices Become Growth and Risk?

- Question of the seminar
  - **how do we turn a price series into returns, growth and risk, measure the largest loss from a peak (drawdown) and check that the prices are right?**
- Deliverables
  - the answers to the A exercises
  - one figure or table per B exercise
  - one interpretation sentence per exercise
- Assessment criteria
  - correct calculations, with the units and the annualisation convention stated
  - data checked against an independent source before any calculation
  - one interpretation for every result
  - the critique of the AI (artificial intelligence) answer, documented in `AI_ERRORS.md`
- Notebook for today: chapter0_seminar_notebook.ipynb in Colab
  - first step: Setup (save a copy, load the data); each task names its notebook section
- The exercises, their types and their order: the exercise map
- Seminar: Prof. Daniel Traian Pele, danpele@ase.ro; Antoaneta Amza, doctoral candidate, antoaneta.amza@csie.ase.ro

# Exercise Map

| Exercise | Question | Type | Model |
|---|---|---|---|
| **Core** (in class) | | | |
| A1–A4 | returns, compound annual growth, annualised volatility, drawdown on paper | Solved | – |
| B1 | is one EUR/RON quote a real price? | Solved | – |
| B2 | how did the SPY fund, gold and Bitcoin do since 2015? | Solved | – |
| A5, A6 | your turn: the same calculations on new numbers | Proposed | A1–A4 |
| **Homework** | | | |
| B3, B4 | two more quotes; three other assets | Proposed | B1, B2 |
| B5 | how deep and long are drawdowns? | Proposed | A4, B2 |
| C1 | what is wrong in an AI answer? | Proposed | A2–A3, B2 |
| **After the lecture** (optional, at the end of this PDF) | | | |
| Packs 1–8 | full data audit, calendars and correlations, precision of means and Sharpe ratios, extensions | Solved or Proposed (marked on each) | named on each task |

- **[Solved]**: full solution in the slides (core) or in the notebook (after the lecture), a model to copy; **[Proposed]**: you solve it, the solution is discussed in class.

## Data Used Today and in the Homework

- ETF (exchange-traded fund): a fund whose shares trade like a stock; SPY (*SPDR S&P 500 ETF Trust*) holds the stocks of the S&P 500 (Standard & Poor's 500), RSP (*Invesco S&P 500 Equal Weight ETF*) the same stocks equally weighted, TLT (*iShares 20+ Year Treasury Bond ETF*) American Treasury bonds over 20 years

| Series | Price used | Currency | Calendar | Exercise |
|---|---|---|---|---|
| S&P 500 | SPY, adjusted close | USD | US trading days | B2, B4 |
| Gold | spot gold, close | USD | weekdays only | B2 |
| Bitcoin | BTC/USD, close | USD | daily, 24:00 Coordinated Universal Time (UTC) | B2 |
| RSP, TLT | ETFs, adjusted close | USD | US trading days | B4 |
| EUR/RON, official | National Bank of Romania (BNR), reference rate | RON per EUR | Romanian working days, 13:00 | B1, B3 |
| EUR/RON, market | EODHD (EOD Historical Data) series, close | RON per EUR | every day with a quote | B1, B3 |

- Daily prices of the course data; sample 2 Jan 2015 – 18 Sep 2026
- Adjusted close: the closing price corrected for dividends and splits, so that its changes are total returns

## Setup (1/2): Open the Notebook and Save a Copy

- Colab: Google's free notebook service; Python runs on Google's servers, nothing to install, a Google account is enough
- **Tasks**
  1. Open the seminar notebook in Colab.
  2. Choose *File* → *Save a copy in Drive*, or run the Drive-save cell under the banner: without a copy your changes are lost when the tab closes.
  3. Run the cells of the section *Setup* in order (Shift + Enter); they import the libraries and define the data loader.
  4. If a cell fails, choose *Runtime* → *Restart session* and run again from the top.
- **Report**: nothing; you are ready when the setup check of the next slide prints the expected numbers

# Setup (2/2): Load the Data

```
# section "Setup" defines load(name, start):
# one daily series by its name,
# with the right price field
spy = load('SPY', start='2015-01-02')
bnr = load('EUR/RON', start='2015-01-01')
mkt = load('EUR/RON (market)',
           start='2015-01-01')
print(len(spy), spy.index[0].date(),
      spy.index[-1].date())
print(len(bnr), len(mkt))
```

- `load` calls `fetch_daily(symbol, source, field)`; SERIES in the notebook lists every name

- Expected: 2 945 SPY prices, 2 Jan 2015 to 18 Sep 2026; 2 939 BNR (Banca Națională a României, the National Bank of Romania) rates; 3 115 market quotes

| date | close | adjusted |
|---|---|---|
| 2015-01-02 | 205.43 | 169.26 |
| 2015-01-05 | 201.76 | 166.24 |
| 2015-01-06 | 199.82 | 164.64 |

- The two SPY columns differ: dividends are in the adjusted close only
- If your count or first value differs, stop and check the name and the start date

## Prices and Returns: Definitions (1/2)

- $P_t$: the price on day $t$ (here the close, or the adjusted close for ETFs)
- **Simple return** and **log return** of day $t$

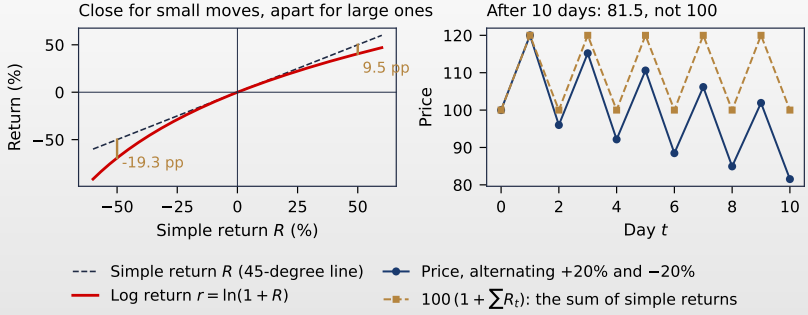
$$R_t = \frac{P_t}{P_{t-1}} - 1, \qquad r_t = \ln \frac{P_t}{P_{t-1}} = \ln(1 + R_t)$$

  - $R_t$: the percentage gain of one day, what an investor earns
  - $r_t \approx R_t$ for small moves; $r_t < R_t$ for any $R_t \neq 0$, and the gap grows with the size of the move
- Over several days: simple returns **compound**, log returns **add**

$$\frac{P_n}{P_0} - 1 = \prod_{t=1}^{n}(1 + R_t) - 1, \qquad \ln \frac{P_n}{P_0} = \sum_{t=1}^{n} r_t$$

- Worked example, two days
  - $P_0 = 100$, $P_1 = 102$: $R_1 = 102/100 - 1 = 2.00\%$, $r_1 = \ln 1.02 = 1.98\%$
  - $P_2 = 96.9$: $R_2 = 96.9/102 - 1 = -5.00\%$, $r_2 = \ln 0.95 = -5.13\%$

# Prices and Returns: Definitions (2/2) — Simple and Log Returns on a Chart



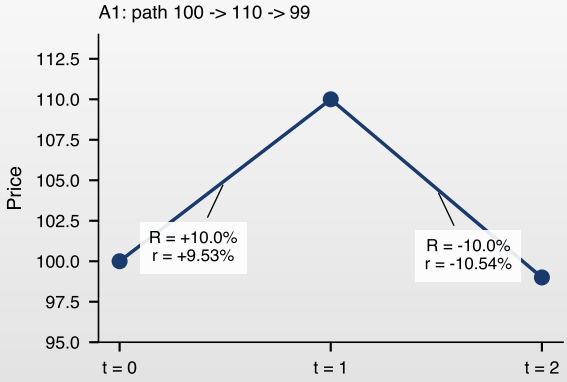

- Left: $r = \ln(1+R)$ against $R$; the gap is 9.5 percentage points (pp) at $R = +50\%$ and 19.3 pp at $R = -50\%$
- Right: a price that alternates $+20\%$ and $-20\%$ for 10 days (simulated)
  - the sum of the simple returns is 0, but the price ends at 81.5: $100 \times (1.2 \times 0.8)^{\mathbf{5}} = 81.5$
  - the sum of the log returns, $5(\ln 1.2 + \ln 0.8) = -20.4\%$, equals $\ln(81.5/100)$

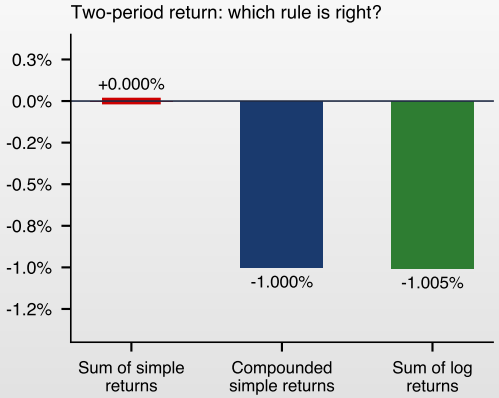
# A1 [Solved]: Adding Returns — Task

- **Question**: can two daily returns simply be added to get the two-day return?
- **Inputs**: prices $P_0 = 100$, $P_1 = 110$, $P_2 = 99$; no data file
- **Tasks**
  1. Compute the simple returns $R_1, R_2$ and the log returns $r_1, r_2$.
  2. Add the two simple returns and the two log returns.
  3. Compute the two-period returns $P_2/P_0 - 1$ and $\ln(P_2/P_0)$, and compare them with the sums of step 2.
- **Report**: a table with columns day $t$, price, simple return, log return, and one sentence on which returns add up
- Notebook: section A1–A4

Back Exercise list

# A1 [Solved]: The Price Path and Its Returns




MFM_ch0_seminar

# A1 [Solved]: Adding Log Returns, Compounding Simple Returns

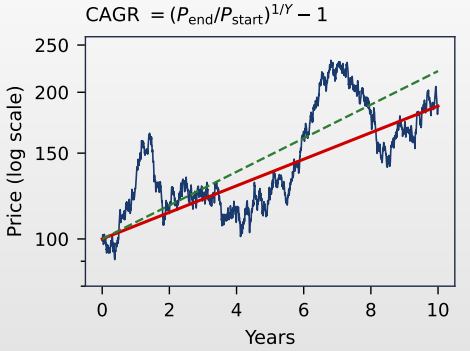
| $t$ | $P_t$ | $R_t$ | $r_t$ |
|---|---|---|---|
| 0 | 100 | – | – |
| 1 | 110 | +10.00% | +9.53% |
| 2 | 99 | −10.00% | −10.54% |
| sum | | 0.00% | −1.005% |
| $0 \to 2$ | | −1.00% | −1.005% |

- Interpretation: adding simple returns reports 0% for a path that lost 1%

- Log returns add up: $\ln P_2 - \ln P_1 + \ln P_1 - \ln P_0 = \ln(P_2/P_0)$, because the intermediate term $\ln P_1$ cancels
- Simple returns compound: $(1 + R_1)(1 + R_2) - 1 = 1.1 \times 0.9 - 1 = -1\%$

MFM_ch0_seminar

# CAGR and Annualised Volatility: Definitions (1/3) — CAGR




- $Y$: the length of the sample in years

$$Y = \frac{\text{calendar days between the first and the last price}}{365.25}$$

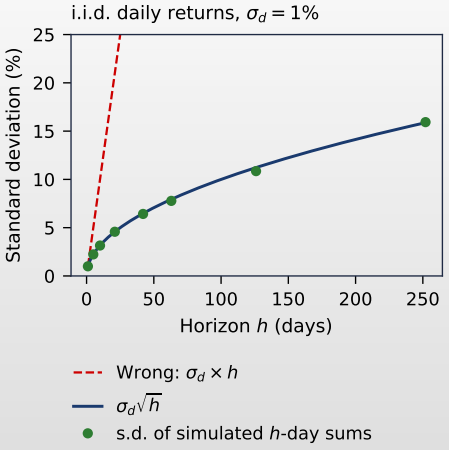
- **Growth multiple** $P_{\text{end}}/P_{\text{start}}$ and **CAGR** (compound annual growth rate)

$$\text{CAGR} = \left(\frac{P_{\text{end}}}{P_{\text{start}}}\right)^{1/Y} - 1 = \exp\left(\frac{\ln(P_{\text{end}}/P_{\text{start}})}{Y}\right) - 1$$

  - the constant yearly growth that turns $P_{\text{start}}$ into $P_{\text{end}}$ (red line)
  - worked number: 100 becomes 200 in 10 years: $2^{1/10} - 1 = 7.18\%$, not $100\%/10 = 10\%$

- Chart: a simulated price over 10 years: CAGR 6.5%
  - growing at the mean simple return (8.3% a year, dashed) overshoots the final price: CAGR is geometric growth, not the average return

# CAGR and Annualised Volatility: Definitions (2/3) — Volatility




- **Daily volatility**: the sample standard deviation of the $n$ daily log returns

$$\hat{\sigma}_d = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2}, \qquad \bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_t$$

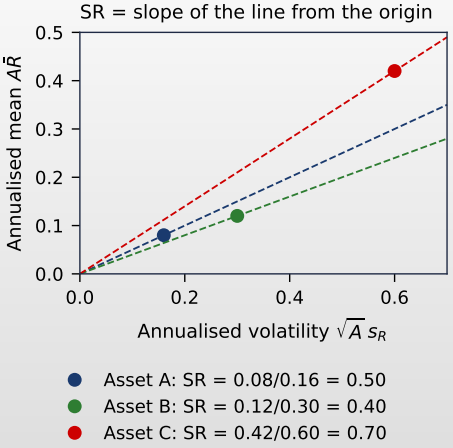
- **Return periods per year** $A = n/Y$: about 252 for US (United States) stocks, 260 for gold (weekdays), 365 for Bitcoin (every day)
- **Annualised volatility**

$$\hat{\sigma}_A = \hat{\sigma}_d\sqrt{A}$$

  - worked number: $\hat{\sigma}_d = 1.2\%$, $A = 252$: $1.2\% \times 15.87 = 19.05\%$

- $\sqrt{A}$, not $A$: for uncorrelated returns with equal variance, $\mathrm{Var}(r_1 + \cdots + r_h) = h\,\sigma_d^2$, so the s.d. of an $h$-day return is $\sigma_d\sqrt{h}$
  - chart: simulated sums of $h$ daily returns lie on $\sigma_d\sqrt{h}$; one year: 15.9%, not 252%

# CAGR and Annualised Volatility: Definitions (3/3) — Mean and Sharpe Ratio




- **Annualised mean** of the $n$ daily simple returns

$$\bar{R} = \frac{1}{n}\sum_{t=1}^{n} R_t, \qquad \hat{\mu}_A = A\,\bar{R}$$

  - the mean scales with $A$, the standard deviation with $\sqrt{A}$
- **Sharpe ratio** (SR; risk-free rate 0): the annualised mean per unit of annualised volatility

$$\mathrm{SR} = \frac{A\,\bar{R}}{\sqrt{A}\,s_R} = \sqrt{A}\,\frac{\bar{R}}{s_R}$$

  - $s_R$: the sample s.d. of the daily simple returns (divisor $n-1$)
  - reading: $\mathrm{SR} = 0.5$ means 0.5 pp of mean return per pp of volatility; a negative mean gives $\mathrm{SR} < 0$
  - chart: the slope of the line from the origin to each (hypothetical) asset; the steepest line has the best return per unit of risk

# A2–A3 [Solved]: Annualisation and CAGR — Task

- **Question**: how do we turn daily numbers into annual ones?
- **A2 inputs**: daily volatility $\sigma_d = 1\%$; $A = 252$ return intervals a year for an equity index, $A = 365$ for Bitcoin
  1. Compute the annualised volatility $\sigma_d\sqrt{A}$ for both values of $A$.
  2. State the two assumptions under which $\sqrt{A}$ scaling is exact.
- **A3 inputs**: S&P 500 price index, multiplied by 3.72 between 2 Jan 2015 and 18 Sep 2026 (4 277 calendar days); over the same period SPY has a CAGR of 13.7%
  3. Compute $Y = 4\,277/365.25$ and the CAGR $= 3.72^{1/Y} - 1$.
  4. Explain why this CAGR is lower than SPY's CAGR.
- **Report**: two volatilities, one CAGR, two sentences

Back Exercise list

## A2–A3 [Solved]: Annual Growth and Volatility

- A2
  1. $\sqrt{252} = 15.87$, so $1\% \times 15.87 = 15.87\%$; $\sqrt{365} = 19.10$, so 19.10%.
  2. Exact when daily returns are uncorrelated and have constant variance; A9 (Pack 3, after the lecture) shows what happens otherwise.
- A3
  3. $Y = 4\,277/365.25 = 11.71$;
     $\text{CAGR} = \exp(\ln 3.72/11.71) - 1 = \exp(1.3137/11.71) - 1 = e^{0.1122} - 1 = 11.87\%$.
  4. The S&P 500 index is a price index: it excludes dividends; SPY's adjusted close reinvests them (13.7% a year).
- Interpretation: CAGR is geometric growth per year, not the average annual return; for the same data the arithmetic mean is higher (A8, Pack 3)

# Drawdown: Definitions (1/2)

- **Running peak**, **drawdown** and **maximum drawdown** (MDD)

$$M_t = \max_{s \le t} P_s, \qquad DD_t = \frac{P_t}{M_t} - 1 \le 0, \qquad \text{MDD} = \min_t DD_t$$

  - $M_t$: the highest price so far; $DD_t$: the loss of an investor who bought at that peak
  - MDD: the worst such loss in the sample; it comes with a peak date and a trough date (the day of the minimum)

- **Gain needed to recover**: after a loss $L = -DD_t$ the price is $M_t(1 - L)$

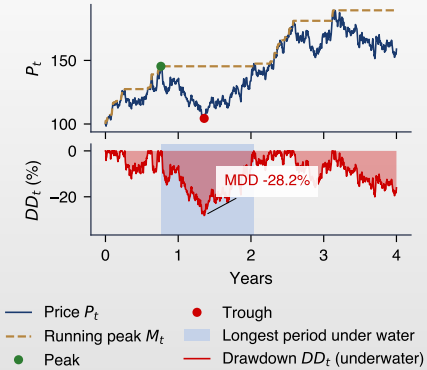
$$\frac{M_t}{M_t(1-L)} - 1 = \frac{L}{1-L}$$

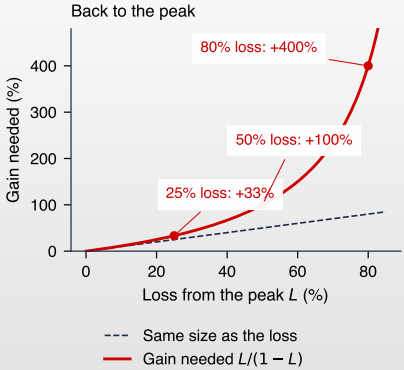
  - worked number: peak 120, later price 90: $DD = -25\%$; back to the peak needs $120/90 - 1 = 0.25/0.75 = +33.3\%$

- **Underwater chart**: $DD_t$ over time; 0 at every new peak, negative in between

- **Period under water**: from the day a peak is set to the first later close at or above it
  - a period still open at the end of the sample is **right-censored**: its true length is unknown, only a lower bound

# Drawdown: Definitions (2/2) — Drawdown and Recovery on a Chart




- Left: a simulated price over 4 years; MDD $-28.2\%$ (green peak to red trough); the longest period under water lasts 321 trading days
- Right: the gain needed grows faster than the loss: $-50\%$ needs $+100\%$, $-80\%$ needs $+400\%$

# A4 [Solved]: Drawdown — Task

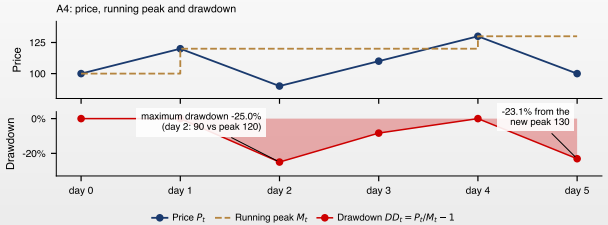
- **Questions**
  - what is the worst loss from a previous peak?
  - when did it happen?
- **Inputs**: prices on days 0–5: $100, 120, 90, 110, 130, 100$
- **Tasks**: complete the worksheet
  1. Running peak $M_t = \max_{s \le t} P_s$.
  2. Drawdown $DD_t = P_t / M_t - 1$, reported as a negative number.
  3. Maximum drawdown $\min_t DD_t$ and its day.

| Day | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P_t$ | 100 | 120 | 90 | 110 | 130 | 100 |
| $M_t$ | 100 | 120 | ... | ... | ... | ... |
| $DD_t$ | 0% | 0% | ... | ... | ... | ... |

- **Report**: the completed table and the maximum drawdown

Back Exercise list

# A4 [Solved]: Drawdown from Each Running Peak




| Day | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $P_t$ | 100 | 120 | 90 | 110 | 130 | 100 |
| $M_t$ | 100 | 120 | 120 | 120 | 130 | 130 |
| $DD_t$ | 0% | 0% | $-25.00\%$ | $-8.33\%$ | 0% | $-23.08\%$ |

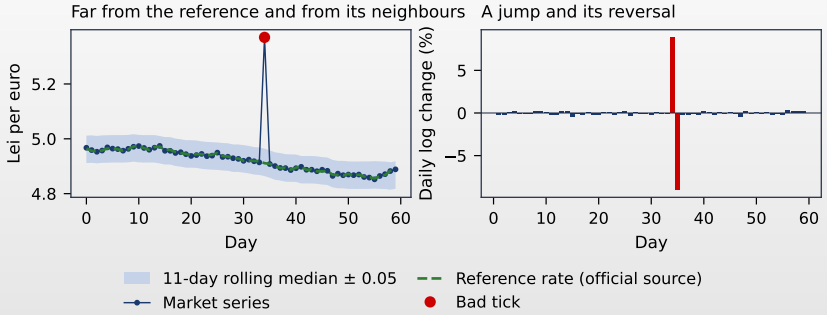
- Maximum drawdown $-25\%$ on day 2; the later fall to 100 is smaller, $-23.08\%$, because it is measured from the new peak of 130

MFM_ch0_seminar

# Bad Ticks and the BNR Reference Rate (1/2)

- **Bad tick**: a recorded price at which nobody traded, for example a typing error, a wrong decimal or a quote from another instrument
- **BNR reference rate**: the official rate of the euro (EUR) in Romanian lei (RON), published by the BNR every Romanian working day at about 13:00; an independent source to check a market series against
- The **EUR/RON series from EODHD (EOD Historical Data)**: aggregates dealer quotes
  - it is not an official rate
  - it is recorded at another time of day than the BNR rate
- Three signs of a bad tick
  1. a one-day move far larger than usual for the series;
  2. the move reverses the next day;
  3. the quote is far from the independent source of the same day.
- A gap to BNR alone is not enough: on a day of real movement the two series differ because they are recorded at different times
- One bad tick creates two huge returns, the jump and its reversal, and inflates the volatility

# Bad Ticks and the BNR Reference Rate (2/2): a Bad Tick on a Simulated Series

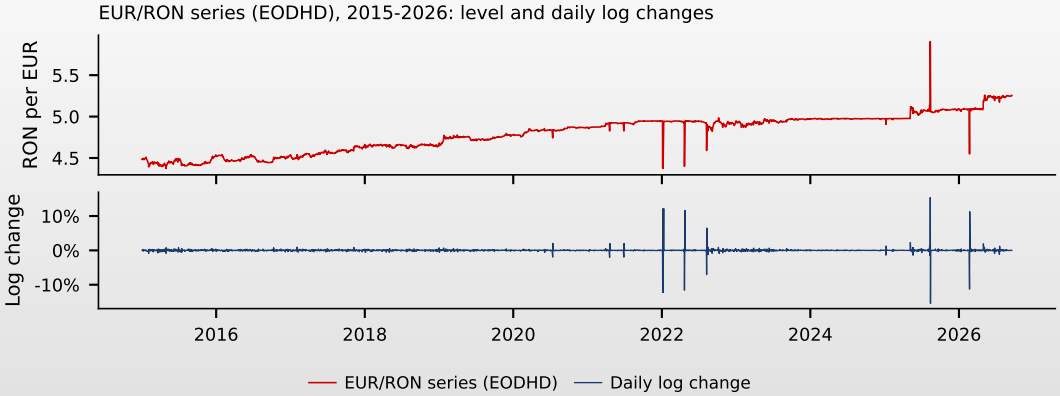



- Two checks of a quote $P_t$: against an independent reference rate $B_t$ and against its own neighbours

$$|P_t - B_t| > 0.05 \text{ lei}, \qquad |P_t - \mathrm{med}(P_{t-\mathbf{5}}, \ldots, P_{t+\mathbf{5}})| > 0.05 \text{ lei}$$

  - $\mathrm{med}$: the median of the 11 quotes centred on day $t$; one bad quote cannot move it
- Right: the bad price gives a jump of $+8.9\%$ and a reversal of $-9.0\%$; annualised volatility 26.4% with it, 2.3% without it

# B1 [Solved]: One Suspicious Quote — Task (1/2)




MFM_ch0_seminar

# B1 [Solved]: One Suspicious Quote — Task (2/2)

- **Question**: is the market EUR/RON quote of 13 Aug 2025 a real price?
- **Data**: EUR/RON series from EODHD (close), 1 Jan 2015 – 18 Sep 2026, 3 115 rows; BNR reference rate
- **Tasks**
  1. Compute the daily log changes of the market series and find the largest one in absolute value.
  2. For 13 Aug 2025, report the previous quote, the quote and the next quote, and the two log changes.
  3. Compare the quote with the BNR rate of 13 Aug 2025, in RON and in percent.
  4. Compute the annualised volatility of the market series, of the same series without this quote and of the BNR rate, each with its own $A$.
- **Report**: the three quotes, the gap to BNR and the three volatilities
- Look first: spikes that reverse the next day are the suspects
- Notebook: section B1
- **Interpretation**: which of the three signs are present?

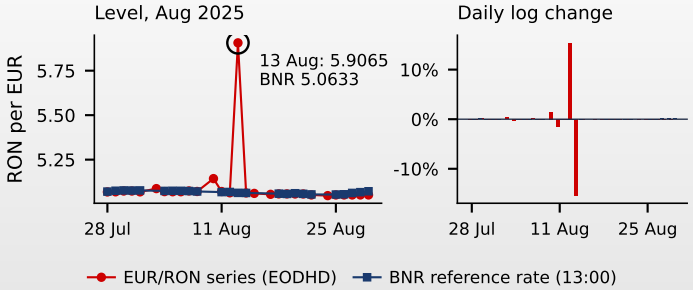
MFM_ch0_seminar

Back Exercise list

## B1 [Solved]: Solution (1/2) — Step by Step

- Step 1: the largest daily log change is $-15.41\%$, on 14 Aug 2025; the second largest, $+15.40\%$, is on 13 Aug 2025
- Step 2: 12 Aug 5.0636 $\rightarrow$ 13 Aug 5.9065 $\rightarrow$ 14 Aug 5.0629
  - $\ln(5.9065/5.0636) = +15.40\%$, then $\ln(5.0629/5.9065) = -15.41\%$: the move reverses the next day
- Step 3: BNR on 13 Aug 2025: 5.0633; gap $5.9065 - 5.0633 = 0.843$ RON, 16.7% above the official rate
- Step 4: annualised volatility
  - market series 12.4% ($A = 265.9$, it has some weekend quotes); without the quote of 13 Aug 2025: 10.6%
  - BNR rate 2.1% ($A = 251.1$)

# B1 [Solved]: Solution (2/2) — Chart and Interpretation

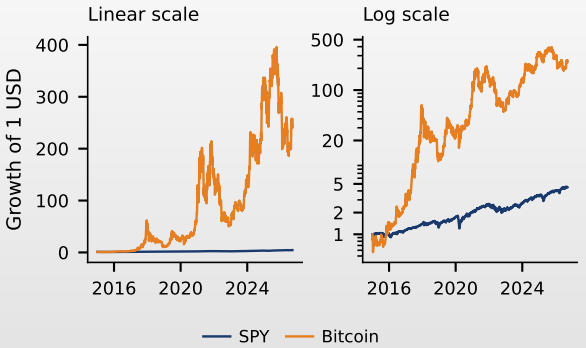



- All three signs are present: a 15% move where the BNR rate moves about 0.13% a day ($2.1\%/\sqrt{251}$), reversed the next day, 0.84 RON away from BNR: a bad tick
- One quote adds 1.8 pp to the annualised volatility; the market series is still five times as volatile as the BNR rate: other bad ticks remain, the full audit is Pack 1
- Interpretation: check a market series against an official source before computing anything; for EUR/RON use the BNR rate

MFM_ch0_seminar

# Growth of 1 USD and the Log Scale

- **Growth of 1 USD** (US dollar): $G_t = P_t / P_0$, the value on day $t$ of 1 USD invested on day 0; it puts assets with very different prices on one chart
- **Log scale**: the vertical axis shows $\ln G_t$ with the labels of $G_t$
  - equal distances are equal ratios: $1 \to 2$ and $100 \to 200$ have the same height
  - a constant growth rate is a straight line
- On a linear scale (left) Bitcoin's 257-fold growth flattens SPY into a line at the bottom



# B2 [Solved]: Three Assets on One Chart — Task

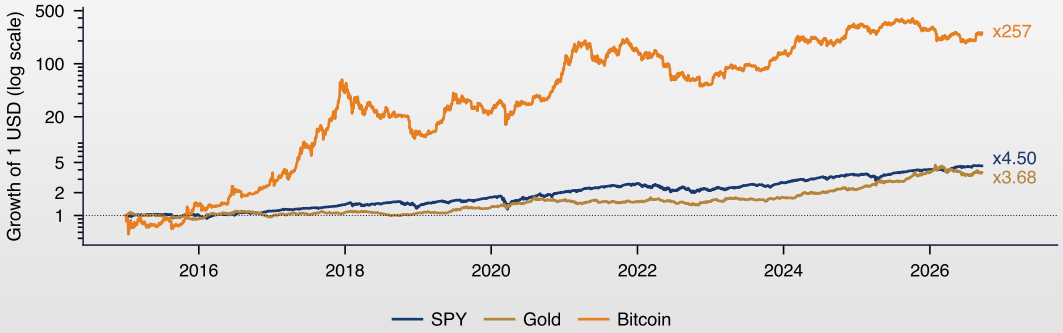
- **Questions**
  - how much did 1 USD grow in SPY, gold and Bitcoin since 2015?
  - how volatile was each?
  - how deep was the worst fall?
- **Data**: SPY (adjusted close), gold (close, weekdays), Bitcoin (close, every day), 2 Jan 2015 – 18 Sep 2026, all in USD; each asset on its own calendar
- **Tasks**
  1. Plot the growth of 1 USD, $P_t/P_0$, on a log scale.
  2. For each asset compute the multiple, $Y$ and the CAGR.
  3. Compute $n$, $A = n/Y$, the daily volatility $\hat{\sigma}_d$ and the annualised volatility $\hat{\sigma}_d\sqrt{A}$.
  4. Plot the drawdown curves and report each maximum drawdown with its peak and trough dates.
- **Report**: the two charts and a table (multiple, CAGR, $A$, volatility, MDD)
- **Interpretation**: does the asset with the highest CAGR also have the highest volatility and the deepest drawdown?
- Notebook: section B2

Back Exercise list

## B2 [Solved]: Solution (1/4) — SPY Step by Step

- Prices: 169.26 on 2 Jan 2015, 761.69 on 18 Sep 2026; $Y = 4\,277/365.25 = 11.71$
- Step 2: multiple $761.69/169.26 = 4.50$; CAGR $= 4.50^{1/11.71} - 1 = e^{1.504/11.71} - 1 = 13.7\%$
- Step 3: $n = 2\,944$ returns, $A = 2\,944/11.71 = 251.4$; $\hat{\sigma}_d = 1.108\%$; volatility $1.108\% \times \sqrt{251.4} = 17.6\%$
- Step 4: running peak 307.63 on 19 Feb 2020, trough 203.91 on 23 Mar 2020:
  $\text{MDD} = 203.91/307.63 - 1 = -33.7\%$
- The same code runs for gold and Bitcoin: only the name changes; $A$ follows each calendar (260.5 for gold, 365.25 for Bitcoin)

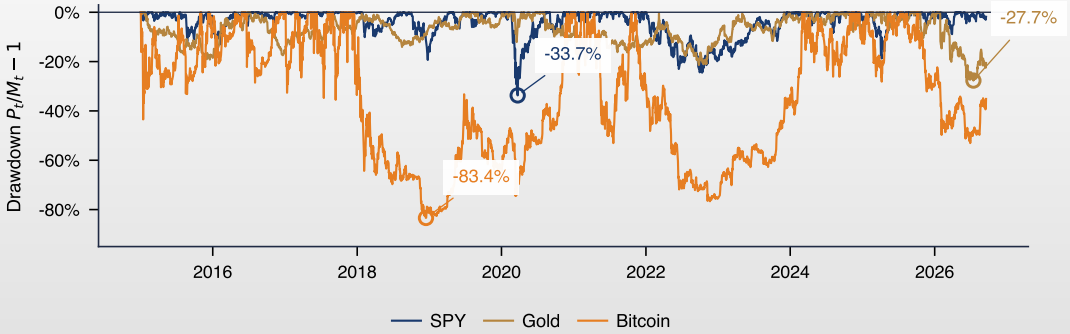
# B2 [Solved]: Solution (2/4) — Growth of 1 USD




- Log scale: Bitcoin's falls of 2018 and 2022 are as visible as its rises; end labels are the growth multiples

MFM_ch0_seminar

# B2 [Solved]: Solution (3/4) — Drawdowns




- Circles mark each maximum drawdown; on 18 Sep 2026 SPY is 1.8% below its peak, gold 20.6%, Bitcoin 35.2%

MFM_ch0_seminar

# B2 [Solved]: Solution (4/4) — Table and Interpretation

| | Multiple | CAGR | $A$ | Volatility | MDD (peak → trough) |
|---|---|---|---|---|---|
| SPY | 4.50 | 13.7% | 251.4 | 17.6% | −33.7% (19 Feb 2020 → 23 Mar 2020) |
| Gold | 3.68 | 11.8% | 260.5 | 16.3% | −27.7% (28 Jan 2026 → 16 Jul 2026) |
| Bitcoin | 256.8 | 60.6% | 365.25 | 66.7% | −83.4% (16 Dec 2017 → 15 Dec 2018) |

- Yes in this sample: Bitcoin has the highest CAGR, about 3.8 times the volatility of SPY and 4.1 times that of gold, and lost 83% from its 2017 peak
- SPY and gold have similar volatility; gold's worst fall is recent and not yet recovered
- One sample of 11.7 years: whether such differences are reliable is a question for the lecture and Pack 2
- **What do you think**: if we extend the SPY sample back before 2015, can the maximum drawdown become smaller than 33.7% in absolute value?
  - **Answer**: no: the added days can only raise the running maximum or add new troughs, so the MDD can only get deeper
  - hence we compare maximum drawdowns only over the same window

# A5 [Proposed]: Returns, CAGR and Volatility — Task

- **Model**: A1 and A2–A3 [Solved]
- **Question**: do the rules of A1 and A2–A3 hold on new numbers?
- **Inputs**: a stock at $P_0 = 50$, $P_1 = 55$, $P_2 = 44$; a fund that grew from 100 to 250 in 2 922 calendar days; a daily volatility of 1.5%
- **Tasks**
  1. Compute $R_1, R_2, r_1, r_2$, the two sums and the two-day returns $P_2/P_0 - 1$ and $\ln(P_2/P_0)$.
  2. Compute $Y$ and the CAGR of the fund.
  3. Annualise the daily volatility with $A = 252$ and with $A = 365$.
- **Report**: a small table and one sentence on which sum equals the two-day return

Back Exercise list

# A6 [Proposed]: A Drawdown on Paper — Task

- **Model**: A4 [Solved]
- **Questions**
  - when was the investor furthest below the previous peak?
  - when did the price recover?
- **Inputs**: prices on days 0–5: $200, 180, 220, 165, 198, 230$
- **Tasks**
  1. Compute the running peak $M_t$ and the drawdown $DD_t$ for every day.
  2. Report the maximum drawdown and its day.
  3. Recovery after the trough:
     - on which day is the price back at or above the peak before the trough?
     - what percentage gain from the trough brings the price exactly back to the previous peak?
- **Report**: the completed table, the MDD and the recovery day with the gain needed

Back Exercise list

# B3 [Proposed]: Two More Suspicious Quotes — Task

- **Model**: B1 [Solved]
- **Question**: the market EUR/RON quotes of 22 Apr 2022 and of 6 May 2025 are both more than 0.05 RON from the BNR rate; which one is a bad tick?
- **Data**: EUR/RON series from EODHD and BNR reference rate, as in B1
- **Tasks**
  1. For each date, report the previous quote, the quote and the next quote, and the two daily log changes.
  2. Compare each quote with the BNR rate of that day; if BNR published no rate that day (a public holiday), use its last earlier rate.
  3. Apply the two numeric checks: distance from BNR and from the centred 11-quote median, each above 0.05 RON. Relate them to the three signs of a bad tick, then decide: remove or keep.
  4. Plot both episodes, market series against BNR, over about three weeks.
- **Report**: one table row per date (quotes, changes, BNR, gap, signs present, decision) and the chart
- **Interpretation**: why is a gap to BNR alone not enough to remove a quote?

Back Exercise list

## B4 [Proposed]: Three Other Assets — Task

- **Model**: B2 [Solved]
- **Question**: how did the S&P 500 with equal weights (RSP) and long US Treasuries (TLT) do against SPY since 2015?
- **Data**: SPY, RSP and TLT (adjusted close), 2 Jan 2015 – 18 Sep 2026, USD, US trading days
- **Tasks**
  1. Plot the growth of 1 USD of the three ETFs on a log scale.
  2. Compute the multiple, the CAGR, $A$ and the annualised volatility of each.
  3. Plot the drawdown curves and report each MDD with its peak and trough dates and the drawdown on 18 Sep 2026.
- **Report**: the two charts and the table of B2 for the three ETFs
- **Interpretation**: are long Treasuries a safe asset in this sample?

Back Exercise list

# B5 [Proposed]: Drawdowns and Recovery Time — Task

- **Model**: A4 and B2 [Solved]
- **Questions**
  - how deep are the drawdowns of the S&P 500 and of Bitcoin?
  - how long are they?
- **Data**: S&P 500 (close) from 3 Jan 2000 and Bitcoin (close) from 17 Sep 2014, both to 18 Sep 2026
- **Conventions**
  - a spell starts on the last day the running peak was set (equal peaks: the latest) and ends on the first close at or above that peak
  - a spell still open on 18 Sep 2026 is right-censored
- **Tasks**
  1. Compute $DD_t$ for both assets and plot the two underwater curves.
  2. Report the maximum drawdown with its peak, trough and recovery dates and prices.
  3. Report the longest spell (start, end, calendar and trading days) and the current drawdown.
  4. Repeat 2–3 for the S&P 500 from 17 Sep 2014, the Bitcoin horizon.
- **Report**: the underwater chart; for each asset and horizon, the MDD with its three dates and prices, the longest spell below the peak and the current drawdown
- **Interpretation**: why must the two MDDs be compared on the same horizon?

Back Exercise list

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

- **Prompt** sent to an AI assistant: "Using the daily prices of the course data, compute the annualised mean, the annualised volatility and the Sharpe ratio ($r_f = 0$) of SPY and Bitcoin from daily simple returns"
- **Answer**: the code below and the conclusion

```
import numpy as np, pandas as pd
URL = 'https://raw.githubusercontent.com/danpele/MFM/main/data/market/'
def load(sym, field='close'):
    return pd.read_csv(URL + sym + '.csv', index_col='date',
                       parse_dates=True)[field].dropna()
spy, btc = load('SPY.US'), load('BTC-USD.CC')
A = 252                                  # trading days per year
for name, p in [('SPY', spy), ('BTC', btc)]:
    r   = p.pct_change().dropna()        # daily simple returns
    mu  = r.mean() * A                   # annualised mean
    vol = r.std() * A                    # annualised volatility
    print(name, mu, vol, mu / vol)       # Sharpe ratio, rf = 0
```

- **Conclusion of the AI**: "Sharpe ratios: SPY 0.03, Bitcoin 0.07. Both assets barely reward their risk."

Back Exercise list

# C1 [Proposed]: Critique an AI Answer — Design before Code

- **Model**: A2–A3 and B2 [Solved]
- **Sharpe ratio**: the definition in CAGR and Annualised Volatility: Definitions (3/3)
- **Data**: the full history, SPY from 29 Jan 1993 and Bitcoin from 13 Jul 2010, both to 18 Sep 2026
- **Step 1, research design** (write it down before you run any code)
  1. State the hypothesis behind the AI's conclusion as a testable sentence about the two Sharpe ratios.
  2. Fix the sample (dates, price field, calendar of each asset) and the comparison you will make.
  3. Name one rival explanation for any difference you will find (for example, the different sample periods).
  4. Say what the result cannot establish, whatever the numbers turn out to be.
- **Report**: four sentences, written before the code

# C1 [Proposed]: Find and Log the Planted Errors

- **Step 2, the code**
  5. Find the three implementation errors planted in the code.
  6. For each, say how you would detect it: a check on the output, a published number, the lecture or the documentation.
  7. Check the corrected code on 2 Jan 2015 – 18 Sep 2026 against B2; then report the annualised mean, volatility and Sharpe ratio on each asset's full history. Interpret differences in light of the sample dates.
  8. Check the corrected result against Step 1:
     - is your rival explanation ruled out?
     - which limitation remains even with the code fixed?
  9. Log every error you actually found, in the format of `AI_ERRORS.md`: request, answer, error, how found, correction.
- **Report**: the three errors with how each was detected, the mean / volatility / SR table for the two samples, the answers to task 8, the `AI_ERRORS.md` entries
- **Interpretation**: does the conclusion "both assets barely reward their risk" survive?
- AI code usually runs without an error message: the mistakes are in the finance, not in the syntax

# After the Lecture: Packs, Data and Notation

| Pack | Exercises | Prerequisites |
|---|---|---|
| 1. Data and calendars | B1 Full Audit, B2 Six Assets, B6 [all Solved] | core B1 and B2 |
| 2. Precision | A7, B2 Extended [Solved] | Pack 1, the Chapter 0 lecture |
| 3. Return arithmetic | A8 [Solved], A9, A10 | Pack 2 |
| 4. Correlation inference | B6 Extended, A11, A12, C2 | Packs 1–2 |
| 5. Drawdown extremes | B5 Extended | B5, Pack 2 |
| 6. Romanian market | B7 [Solved], B8, B9, B9 Extended, B1 Extended | Pack 1, linear regression |
| 7. Stablecoins | B10, B10 Extended | A2–A3, B7 |
| 8. Passive investing lab | C3 | the lecture case study, Pack 2 |

- For after the Chapter 0 lecture, which introduces the standard errors, the bootstrap and the tests used here; one notebook section per pack, in this order; C2 and C3 are possible project starts
- Data (course data, daily, 2 Jan 2015 – 18 Sep 2026 unless stated)
  - SPY and TLT (adjusted close, total return)
  - S&P 500 index, gold (weekdays), Bitcoin (every day, 24:00 UTC), Euro Stoxx 50 and Nikkei 225 (close, converted to USD with FRED (Federal Reserve Economic Data) rates, series `DEXUSEU`, `DEXJPUS`)
  - BNR EUR/RON and the EUR/RON series from EODHD
  - BET (Bucharest Exchange Trading index), BET-TR (dividends reinvested), Banca Transilvania, BRD–Groupe Société Générale
  - DefiLlama stablecoin supply
- Notation
  - $P_t$, $R_t$, $r_t$, $n$ (also $T$), $Y$, $A = n/Y$, CAGR and $\hat{\sigma}_d\sqrt{A}$ as in the core
  - Sharpe ratio $\mathrm{SR} = \sqrt{A}\,\bar{R}/s_R$, defined in the core
  - iid: independent and identically distributed
  - subscript $A$ marks an annualised quantity ($\hat{\mu}_A$, $\hat{\sigma}_A$, $\mathrm{SR}_A$), subscript $d$ a daily one

# Pack 1: B1 Full Audit [Solved]: All Bad Ticks

- **Task**
  - which quotes of the EUR/RON series from EODHD (close, 1 Jan 2015 – 18 Sep 2026, 3 115 rows) are errors?
  - how much do they distort volatility?
  - external check: the BNR reference rate

  1. Drop weekend rows.
  2. Flag weekday quotes more than 0.05 RON from the 11-day centred rolling median.
  3. Compare each flag with the BNR rate of that day (on a holiday, the last earlier fixing); remove only quotes more than 0.05 RON from BNR.
  4. Compute the annualised volatility of the raw market series, the cleaned series and the BNR rate, each with its own $A$.
- **Report**: an audit table (date, quote, median, BNR, gap, decision) and the three volatilities
- **Worked solution**: notebook section *B1 full audit*

# Pack 1: B2 Six Assets [Solved]: A Cross-Asset Dashboard

- **Task**
  - how much did 1 USD grow since 2015 in SPY, Euro Stoxx 50, Nikkei 225, gold, TLT and Bitcoin?
  - how volatile was each?
  - period: 2 Jan 2015 – 18 Sep 2026, all in USD
  1. Convert Euro Stoxx 50 and Nikkei 225 to USD: $P^{USD} = P^{EUR} \times$ (USD per EUR), $P^{USD} = P^{JPY}/$(JPY per USD), with the FRED rate of the same day or, if missing, the last rate at most 5 calendar days earlier.
  2. Plot the growth of 1 USD, $P_t/P_{\mathbf{0}}$, on a log scale.
  3. For each asset compute $n$, $Y$, $A = n/Y$, the multiple, the CAGR and the volatility $\hat{\sigma}_d\sqrt{A}$.
  4. Recompute each volatility with $A = 252$.
- **Report**
  - the chart and a table ($A$, multiple, CAGR, volatility, volatility with 252)
  - for which assets does the choice of $A$ matter?
  - why does it matter for them?
- **Worked solution**: notebook section *B2 six assets*

# Pack 1: B6 [Solved]: Correlations and Calendars (1/2)

- **Task**
  - how much do the calendar and the estimator change a correlation?
  - is the change of the stock–bond correlation larger than sampling error?
  - data: S&P 500 and Bitcoin (close), TLT (adjusted close), daily log returns; Bitcoin – S&P 500 in 1 Jan 2022 – 18 Sep 2026; S&P 500 – TLT in 2010–2020 ($n_1$ days) and 2022–2026 ($n_2$ days)
- **Calendars**
  - join the *prices* on common days first, then compute returns (Friday → Monday for both)
  - returns first and join later pairs the S&P 500 Friday → Monday return with Bitcoin's Sunday → Monday return and loses Bitcoin's weekend move
  - forward-filling the S&P 500 creates zero returns paired with Bitcoin's weekend moves

1. Bitcoin – S&P 500 whole-period correlation on common days (prices joined first).
2. The same with the S&P 500 forward-filled on weekends, and with weekly returns: last common-day price of each week ending on Friday (Thursday if Friday is a holiday); an incomplete final week is dropped.
3. The mean of the 252-day rolling correlation over 2022–2026 (windows may start in 2021).
4. S&P 500 – TLT in both periods, with both estimators.

Back Exercise list

# Pack 1: B6 [Solved]: Correlations and Calendars (2/2)

5. Compute $\hat{\rho}_{\mathbf{1}}$, $\hat{\rho}_{\mathbf{2}}$ unrounded, $z_i = \operatorname{atanh} \hat{\rho}_i$, the standard error (SE) $\mathrm{SE}_z = \sqrt{1/(n_{\mathbf{1}} - 3) + 1/(n_{\mathbf{2}} - 3)}$ and $z = (z_{\mathbf{2}} - z_{\mathbf{1}})/\,\mathrm{SE}_z$ (Fisher (1915); iid pairs from a bivariate Normal distribution).
6. Compute the iid 95% interval of $\rho_{\mathbf{2}} - \rho_{\mathbf{1}}$: $(\hat{\rho}_{\mathbf{2}} - \hat{\rho}_{\mathbf{1}}) \pm 1.96\sqrt{(1 - \hat{\rho}_{\mathbf{1}}^{\mathbf{2}})^{\mathbf{2}}/(n_{\mathbf{1}} - 3) + (1 - \hat{\rho}_{\mathbf{2}}^{\mathbf{2}})^{\mathbf{2}}/(n_{\mathbf{2}} - 3)}$, where $(1 - \hat{\rho}^{\mathbf{2}})^{\mathbf{2}}/(n - 3)$ is the approximate variance of one $\hat{\rho}$.

⊡ **Report**

- a plot of the rolling correlations; a table method / sample size / correlation; $z$, its $p$-value, the change and its interval
- which estimator answers "what was the correlation in 2022–2026"?
- what can the interval not establish?

⊡ **Worked solution**: notebook section *B6*

# Pack 2: A7 [Solved]: How Precise Is a Mean Return?

- **Task**
  - SPY returned about 14% a year in 2015–2026
  - how precisely does this sample pin down the expected annual return?
  - SPY (adjusted close), daily simple returns $R_t$, 2 Jan 2015 – 18 Sep 2026
  - $A$ and $Y$ as in B2 Six Assets

  1. Compute the annualised mean $\hat{\mu}_A = A\,\bar{R}$ and the annualised volatility $\hat{\sigma}_A = s_R\sqrt{A}$.
  2. With iid returns, $\mathrm{SE}(\hat{\mu}_A) = A\,s_R/\sqrt{n} = \hat{\sigma}_A/\sqrt{Y}$ (since $n = AY$): compute it and the 95% interval $\hat{\mu}_A \pm 1.96\,\mathrm{SE}$.
  3. How many years of data would give a 95% interval of $\pm 2$ pp with the same volatility?

- **Report**: $\hat{\mu}_A$, $\hat{\sigma}_A$, the SE, the interval and the number of years; why does sampling daily instead of monthly not make the mean more precise?

- **Worked solution**: notebook section *A7*

Back Exercise list

# Pack 2: B2 Extended [Solved]: How Precise Is a Sharpe Ratio?

- **Task**
  - is SPY's Sharpe ratio precise enough to rank it above gold?
  - the six series of B2 Six Assets, daily simple returns $R_t$, risk-free rate 0
- **Stationary bootstrap** (Politis & Romano (1994))
  - blocks from random start days, continued with probability $1 - 1/b$ (mean block $b = 20$), wrapping around, until $T$ days are drawn
  - each drawn day keeps the returns of all assets
  - $A$, $Y$ fixed
  - `default_rng(42)`
  1. Compute $\mathrm{SR}_A = \sqrt{A}\,\bar{R}/s_R$ for SPY and the daily $\mathrm{SR}_d = \bar{R}/s_R$.
  2. Compute the asymptotic iid SE for returns following the Normal distribution (Lo (2002)), $\mathrm{SE} = \sqrt{A(1 + \mathrm{SR}_d^2/2)/T}$, and the 95% interval $\mathrm{SR}_A \pm 1.96$ SE.
  3. Bootstrap (block 20, 2 000 resamples): 95% percentile interval of SPY's SR and of the other five.
  4. On the common days of SPY and gold, bootstrap the paired difference of their Sharpe ratios (both with $A = 252$) and report its 95% interval.
- **Report**: the SPY calculation, a chart of the intervals, the difference with its interval; does B2 Six Assets rank SPY above gold?
- **Worked solution**: notebook section *B2 extended*

# Pack 3: A8 [Solved]: Volatility Drag

- **Task**
  - can a positive average return go together with falling wealth?
  - daily $m = E[r]$, $a = E[R]$, $v_R = \mathrm{Var}(R)$, $v_r = \mathrm{Var}(r)$, $r = \ln(1 + R)$
  - Bitcoin close in USD, 17 Sep 2014 (457.33) to 18 Sep 2026 (80 901.46), 4 384 daily returns, 365 a year

  1. Expand $\ln(1 + R)$ to second order and show $m \approx a - (v_R + a^2)/2 \approx a - v_r/2$ for small daily returns.
  2. For Bitcoin, fill a daily-to-annual table: log mean $\times 365$, drag $v_r \times 365/2$, implied simple mean, CAGR $= e^{365m} - 1$; compare with the CAGR from the two endpoint prices.
  3. An asset has an annual arithmetic mean of 10% and a volatility of 60%: approximate its expected log growth and CAGR.
- **Report**: the expansion, the table, the two numbers of task 3; why can the typical path lose money when the mean return is positive?
- **Worked solution**: notebook section *A8*

Back Exercise list

# Pack 3: A9 [Proposed]: When the $\sqrt{A}$ Rule Fails

**Task**

- daily log returns, stationary, variance $\sigma^2$, autocorrelations $\rho_j$
- data: BNR EUR/RON, 4 Jul 2005 – 18 Sep 2026, 5 352 daily log changes (models: A2–A3, A8)

1. For $k = 3$, write $\mathrm{Var}(r_1 + r_2 + r_3)$ as the sum of the nine covariances; then generalise to $\mathrm{Var}(\sum_{t=1}^{k} r_t) = k\sigma^2[1 + 2\sum_{j=1}^{k-1}(1 - j/k)\rho_j]$.
2. For a *hypothetical* autoregressive AR(1) process, $\rho_j = \rho^j$, find the limit of the bracket as $k \to \infty$.
3. Compute $\hat{\rho}_1$, the $\sqrt{252}$ volatility and its hypothetical AR(1) correction.
4. Compute the autocorrelations at lags 1–10 and the volatility from month-end changes $\times\sqrt{12}$.
5. The raw EUR/RON series from EODHD (B1), bad ticks included, has $\hat{\rho}_1 = -0.57$:
   - what would the AR(1) factor be?
   - why is it the wrong fix?

**Report**

- the two derivations, the volatilities and the autocorrelations
- does positive autocorrelation hide risk or invent it?
- is $\hat{\rho}_1$ enough?

Back Exercise list

# Packs 3–4: A10 [Proposed]: The Standard Error of a Sharpe Ratio

**Task**

- $T$ iid daily returns with mean $\mu$, variance $\sigma^2$, finite central moments $\mu_3, \mu_4$, skewness $\gamma_3 = \mu_3/\sigma^3$, ordinary (not excess) kurtosis $\gamma_4 = \mu_4/\sigma^4$, daily $\mathrm{SR} = \mu/\sigma$
- SPY: $\mathrm{SR}_A = 0.82$, $A = 251.4$, $T = 2\,944$, $Y = 11.71$, $\gamma_3 = -0.31$, $\gamma_4 = 17.1$ (models: A8, A7, B2 Extended)

1. Gradient: for $g(\mu, \sigma^2) = \mu/\sigma$, compute $\nabla g$.
2. Quadratic form: $\sqrt{T}(\hat{\mu} - \mu, \hat{\sigma}^2 - \sigma^2) \to N(0, \Sigma)$, $\Sigma = \begin{pmatrix} \sigma^2 & \mu_3 \\ \mu_3 & \mu_4 - \sigma^4 \end{pmatrix}$; compute $\nabla g^\top \Sigma \nabla g$ and its Normal case (Lo (2002)).
3. Normal case, annual scaling: with $\mathrm{SR}_A = \sqrt{A}\,\mathrm{SR}_d$ and $T = AY$, show $\mathrm{Var}(\widehat{\mathrm{SR}}_A) \approx (1 + \mathrm{SR}_A^2/(2A))/Y$ (large-$T$ approximation, not an exact finite-sample result).
4. SPY's asymptotic SE three ways: Normal daily, general daily (with $\gamma_3$, $\gamma_4$), and the Normal formula with $A = 1$, $\sqrt{(1 + \mathrm{SR}_A^2/2)/Y}$, as if the span were observed as annual returns; state the sampling assumption of each.

**Report**: the derivations and the three SEs; why does sampling more often than monthly barely help?

# Packs 3–4: B6 Extended [Proposed]: The Stock–Bond Change in the Presence of Dependence

- **Task**
  - S&P 500 – TLT daily log returns on common days, as in B6: 2010–2020 ($n_1 = 2\,769$) and 2022 – 18 Sep 2026 ($n_2 = 1\,182$)
  - bootstrap of B2 Extended

1. Stationary bootstrap each period separately (block 20, 2 000 resamples, `default_rng(42)`): recompute $\hat{\rho}_2^* - \hat{\rho}_1^*$ on each pair of resamples.
2. Report the 95% percentile interval of $\rho_2 - \rho_1$ and plot the histogram of the bootstrap changes with both intervals.
3. Compare its width with the width of the iid interval of B6 (task 6).

- **Report**
  - the bootstrap interval, the width ratio and the histogram
  - why is the bootstrap interval wider?
  - does the conclusion of B6 survive?

# Pack 4: A11 [Proposed]: Why Is the Weekly Correlation Lower?

- **Task**
  - $x_t$ = S&P 500, $y_t$ = Bitcoin, daily log returns on common days, 2022 – 18 Sep 2026, $n = 1\,182$
  - $\gamma(j) = \mathrm{Cov}(x_t, y_{t-j})$
  - $\sigma_x$, $\sigma_y$: daily standard deviations
  - $\rho_{x,j}$: autocorrelation of $x$ at lag $j$
  - $\rho_{\mathbf{1}}$, $\rho_{\mathbf{5}}$: correlation of daily and of 5-day returns (model: B6)

  1. For sums $X, Y$ of 5 consecutive common-day returns, draw the $5 \times 5$ grid of covariances and show $\mathrm{Cov}(X, Y) = 5\gamma(0) + \sum_{j=1}^{4}(5-j)[\gamma(j) + \gamma(-j)]$ and $\mathrm{Var}(X) = 5\sigma_x^2 f_x$, $f_x = 1 + 2\sum_{j=1}^{4}(1 - j/5)\rho_{x,j}$; hence $\rho_{\mathbf{5}} = \{\rho_{\mathbf{1}} + \sum_{j=1}^{4}(5-j)[\gamma(j) + \gamma(-j)]\}/(5\sigma_x\sigma_y)\}/\sqrt{f_x f_y}$.
  2. Check the shortcut $f_x = f_y = 1$: own autocorrelations of $x$ and $y$ at lags 1–4, and $f_x$, $f_y$.
  3. Compute the lag term, the implied $\rho_{\mathbf{5}}$ with and without $\sqrt{f_x f_y}$, and the correlation of overlapping 5-day sums.
  4. Compute the weekly correlation for weeks anchored on Monday, . . . , Friday, each with a Fisher 95% interval.
  5. Report the joint week-block bootstrap interval of daily minus weekly (notebook section B6 extended).

- **Report**: the derivation and the numbers of tasks 2–5; is the gap 0.42 vs 0.30 a property of the markets or of the sampling? the Epps (1979) effect (correlations shrink as the return interval shortens) predicts the opposite sign

Back Exercise list

# Pack 4: A12 [Proposed]: Asynchronous Closes

- **Task**
  - log prices $X(s) = W(s) + U(s)$ (Bitcoin), $Y(s) = W(s) + V(s)$ (S&P 500), $W, U, V$ independent Brownian motions with variance rates $\sigma_W^2, \sigma_U^2, \sigma_V^2$, $\rho = \sigma_W^2 / \sqrt{(\sigma_W^2 + \sigma_U^2)(\sigma_W^2 + \sigma_V^2)}$
  - $X$ closes at integer times $t$ (24:00 UTC), $Y$ a fraction $\delta$ of a day earlier (20:00 UTC in summer, 21:00 in winter)
  1. Draw the intervals covered by $x_t$, $y_t$ and $y_{t+\mathbf{1}}$ on a time line and show $\mathrm{Corr}(x_t, y_t) = \rho(1-\delta)$, $\mathrm{Corr}(x_t, y_{t+\mathbf{1}}) = \rho\delta$, all other cross-correlations 0.
  2. Deduce that the lag $-1$, 0, $+1$ correlations sum to $\rho$ (Scholes & Williams (1977)).
  3. Compute $\delta$ in summer and winter and the implied correlations for $\rho = 0.45$.
  4. Tabulate the observed $\mathrm{Corr}(x_t, y_{t+j})$, $j = -1, 0, 1$, for 2022–2026 and compare.
- **Report**: the derivation and the two tables; is the asynchronous-close bias important for this pair? Treat the sum as a diagnostic

# Packs 4–5: C2 [Proposed]: Did Spot ETFs Change Bitcoin's Correlation with Equities?

- **Task**
  - US spot Bitcoin ETFs started trading on 11 Jan 2024
  - did the Bitcoin – S&P 500 correlation change at that date?
  - daily log returns on common days, 2018 – 18 Sep 2026 (models: B6, B2 Extended)

  1. Correlations in 2018–2019, 2020–2021, 1 Jan 2022 – 10 Jan 2024 (pre) and 11 Jan 2024 – 18 Sep 2026 (post), each with a stationary-bootstrap 95% interval (block 20).
  2. Test $H_0 : \rho_{post} - \rho_{pre} = 0$ with a bootstrap interval of the difference; plot the rolling correlation with the ETF date.
  3. Project extensions (optional): a counterfactual asset without an ETF event (state why it is comparable), an unknown-date break search, other calendars and windows.

- **Report**: a table period / $n$ / $\hat{\rho}$ / interval, the difference, one figure, one limitation; a causal claim needs the extensions

# Packs 4–5: B5 Extended [Proposed]: Is the 2009 Drawdown Extreme?

- **Task**
  - S&P 500 daily log returns $r_t$, 3 Jan 2000 – 18 Sep 2026, $n = 6\,717$
  - your observed MDD and longest spell from B5, in trading days (open spells censored at the last day)
  - the resampled returns include the 2008–2009 crisis itself: the result is a diagnostic, not a forecast

  1. Build one path: resample $n$ returns, set $\ln P_t = \ln P_{\mathbf{0}} + \sum_{s \le t} r_s^*$ ($r_s^*$: the resampled returns), compute its MDD and longest spell; plot it.
  2. Repeat 5 000 times with the iid bootstrap and 5 000 times with the stationary bootstrap (block 20); seed 42.
  3. Plot both distributions of the MDD and of the longest spell with the observed values; report the medians, the adverse-tail quantiles (5% for the MDD, 95% for the spell) and the tail shares $P(\mathrm{MDD}^* \le \mathrm{MDD}_{obs})$, $P(\text{spell}^* \ge \text{spell}_{obs})$.
  4. Compute $\hat{\rho}_{\mathbf{1}}$ and the variance ratio (VR), $\widehat{\mathrm{VR}}(20) = \widehat{\mathrm{Var}}(r_t + \cdots + r_{t-\mathbf{19}})/(20\,\widehat{\mathrm{Var}}(r_t))$ from all overlapping 20-day sums (divisor: number of terms $-1$, no small-sample correction).

- **Report**
  - the charts and the numbers of task 3, $\hat{\rho}_{\mathbf{1}}$ and $\widehat{\mathrm{VR}}(20)$
  - is the observed MDD extreme for this return distribution?

## Abnormal Returns and Their Cumulative Sum

- **Abnormal return** (AR) in an event study: the observed return minus the return predicted by the benchmark model.
- Here AR is not the autoregressive AR(1) model of A9.
- **Cumulative abnormal return** (CAR): the sum of the abnormal returns over the event window.
- If the observed return is 2% and the benchmark predicts 0.5%, the abnormal return is 1.5 percentage points.
- Abnormal returns of 1.5 and $-0.5$ percentage points give a CAR of 1 percentage point.
- CAR measures the cumulative departure from the benchmark; its interpretation depends on the model and event window.

# Pack 6: B7 [Solved]: Event Study, the Bank Tax of December 2018

- **Task**
  - how much did Banca Transilvania lose on the bank-tax news, beyond the market move?
  - the tax was announced on the evening of 18 Dec 2018, so day 0 is 19 Dec 2018
  - TLV (the ticker of Banca Transilvania, adjusted close) and the BET index (close), daily log returns
  - estimation window $[-250, -11]$ (18 Dec 2017 – 4 Dec 2018, $T = 240$), event window $[0, +5]$ ($L = 6$)
  - method: MacKinlay (1997)
  - errors iid Normal, independent of the market return, with the same variance $\sigma_\varepsilon^2$ in both windows and independent across windows

  1. Estimate $r_{i,t} = a + b\, r_{m,t} + \varepsilon_t$ by ordinary least squares (OLS): $\hat{a}$, $\hat{b}$, $\hat{\sigma}_\varepsilon$, $\bar{r}_m$, $S_{mm} = \sum (r_{m,t} - \bar{r}_m)^2$.
  2. $AR_\tau = r_{i,\tau} - (\hat{a} + \hat{b}\, r_{m,\tau})$; table day, $r_i$, $r_m$, prediction, $AR$; $CAR[0, +5] = \sum_{\tau=0}^{5} AR_\tau$.
  3. Test $AR_0 = 0$ and $CAR = 0$ with the prediction-error variances $\hat{\sigma}_\varepsilon^2[1 + 1/T + (r_{m,0} - \bar{r}_m)^2 / S_{mm}]$ and $\hat{\sigma}_\varepsilon^2[L + L^2/T + (\sum_\tau r_{m,\tau} - L\bar{r}_m)^2 / S_{mm}]$; Student $t$ with $T - 2$ degrees of freedom under the assumptions above.

- **Report**: the estimates, the AR table, CAR, $t$ and $p$; BET contains both big banks: what does this do to $\hat{b}$ and to the AR?
- **Worked solution**: notebook section *B7*

Back Exercise list

# Pack 6: B8 [Proposed]: The Same Event for BRD

- **Task**
  - the event of B7 for BRD (adjusted close) against the BET index, same windows
  - reuse the B7 function, only the symbol changes

  1. Report $\hat{b}$, $\hat{\sigma}_\varepsilon$, the AR table and $CAR[0,+5]$ with $t$ and $p$.
  2. Build a two-bank table: $AR_{\mathbf{0}}$, its prediction-error $t$, $CAR$ and its $t$.
  3. Plot the two CAR paths with $\pm 1.96$ SE bands on one chart.
  4. Which bank was hit harder relative to its normal volatility? Use the prediction-error $t$ of $AR_{\mathbf{0}}$.

- **Report**: the two tables and the chart; the reversal can mean overreaction or a partly withdrawn policy: what dated evidence would separate the two?

Back Exercise list

# Pack 6: B9 [Proposed]: BET-TR in RON and in EUR

- **Data**: the BET-TR index (close) and BNR EUR/RON $S_t$ (RON per EUR), prices joined on common days, 5 Jan 2015 – 18 Sep 2026 (model: the conversion of B2 Six Assets)
  1. Convert: $P_t^{EUR} = P_t^{RON}/S_t$; work one date step by step (13 Aug 2025).
  2. Show $r_t^{EUR} = r_t^{RON} - \Delta s_t$, with $s_t = \ln S_t$.
  3. Compute the CAGR and the volatility (observed $A$) in RON and in EUR, and the correlation of $r^{RON}$ with $\Delta s$.
  4. Plot both wealth paths.
- **Report**: the two CAGRs, two volatilities, the correlation, the chart; for a euro investor, is the leu's depreciation a return drag, a risk, or both?

# Pack 6: B9 Extended [Proposed]: Is the Currency Effect Significant?

- **Data**: the data of B9, $n = 2\,925$ common-day returns, $A = n/Y$ from the joined prices of B9 (models: B2 Extended, B7)
  1. Show $\mathrm{Var}(r^{EUR}) = \mathrm{Var}(r^{RON}) + \mathrm{Var}(\Delta s) - 2\,\mathrm{Cov}(r^{RON}, \Delta s)$; report each term annualised ($\times A$) in decimal units, then take square roots.
  2. Regress $r_t^{RON}$ on a constant and $\Delta s_t$ with Newey–West HAC (heteroskedasticity and autocorrelation consistent) standard errors (Newey & West (1987), $L = 8$ lags): slope, $t$, $p$; add a stationary-bootstrap 95% interval for the correlation.
  3. Estimate the annual log drift of the leu, $A\times$ mean of $\Delta s_t$, with its HAC SE, $t$ and 95% interval.
  4. Show that $\ln(1 + \mathrm{CAGR}^{RON}) - \ln(1 + \mathrm{CAGR}^{EUR}) = \ln(S_{end}/S_{start})/Y$, the drift of task 3.
- **Report**: the variance terms, the regression, the drift with its interval; the hypothesis of task 3 is "zero log drift": is the gap between your two CAGRs of B9 distinguishable from zero?

# Pack 6: B1 Extended [Proposed]: The Standard Hampel Filter

**Task**

- standard Hampel filter (Pearson, Neuvo, Astola & Gabbouj (2016)): window $x_{k-K}, \ldots, x_{k+K}$ with $K = 5$ (11 quotes)
- $m_k$ = window median, $S_k = 1.4826 \times \text{median}_j\, |x_{k-j} - m_k|$
- $x_k$ is flagged if $|x_k - m_k| > t\, S_k$
- the first and last 5 quotes are not tested
- data: BNR EUR/RON 2015–2026 (clean reference) and the weekday market series of B1 Full Audit

1. Work the window around 13 Aug 2025 step by step: $m_k$, median absolute deviation (MAD), $S_k$ and $|x_k - m_k|/S_k$; for which $t$ is the quote flagged?
2. On BNR: share of implosion windows ($S_k = 0$) and share of days flagged on the grid $t = 0, 0.5, \ldots, 21$, with a plot; compare with the largest $|x_k - m_k|/S_k$ the bound of Pearson, Neuvo, Astola & Gabbouj (2016) above which no quote can be flagged.
3. On the market series, apply $t = 2$ and $t = 5$: flags, overlap with the 11 flags and 9 errors of B1 Full Audit, volatility after removing the flags.

**Report**

- the worked window, the two shares with the plot, the flags and the volatility
- why MAD, and why 1.4826?
- what happens in an implosion window?
- can a threshold replace the BNR check of B1 Full Audit?

# Pack 7: B10 [Proposed]: Stablecoin Growth

- **Data**: DefiLlama, daily total circulating value of USD-pegged stablecoins, in bn USD (divided by $10^9$), 31 Dec 2019 – 18 Sep 2026 (models: A2–A3, B2)
  1. Extract year-end values 2019–2025 and the calendar-year growth; report 2026 separately as year-to-date growth to 18 Sep.
  2. Find the 2022 peak and the following trough up to the end of 2023, with dates, and the peak-to-trough change.
  3. Compute the CAGR from 1 Jan 2020 to 18 Sep 2026 with the exact endpoints.
  4. Plot the series with the year-end points.
- **Report**: the year-end table, the peak, trough and change, the CAGR, the chart; supply is a USD value in circulation, not a price: can it fall without any token being burned?

Back Exercise list

# Pack 7: B10 Extended [Proposed]: Did the Terra Collapse Change the Growth Regime?

**Task**

- daily log growth $g_t = \Delta \ln(\text{supply}_t)$, 2 Jan 2020 – 18 Sep 2026, $n = 2\,452$
- Terra collapse on 9 May 2022 (model: the OLS regression and $t$ test of B7)

1. Known date (Chow (1960)): $g_t = a + b\,D_t + e_t$, $D_t = 1\{t \geq 9 \text{ May } 2022\}$, HAC SE (Newey & West (1987), $L = 8$ lags); $a$: daily growth before the date, $b$: its change after; report $a$, $b$ (annualised $\times 365$) and $t$.
2. Unknown date (Andrews (1993)): the HAC Wald statistic $(\hat{b}/\,\mathrm{SE}_{\mathrm{HAC}}(\hat{b}))^2$ for every candidate date in the central 70% (3 Jan 2021 – 15 Sep 2025); compare its maximum with 8.85, the 5% critical value for this trimming.
3. Least-squares date: the date minimising the sum of squared residuals (SSR); plot the Wald and SSR profiles.
4. Check the residual variance before and after the date, and say whether the break-date interval of Bai (1997), which assumes the same error distribution in both regimes, can be reported.

**Report**

- $a$, $b$, $t$, the two dates and the profiles
- did the slowdown start with Terra?
- is the break in the level or in the drift of supply?

# Pack 8: C3 [Proposed]: Passive Investing and Demand Elasticity (1/2)

**Task**

- is the one-third pass-through of Haddad, Huebner & Loualiche (2025) (the case study of Chapter 0) robust to who turns passive?
- stock $k$, active investor $i$
- $w_{ik}$ = shares held by $i$ as a fraction of all shares of $k$
- active share $\alpha_k = \sum_i w_{ik}$
- own elasticity $\bar{\mathcal{E}}_{ik}$
- $\chi$: strength of the strategic response to the elasticity of the other investors in the stock (estimate of Haddad, Huebner & Loualiche (2025): $\chi = 2.97$)
- aggregate elasticity $\mathcal{E}_{agg,k} = \sum_i w_{ik}\, \bar{\mathcal{E}}_{ik}/(1 + \chi\alpha_k)$

1. Two investors, $w_{\mathbf{1}} = 0.4$ with $\bar{\mathcal{E}}_{\mathbf{1}} = 1$, $w_{\mathbf{2}} = 0.3$ with $\bar{\mathcal{E}}_{\mathbf{2}} = 3$, $\chi = 2.97$; $\alpha$ falls from 0.70 to 0.60 (0.1 of all shares turns passive): compute $\mathcal{E}_{agg}$ before, and after the switch when both reduce in proportion, when investor 2 (elastic) switches, when investor 1 switches.
2. Compare each $\Delta \log \mathcal{E}_{agg}/\Delta \log \alpha$ with $1/(1 + \chi\alpha) = 0.325$ at the initial $\alpha = 0.70$ (a hypothetical example, distinct from the time average 0.68 in Haddad, Huebner & Loualiche (2025)).
3. FRED `BOGZ1LM563064100Q` (ETF corporate equities) over `BOGZ1LM893064105Q` (all sectors), 2001 Q1 – 2026 Q2: the ETF share in 2001 Q1, 2020 Q4, 2026 Q2 and the implied changes with 81% active and pass-through 0.33. Explain why the ETF share is not the passive share of Haddad, Huebner & Loualiche (2025): ETFs include active funds and omit index mutual funds.

Back Exercise list

# Pack 8: C3 [Proposed]: Passive Investing and Demand Elasticity (2/2)

4. Simulation (starter code in the notebook, `numpy.random.default_rng(seed)`): 2 000 stocks, $\alpha_k \sim U(0.55, 0.85)$; 40 active investors per stock with $w_{ik} = \alpha_k \times$ weights drawn from the Dirichlet$(1, \ldots, 1)$ distribution (uniform on the 40-dimensional simplex); $\mathcal{E}_{ik}$ log-Normal with median 1.6 and log standard deviation 0.6; $\chi = 2.97$.
5. In each stock a fraction $f_k \sim U(0, 0.2)$ of $\alpha_k$ turns passive under three rules: all in proportion; most elastic first; least elastic first (whole holdings in order, the last one partly, until $f_k \alpha_k$ is reached).
6. Regress $\Delta \log \mathcal{E}_{agg,k}$ on a constant and $\Delta \log \alpha_k$, as in Haddad, Huebner & Loualiche (2025).
7. Repeat for seeds 1–10 with the same draws for all rules, and for $\chi = 0$.

⊡ **Report**

- the ETF shares and implied changes; a table of slopes (mean, min, max over seeds; simulation variability, not a sampling confidence interval) next to the mean of $1/(1 + \chi\alpha_k)$
- why is the slope one-third only for representative switchers?
- Haddad, Huebner & Loualiche (2025) attribute a fall of about 11% in aggregate elasticity to the rise of passive investing: how should this number be read if the most elastic investors turned passive first?

# References (1/2)


- Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61(4), 821–856.
- Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4), 551–563.
- 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.
- Campbell, J.Y., Lo, A.W., MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*, Ch. 4 of the book, Event-Study Analysis. Princeton University Press.
- Chow, G.C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28(3), 591–605.
- Epps, T.W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association*, 74, 291–298.
- Fisher, R.A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4), 507–521.
- Forbes, K.J., Rigobon, R. (2002). No contagion, only interdependence: measuring stock market comovements. *Journal of Finance*, 57(5), 2223–2261.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction* (5th ed.). Springer.
- Haddad, V., Huebner, P., Loualiche, E. (2025). How competitive is the stock market? Theory, evidence from portfolios, and implications for the rise of passive investing. *American Economic Review*, 115(3), 975–1018.

# References (2/2)


- Härdle, W.K., Harvey, C.R., Reule, R.C.G. (2020). Understanding cryptocurrencies. *Journal of Financial Econometrics*, 18(2), 181–208.
- Koijen, R.S.J., Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4), 1475–1515.
- Lo, A.W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36–52.
- MacKinlay, A.C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13–39.
- Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708.
- Opdyke, J.D. (2007). Comparing Sharpe ratios: so where are the p-values? *Journal of Asset Management*, 8(5), 308–336.
- Pearson, R.K., Neuvo, Y., Astola, J., Gabbouj, M. (2016). Generalized Hampel filters. *EURASIP Journal on Advances in Signal Processing*, 2016, article 87.
- Politis, D.N., Romano, J.P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303–1313.
- Scholes, M., Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5(3), 309–327.
- Public data: FRED, BNR reference rates, DefiLlama
- Code and data: github.com/danpele/MFM

# Acronyms used in this material

- **ACF**: AutoCorrelation Function
- **AI**: Artificial Intelligence
- **AR**: AutoRegressive (model); in event studies: Abnormal Return
- **BET**: Bucharest Exchange Trading (index)
- **BET-TR**: BET Total Return (index)
- **BNR**: Banca Națională a României — *the National Bank of Romania*
- **CAGR**: Compound Annual Growth Rate
- **CAR**: Cumulative Abnormal Return
- **ETF**: Exchange-Traded Fund
- **FRED**: Federal Reserve Economic Data
- **FX**: Foreign eXchange
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **MAD**: Median Absolute Deviation
- **MDD**: Maximum DrawDown
- **OLS**: Ordinary Least Squares
- **SE**: Standard Error
- **SR**: Sharpe Ratio
- **SSR**: Sum of Squared Residuals
- **US**: United States
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné)
- **VR**: Variance Ratio