








Modelling Financial Markets

Seminar 17: Bubbles and Crashes — 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

Acronyms used in this material

- **ADF**: Augmented Dickey–Fuller (test)
- **AI**: Artificial Intelligence
- **AIC**: Akaike Information Criterion
- **AR**: AutoRegressive (model); in event studies: Abnormal Return
- **BET**: Bucharest Exchange Trading (index)
- **BET-FI**: BET Financial Investment companies index
- **BIC**: Bayesian Information Criterion
- **BSADF**: Backward Sup Augmented Dickey–Fuller (statistic)
- **CI**: Confidence Interval
- **EDGAR**: Electronic Data Gathering, Analysis, and Retrieval (SEC filings database)
- **EM**: Expectation-Maximisation (algorithm)
- **FCLT**: Functional Central Limit Theorem
- **FRED**: Federal Reserve Economic Data
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GSADF**: Generalised Sup Augmented Dickey–Fuller (test)
- **JLS**: Johansen–Ledoit–Sornette (model)
- **LPPL**: Log-Periodic Power Law
- **LPPLS**: Log-Periodic Power Law Singularity (model)
- **LR**: Likelihood Ratio (test)
- **MC**: Monte Carlo (simulation)
- **NVIDIA**: NVIDIA Corporation (company name, US chip designer)
- **OLS**: Ordinary Least Squares
- **PSY**: Phillips–Shi–Yu (bubble test and date-stamping procedure)
- **PWY**: Phillips–Wu–Yu (bubble test and date-stamping procedure)
- **SADF**: Sup Augmented Dickey–Fuller (test)
- **SE**: Standard Error
- **SEC**: Securities and Exchange Commission
- **US**: United States

Can We Distinguish a Bubble from a Price Surge?

- Questions of the seminar
  - **did a price rise faster than a random walk allows?**
  - **since when?**
  - **how sure can we be?**
- Deliverables
  - your own answers to A1 and A2, before opening their solutions
  - one figure and one table per B exercise
  - the corrected artificial intelligence (AI) code in C1
  - one interpretation sentence per exercise
- Textbook: Franke, Härdle & Hafner (2019)
- Notebook for today: open the seminar notebook in Colab
  - each task names its notebook section

Which Series and Units Do We Use?

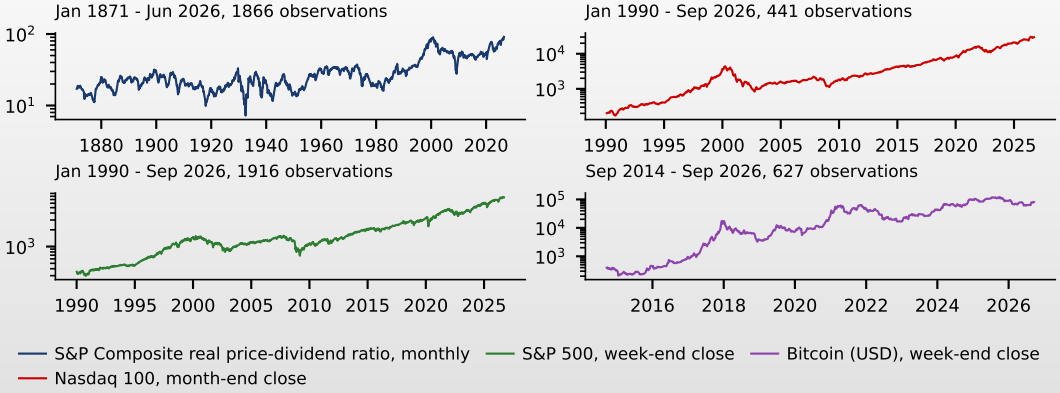
- US means United States and USD is the US dollar. S&P means Standard & Poor's; the S&P Composite and S&P 500 are distinct stock-index series with different coverage.
- BET means Bucharest Exchange Trading index. BET-FI (BET Financial Investment companies index) is the index of financial investment companies in the same family. They describe different sets of Romanian issuers; an episode in one is not automatically an episode in the other.
- FRED means Federal Reserve Economic Data. Its series codes identify the US house-price index and new-home sales in the following table; the links lead to their definitions.
- Bubble tests below use log price levels or the log price-dividend ratio. The regime model instead uses weekly log returns in percent. A persistent level and a volatile return answer different questions.

Data Used

Series	Source, price used	Frequency, sample	Used in
S&P Composite, real price and dividend	Robert J. Shiller, shillerdata.com	monthly, Jan 1871–Jun 2026	B1, C1
S&P 500	course data: daily close	weekly and monthly, 1990–2026	B4, B10
Nasdaq 100	course data: daily close	daily, weekly, monthly, 1990–2026	B7, B10, C2
Bitcoin	course data: daily close in USD	weekly, daily, 2014–2026	B2, B3, B6
BET, BET-FI	course data: daily close	monthly, weekly, 1997–2026	B5
NVIDIA, Cisco	course data: daily close adjusted for dividends and splits	daily, weekly, monthly	C2
US house prices and new-home sales	FRED `CSUSHPINSA`, `HSN1FNSA`	monthly, 1990–2011	C3

- Daily prices of the course data, one series per asset (field used: close / adjusted close), up to 18 September 2026
- Weekly and monthly series: last close of the week or month; tests use log levels $y_t = \ln P_t$ (or $\ln(P_t/D_t)$)

Setup: Run and Verify (1/2)



MFM_ch17_seminar

Setup: Run and Verify (2/2)

```
# notebook section "Data" defines
# price(name, freq, start) and shiller()
pd_ratio = shiller()['PD']        # monthly
ndx = price('ndx', 'M')           # month-end
spx = price('sp500', 'W', '1990-01-01')
btc = price('btc', 'W')           # week-end
for x in (pd_ratio, ndx, spx, btc):
    print(len(x), x.index[0].date(),
          x.index[-1].date())
```

- Expected: 1 866 (P/D, Jan 1871–Jun 2026), 441 (Nasdaq 100), 1 916 (S&P 500), 627 (Bitcoin)
- If a count differs, stop and check the name, the frequency and the start date

MFM_ch17_seminar

Fundamental Value and Bubble

- Details: Chapter 17
- **Pricing equation**: with a constant required return r, today's price is the discounted expected payoff of tomorrow

$$P_t = \frac{\mathbb{E}_t[P_{t+\mathbf{1}} + D_{t+\mathbf{1}}]}{1+r}$$

  - P_t: price at the end of period t; $D_{t+\mathbf{1}}$: dividend paid in period $t+1$; $\mathbb{E}_t$: expectation given the information at t
- **Fundamental value**: present value of the expected dividends; Gordon value if dividends grow at $g < r$

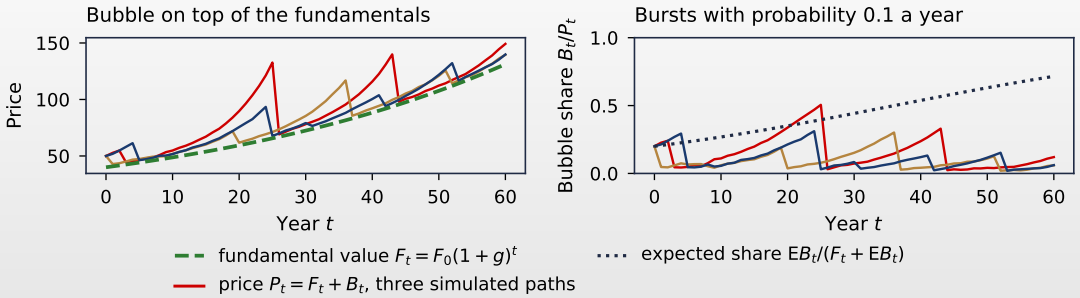
$$F_t = \sum_{i \geq \mathbf{1}} \frac{\mathbb{E}_t D_{t+i}}{(1+r)^i}, \qquad F = \frac{D_{\mathbf{1}}}{r-g}$$

- **Bubble** Blanchard & Watson (1982): the part of the price not explained by dividends

$$B_t = P_t - F_t, \qquad \mathbb{E}_t B_{t+\mathbf{1}} = (1+r)\, B_t$$

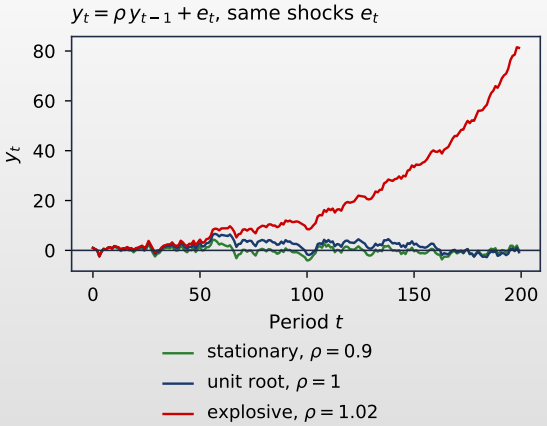
  - The pricing equation allows $B_t \neq 0$ only if it grows at rate r in expectation (derived in A1)

A Bubble That Bursts



- Bubble model Blanchard & Watson (1982): it survives with probability π and then grows faster, to compensate for the crash risk
  - $B_{t+1} = \frac{1+r}{\pi} B_t + \varepsilon_{t+1}$ with probability π, else its value equals the shock; on average $\mathbb{E}_t B_{t+1} = (1+r)B_t$; here $\pi = 0.9$, $B_0 = 10$, $F_0 = 40$
  - Repeated positive restarts in the illustration use the Evans variant of A4. With zero additive shocks, the simpler model stays at zero after its first crash; rolling windows can separate growth from collapse

Stationary, Unit and Explosive Roots



- **Autoregressive model, AR(1)**, for the log price or the log price-dividend ratio y_t

$$y_t = \rho\, y_{t-1} + e_t$$

  - e_t: shock with mean 0; ρ: the root
- Three cases (same shocks in the chart)
  - $|\rho| < 1$: stationary, shocks fade
  - $\rho = 1$: unit root (random walk), shocks persist
  - $\rho > 1$: explosive, each rise feeds the next: the signature of a bubble
- A bubble lives for a while: test $\rho > 1$ on windows, not on the whole sample

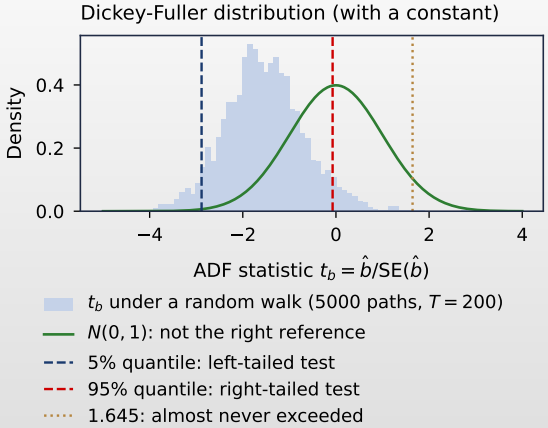
The ADF Regression and the Right Tail

- **Augmented Dickey–Fuller (ADF) regression** on a window, constant, no lags

$$\Delta y_t = a + b\, y_{t-\mathbf{1}} + e_t, \qquad t_b = \frac{\hat{b}}{\mathrm{SE}(\hat{b})}$$

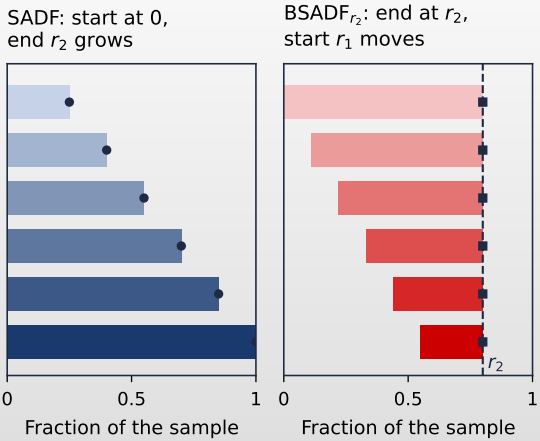
  - $\Delta y_t = y_t - y_{t-\mathbf{1}}$; $b = \rho - 1$; a: constant; $\mathrm{SE}(\hat{b})$: Standard Error (SE) of the Ordinary Least Squares (OLS) estimate $\hat{b}$
- **Two directions** of the same test
  - Usual (left) test: $H_{\mathbf{0}}: b = 0$ against $H_{\mathbf{1}}: b < 0$ (stationarity)
  - Bubble (right) test: $H_{\mathbf{0}}: b = 0$ against $H_{\mathbf{1}}: b > 0$ (explosiveness); reject for **large** t_b
- **Reference distribution**: under $H_{\mathbf{0}}$, t_b is not $N(0,1)$
  - Dickey–Fuller distribution, shifted left: critical values come from simulated random walks (next slide)

Where the Right-Tailed Test Rejects



- 5 000 simulated random walks, $T = 200$; ADF with a constant on each
- Almost all of the distribution lies left of 0
  - The estimated constant absorbs part of the walk and pushes $\hat{b}$ down
- With the Normal critical value 1.645, the right-tailed test is strongly conservative under the null and can lose power in finite samples

SADF, GSADF and BSADF



- $\mathrm{ADF}_{r_1}^{r_2}$: the t_b on the window from fraction r_1 to fraction r_2 of the sample Phillips, Shi & Yu (2015a)
- Supremum statistics over windows of at least r_0: Supremum ADF (SADF), Backward SADF (BSADF), Generalised SADF (GSADF)

$$\mathrm{SADF} = \sup_{r_2 \geq r_0} \mathrm{ADF}_0^{r_2}, \qquad \mathrm{GSADF} = \sup_{r_2 \geq r_0} \mathrm{BSADF}_{r_2}$$

$$\mathrm{BSADF}_{r_2} = \sup_{0 \leq r_1 \leq r_2 - r_0} \mathrm{ADF}_{r_1}^{r_2}$$

- SADF: start fixed (left); BSADF: end fixed at r_2, start moves (right); GSADF: both move
- BSADF uses only data up to r_2: it can date an episode in real time

Window Rules and Episodes

- **Minimum window** Phillips, Shi & Yu (2015a), as a fraction of the T changes and in observations

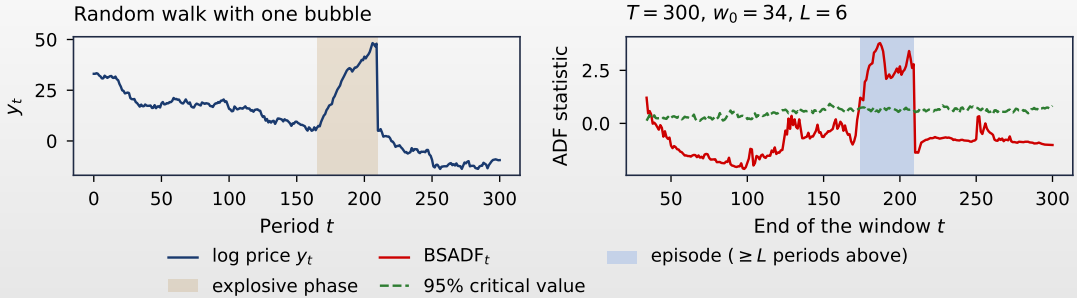
$$r_0 = 0.01 + \frac{1.8}{\sqrt{T}}, \qquad w_0 = \lfloor r_0 T \rfloor$$

  - $\lfloor \cdot \rfloor$: rounded down; $T = 600$ months: $r_0 = 0.01 + 1.8/\sqrt{600} = 0.083$, $w_0 = 50$ months
- **Episode**: BSADF_t above its pointwise 95% critical value for at least L consecutive periods

$$L = \lceil \ln T \rceil$$

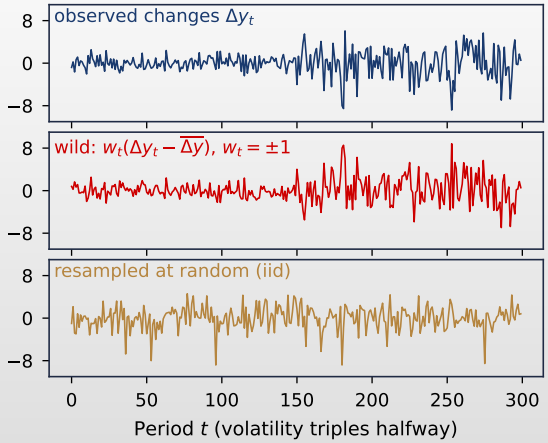
  - $\lceil \cdot \rceil$: rounded up; $T = 600$: $\ln 600 = 6.40$, so $L = 7$ months; shorter runs are not episodes
  - Start dated in hindsight, **confirmed** $L - 1$ periods later; **rise** or **fall**: sign of the price change over the episode

Date-Stamping a Simulated Bubble



- Left: a random walk with one explosive phase ($\rho = 1.04$), then a collapse; right: BSADF_t and its 95% critical value from 300 simulated walks
- Here $T = 300$, so $L = \lceil \ln 300 \rceil = 6$ (not 7 as for $T = 600$); the episode is dated from the first period of a run of at least L exceedances; after the collapse BSADF falls back at once

Critical Values: Monte Carlo and Wild Bootstrap



- **Monte Carlo (MC)**: simulate random walks of the same T
  - $y_t = T^{-1} + y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0,1)$; recompute BSADF and GSADF; 95% quantiles
  - Assumes constant volatility
- **Wild bootstrap** Harvey et al. (2016): keep each observed change, flip its sign at random

$$y_t^* = y_{t-1}^* + w_t\,(\Delta y_t - \overline{\Delta y}), \quad w_t = \pm 1$$

  - w_t: independent, each sign with probability $\frac{1}{2}$; $\overline{\Delta y}$: mean change
  - Large changes stay where they occurred (chart): the critical values follow the volatility

Regimes: the Markov-Switching Model

- **Two regimes** Hamilton (1989): weekly log returns are expressed in percent; the mean and volatility depend on a hidden state s_t

$$r_t = \mu_{s_t} + \sigma_{s_t}\varepsilon_t, \qquad \varepsilon_t \sim N(0,1), \qquad s_t \in \{1,2\}$$

  - $s_t = 1$: calm, $s_t = 2$: turbulent; μ_i, σ_i: mean and volatility of regime i

- **Markov chain**: transition probabilities, expected durations and stationary probabilities

$$p_{ij} = P(s_t = j \mid s_{t-1} = i), \qquad \mathbb{E}[\text{duration}_i] = \frac{1}{1-p_{ii}}, \qquad \pi_2 = \frac{p_{12}}{p_{12}+p_{21}}$$

  - $p_{11} = 0.98$: 50 periods on average; π_2: long-run share of time in regime 2, $\pi_1 = 1 - \pi_2$

- **Unconditional moments** (law of total variance)

$$\mathbb{E}r = \sum_i \pi_i \mu_i, \qquad \mathrm{Var}(r) = \sum_i \pi_i \sigma_i^2 + \sum_i \pi_i (\mu_i - \mathbb{E}r)^2$$

  - Mean variance within regimes plus the variance of the regime means

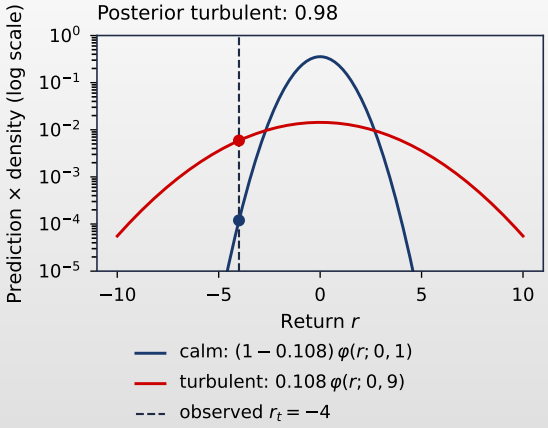
The Hamilton Filter

- **Hamilton filter** = Bayes' rule every period; $P_{t-\mathbf{1}}(i)$: filtered probability of regime i given the data up to $t-1$
- **Predict**, then **update** with the Normal density $\varphi(r_t; \mu_j, \sigma_j^{\mathbf{2}})$

$$\hat{P}_t(j) = \sum_i P_{t-\mathbf{1}}(i)\, p_{ij}, \qquad P_t(j) = \frac{\hat{P}_t(j)\, \varphi(r_t; \mu_j, \sigma_j^{\mathbf{2}})}{\sum_k \hat{P}_t(k)\, \varphi(r_t; \mu_k, \sigma_k^{\mathbf{2}})}$$

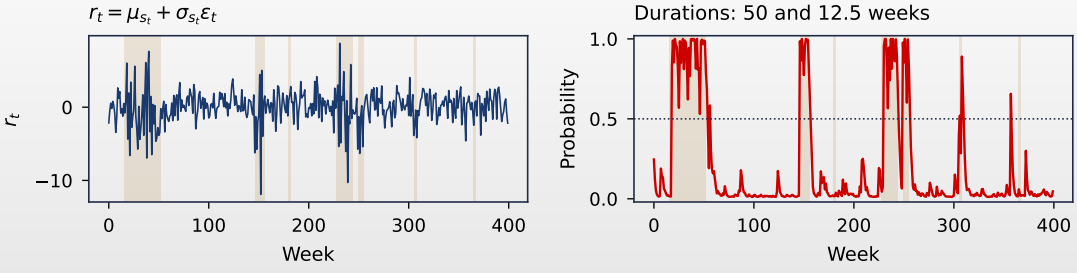
  - The denominator is the density of r_t given the past; the sum of its logs is the log-likelihood $\hat{\ell}$, maximised over the parameters
- **Filtered** and **smoothed** probability
  - Filtered: data up to t (usable in real time); smoothed: the whole sample (hindsight)
- **Tests** (B4)
  - $t = \hat{\mu}_{\mathbf{2}}/\mathrm{SE}(\hat{\mu}_{\mathbf{2}})$, compared with 1.96; Likelihood Ratio (LR): $\mathrm{LR} = 2(\hat{\ell}_{\mathbf{2}} - \hat{\ell}_{\mathbf{1}})$, two regimes against one ($\hat{\ell}_{\mathbf{1}}$, $\hat{\ell}_{\mathbf{2}}$: maximised log-likelihoods)
  - Under one regime the p_{ij} are not identified: LR has no $\chi^{\mathbf{2}}$ distribution, use a parametric bootstrap

One Filter Step in Numbers



- Inputs: $P_{t-\mathbf{1}}(2)=0.10$, $p_{\mathbf{11}}=0.98$, $p_{\mathbf{22}}=0.90$; $r_t=-4$, with $\sigma_{\mathbf{1}}=1$, $\sigma_{\mathbf{2}}=3$ and zero means
- Chart (log scale): the predicted probability of each regime times its density
- Question for the class: starting from 0.10, how far does a single 4% loss move the probability of the turbulent regime?
  - Answer: to 0.98; the calm density at -4 is about 400 times smaller

Regimes in a Simulation



- Simulated weekly returns: calm $\mu_1 = 0.3$, $\sigma_1 = 1.5$; turbulent $\mu_2 = -0.5$, $\sigma_2 = 4$; $p_{11} = 0.98$, $p_{22} = 0.92$
- The filtered probability (right, true parameters) rises within a few weeks of each turbulent spell; isolated large weeks give short false alarms

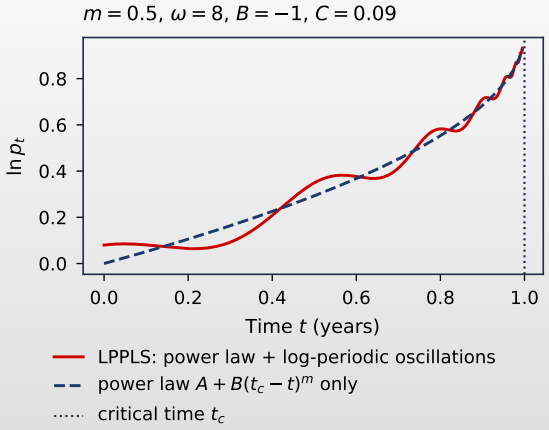
LPPLS: a Crash as a Critical Time

- Log-Periodic Power Law Singularity (LPPLS), the idea of Johansen, Ledoit & Sornette (2000): herding makes the price accelerate faster than exponentially, with ever faster oscillations

$$\ln p_t = A + B(t_c - t)^m + (t_c - t)^m \left[C_1 \cos(\omega \ln(t_c - t)) + C_2 \sin(\omega \ln(t_c - t))\right]$$

  - t_c: **critical time**, the most likely end of the bubble regime, not a crash date; A: log price at t_c
  - $B < 0$, $0 < m < 1$: the slope $-mB(t_c - t)^{m-1}$ grows without bound as $t \to t_c$
  - ω: angular frequency of the log-periodic oscillations; C_1, C_2: their amplitudes, $C = \sqrt{C_1^2 + C_2^2}$
- **Fit** Filimonov & Sornette (2013)
  - For fixed (t_c, m, ω) the model is linear in (A, B, C_1, C_2): OLS gives them; a numerical search covers the three nonlinear parameters

LPPLS Trajectory and Qualified Fits



- Chart: the power law alone (dashed) accelerates towards t_c; the oscillations around it speed up in calendar time. This illustrative curve fails the damping filter ($D \approx 0.69 < 1$, condition below), so it is not a qualified bubble fit
- **Qualified** fit: parameters inside the bounds of the filter of Shu & Zhu (2020a)
  - E.g. $B < 0$, $m \in [0.01, 0.99]$, $\omega \in [2, 25]$, oscillations $O = \frac{\omega}{\pi} \ln \frac{t_c - t_1}{t_c - t_2} \geq 2.5$, damping $D = \frac{m|B|}{\omega|C|} \geq 1$
  - $[t_1, t_2]$: the estimation window
- **Edge fit**: m or ω at the boundary of the search space
  - Not an interior optimum: the data do not pin the parameter down
  - $\hat{t}_c$ moves when the start of the window changes (B3)

Tools for the Extension Exercises

- **Mildly explosive root** Phillips & Magdalinos (2007) (A5): $\rho_T = 1 + c/k_T$, $c > 0$, $k_T \to \infty$, $k_T/T \to 0$
  - The root tends to 1 more slowly than $1/T$
- **Expectation–Maximisation (EM) algorithm** Hamilton (1990) (A6): the E-step computes the smoothed probabilities ξ_{tj}; given them, the M-step re-estimates the parameters of each regime
- **Number of regimes**: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) (B6)
  - $\mathrm{AIC} = -2\hat{\ell} + 2q$, $\mathrm{BIC} = -2\hat{\ell} + q \ln T$ (q: parameters; smaller is better)
  - Bootstrap $p = (1 + \#\{\mathrm{LR}^* \geq \mathrm{LR}\})/(R+1)$; LR^*: bootstrap values, $\#\{\cdot\}$: their count, R: number of samples
- **LPPLS confidence indicator** Demirer et al. (2019); Shu & Zhu (2020a) (B7): share of qualified fits among many windows ending on the same date
  - Linear probability model $\mathrm{fall}_t = a + b\,\mathrm{signal}_t + u_t$ with Newey–West SE; $\mathrm{signal}_t = 1$ if the indicator is > 0, $\mathrm{fall}_t = 1$ if a large fall follows (defined in B7)
- More in Chapter 17: historical bubbles, the evidence on predictability and the case study

Exercise Map

Exercise	Question	Type	Model
A1	can a price exceed the value of the dividends, and how does the gap grow?	Solved	–
A2	how much does one bad week move the regime probability?	Solved	–
B1	is Shiller P/D explosive, and when?	Solved	–
B2	does changing volatility change Bitcoin's episodes?	Solved	B1
B3	how far does the LPPLS critical time move with the window?	Solved	–
B4	when was the S&P 500 in its turbulent regime?	Solved	A2
C1	what is wrong in an AI answer on SADF and LPPLS?	Proposed	B1, B3
B5	is the Bucharest Stock Exchange explosive today?	Proposed, homework	B1, B2
Extension packs	A3–A8, B6–B10, C2, C3 (optional, after the lecture)	Solved and Proposed (marked on each task)	named 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: estimation and inference on real data, each ending with an interpretation question; Part C: open questions and AI critique

A1 [Solved]: Present Value and the Rational Bubble — Task

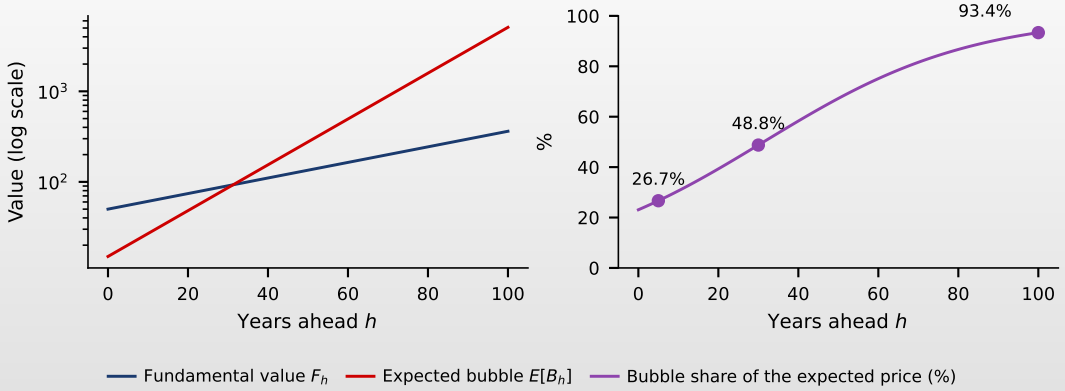
- **Questions**
  - can a price exceed the present value of its dividends in a rational market?
  - what happens to the gap over time?
- **Inputs**: a share with price $P_t \geq 0$ and dividends $D_t \geq 0$ whose discounted expected sum is finite Blanchard & Watson (1982)
  - Model assumption, a constant required return $r > 0$: the pricing equation on the slide "Fundamental Value and Bubble"
  - Numbers for task 4: $D_{\mathbf{1}} = 2$, $g = 2\%$, $r = 6\%$, $P = 65$
- **Tasks**
  1. Substitute the pricing equation into itself h times and write P_t as discounted expected dividends up to $t + h$ plus a remainder.
  2. Let $h \to \infty$, define F_t (the fundamental value) and $B_t = P_t - F_t$, and prove that $\mathbb{E}_t[B_{t+\mathbf{1}}] = (1 + r)B_t$.
  3. Prove that $B_t \geq 0$ and that $B_t = 0$ implies $B_{t+k} = 0$ for every $k \geq 1$ Diba & Grossman (1988).
  4. In the Gordon case compute $F = D_{\mathbf{1}}/(r - g)$, $B = P - F$ and the bubble share $\mathbb{E}[B_h]/(F_h + \mathbb{E}[B_h])$ for $h = 5, 30, 100$ years.
- **Report**: the derivation, F, B and the three shares
- **Interpretation**
  - why must a rational bubble exist from the first trade?
  - why is a growing bubble share the testable implication?
- Notebook: section A1

Back Exercise list

A1 [Solved]: Solution (1/3) — Steps 1–3, the Derivation

- Step 1: write $P_{t+\mathbf{1}} = \mathbb{E}_{t+\mathbf{1}}[P_{t+\mathbf{2}} + D_{t+\mathbf{2}}]/(1+r)$ and substitute; iterated expectations give $\mathbb{E}_t\mathbb{E}_{t+\mathbf{1}} = \mathbb{E}_t$; after h steps $P_t = \sum_{i=\mathbf{1}}^{h} \dfrac{\mathbb{E}_t D_{t+i}}{(1+r)^i} + \dfrac{\mathbb{E}_t P_{t+h}}{(1+r)^h}$
- Step 2: with $D \geq 0$ the partial sums increase to the finite $F_t = \sum_{i\geq \mathbf{1}} \mathbb{E}_t D_{t+i}/(1+r)^i$; the remainder converges to $B_t := P_t - F_t = \lim_h \mathbb{E}_t P_{t+h}/(1+r)^h$
  - $(1+r)F_t = \mathbb{E}_t D_{t+\mathbf{1}} + \mathbb{E}_t F_{t+\mathbf{1}}$ (iterated expectations, monotone convergence); subtract it from $(1+r)P_t = \mathbb{E}_t D_{t+\mathbf{1}} + \mathbb{E}_t P_{t+\mathbf{1}}$: $\mathbb{E}_t B_{t+\mathbf{1}} = (1+r)B_t$
  - Transversality, $\lim_h \mathbb{E}_t P_{t+h}/(1+r)^h = 0$, is exactly $B_t = 0$
- Step 3: B_t is the limit of the remainders $\mathbb{E}_t P_{t+h}/(1+r)^h \geq 0$, so $B_t \geq 0$; if $B_t = 0$, then $\mathbb{E}_t B_{t+\mathbf{1}} = 0$ with $B_{t+\mathbf{1}} \geq 0$, so $B_{t+\mathbf{1}} = 0$ almost surely, and by induction $B_{t+k} = 0$

A1 [Solved]: Solution (2/3) — Step 4, Numbers and Chart



MFM_ch17_seminar

A1 [Solved]: Solution (3/3) — Step 4, Numbers and Chart

- Step 4: $F = 2/(0.06 - 0.02) = 50.0$; $B = 65 - 50.0 = 15.0$, i.e. 23.1% of the price
- $F_h = 50.0 \cdot 1.02^h$, $\mathbb{E}B_h = 15.0 \cdot 1.06^h$
- Bubble share: 26.7% ($h = 5$), 48.8% ($h = 30$), 93.4% ($h = 100$)
- Interpretation: a bubble cannot start from zero later; since $r > g$ its **expected** value eventually dominates the price, so the expected price path is explosive relative to fundamentals: this is what right-tailed tests look for

MFM_ch17_seminar

A2 [Solved]: Regime Durations and One Filter Step — Task

- **Question**: how much does one bad week move the probability of the turbulent regime?
- **Inputs**: weekly log returns in %, the two-regime Markov-switching model Hamilton (1989)
  - Calm: $\mu_1 = 0.3$, $\sigma_1 = 2$; turbulent: $\mu_2 = -0.5$, $\sigma_2 = 5$; $p_{11} = 0.98$, $p_{22} = 0.90$
  - Last week $P(s = 2 \mid \text{data}) = 0.2$; this week $r_t = -6\%$
- **Tasks**
  1. Compute the expected duration of each regime.
  2. Compute the stationary probabilities π_1 and π_2.
  3. Compute the unconditional mean and standard deviation of r_t with the law of total variance.
  4. Do one Hamilton filter step: the predicted probability, the two densities at $r_t = -6$, the updated probability.
- **Report**: the durations, π_1, π_2, mean and standard deviation, and the three probabilities of step 4
- **Interpretation**: why does one week change the regime probability so much?
- Notebook: section A2

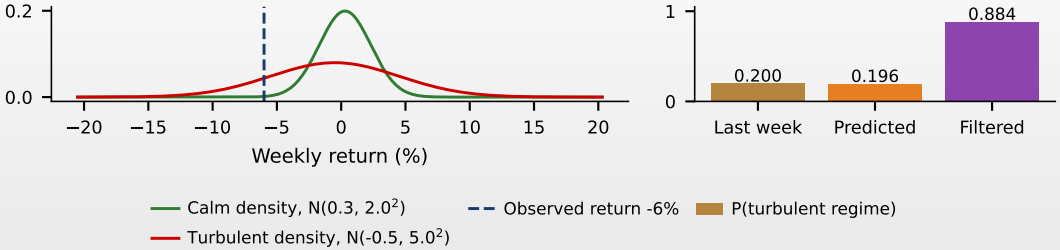
Back Exercise list

A2 [Solved]: Solution — Regime Moments and the Filter Update

- Step 1: $1/(1-0.98) = 50$ weeks (calm), $1/(1-0.90) = 10$ weeks (turbulent)
- Step 2: $\pi_{\mathbf{2}} = 0.02/(0.02+0.10) = 0.167$, $\pi_{\mathbf{1}} = 0.833$
- Step 3: $\mathbb{E}r = 0.833 \cdot 0.3 + 0.167 \cdot (-0.5) = 0.167$; $\mathrm{Var}(r) = \sum_i \pi_i \sigma_i^{\mathbf{2}} + \sum_i \pi_i (\mu_i - \mathbb{E}r)^{\mathbf{2}} = 7.59$, s.d. 2.75
- Step 4: prediction $0.2 \cdot 0.90 + 0.8 \cdot 0.02 = 0.196$
  - Densities $\varphi(-6; 0.3; 2^{\mathbf{2}}) = 0.00140$, $\varphi(-6; -0.5; 5^{\mathbf{2}}) = 0.04357$
  - Update $\dfrac{0.196 \cdot 0.04357}{0.196 \cdot 0.04357 + (1-0.196) \cdot 0.00140} = 0.884$
- Interpretation: -6% lies more than three calm standard deviations below the calm mean, so its calm density is tiny: one such week almost settles the regime

MFM_ch17_seminar

A2: Updating the Regime Probability



MFM_ch17_seminar

B1 [Solved]: Is the S&P Composite Price-Dividend Ratio Explosive? — Task

- **Questions**
  - is the S&P Composite price-dividend ratio explosive?
  - when?
  - does the answer depend on the minimum window?
- **Data**: S&P Composite real price and real dividend, monthly 1871–2026; $y_t = \ln(P_t/D_t)$, $T = 1\,865$ changes
- **Tasks**
  1. Compute GSADF (constant, no lags) with the minimum window $r_0 = 0.01 + 1.8/\sqrt{T}$.
  2. Simulate 1 000 random walks with drift, as on the critical-values slide, and take the 95% quantile of GSADF and, date by date, of BSADF Phillips, Shi & Yu (2015a).
  3. Repeat steps 1–2 with $r_0 = 0.03$ and $r_0 = 0.08$.
  4. For each r_0, date-stamp the episodes (at least $L = \lceil \ln T \rceil$ months above the critical value) and mark each as a rise or a fall of the ratio.
- **Report**: a table (r_0, w_0, GSADF, critical value, episodes) and a chart of BSADF against its critical values
- **Interpretation**
  - which episodes survive every window?
  - is every explosive episode of the ratio a bubble?
- Notebook: section B1

Back Exercise list

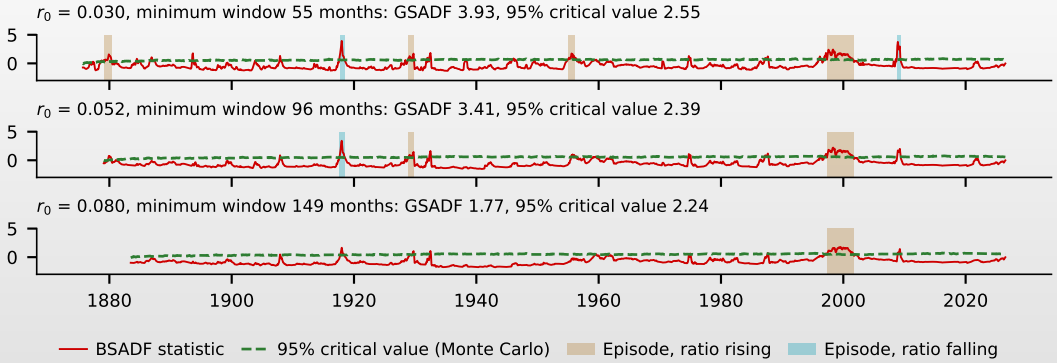
B1 [Solved]: Solution (1/2) — Steps and Table

- Steps 1–2: log ratio; for every end date, ADF on all windows of at least w_0 months; the supremum over start dates is BSADF, over end dates GSADF; the same computation on the simulated walks gives the critical values
- Steps 3–4, results (episodes that are a **fall** of the ratio are marked):

r_0	w_0	GSADF	95% crit. value	Episodes (BSADF)
0.030	55	3.93	2.55	Jan 1879–Apr 1880; Aug 1917–May 1918 (falling); Nov 1928–Sep 1929; Dec 1954–Jan 1956; May 1997–Aug 2001; Sep 2008–Apr 2009 (falling)
0.052	96	3.41	2.39	Aug 1917–Jun 1918 (falling); Nov 1928–Sep 1929; May 1997–Aug 2001
0.080	149	1.77	2.24	May 1997–Aug 2001

- The dot-com episode (1997–2001) is dated with every window; with $r_0 = 0.08$ GSADF is below its critical value, yet BSADF still crosses its pointwise critical values: the global test is stricter than the date-stamping
- The shortest window adds Jan 1879–Apr 1880; Dec 1954–Jan 1956; Sep 2008–Apr 2009 (falling): fewer observations per regression, noisier ADF statistics, more crossings
- 1917–1918 is a **fall** of the ratio (real prices dropped faster than real dividends): explosive, but not a bubble
- Interpretation: only episodes that survive every r_0 and point upwards support the bubble reading

B1 [Solved]: Solution (2/2) — Chart



Red: BSADF; green: its pointwise 95% critical value; shaded: date-stamped episodes, rising (sand) or falling (teal) ratio

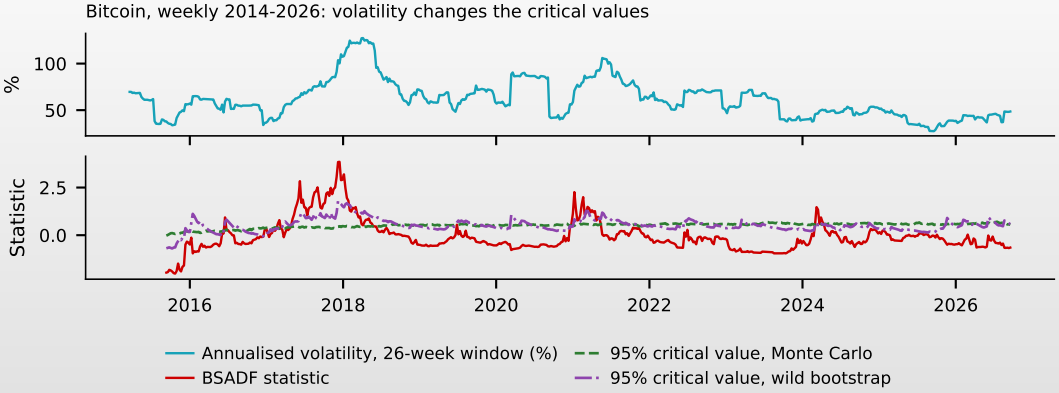
MFM_ch17_seminar

B2 [Solved]: Bitcoin: Simulated vs Bootstrap Critical Values — Task

- **Question**: does Bitcoin's changing volatility change the verdict of the test and the dates of the episodes?
- **Model** (solved exercises to follow): B1
- **Data**: weekly log closes of Bitcoin, September 2014 – September 2026, $T = 626$ weekly changes
- **Tasks**
  1. Compute GSADF with the minimum-window rule and Monte Carlo critical values from 2 000 random walks with Normal shocks.
  2. Wild bootstrap: multiply the demeaned weekly changes by independent random signs $w_t = \pm 1$, rebuild the path and recompute BSADF and GSADF, 1 000 times; take the 95% quantiles Harvey et al. (2016).
  3. Date-stamp the episodes with both critical-value sequences, minimum duration $\lceil \ln T \rceil$ weeks.
  4. Plot BSADF with both sequences, and the 26-week rolling volatility above it.
- **Report**: GSADF, both 95% critical values, both lists of episodes and the chart
- **Interpretation**
  - when are the bootstrap critical values higher?
  - which episode is robust to the choice?
- Notebook: section B2

Back Exercise list

B2 [Solved]: Solution (1/2)



MFM_ch17_seminar

B2 [Solved]: Solution (2/2)

- Steps 1–2: GSADF $= 3.85$; 95% critical value: Monte Carlo 2.27, wild bootstrap 2.82; $3.85 > 2.82$: the random walk is rejected at 5% with either
- Step 3, Monte Carlo episodes: Apr 2017–Jun 2018; Dec 2020–May 2021
  - Wild bootstrap: May 2017–Jan 2018; Dec 2020–Feb 2021
- Step 4: the bootstrap critical values rise in and after the volatile periods, because the resampled changes keep the large shocks where they occurred
  - Largest pointwise critical value: 1.80 with the wild bootstrap vs 0.71 with Monte Carlo
- Interpretation: the 2017 episode survives both, shorter with the bootstrap; the 2021 episode shrinks: its explosiveness is partly a volatility effect

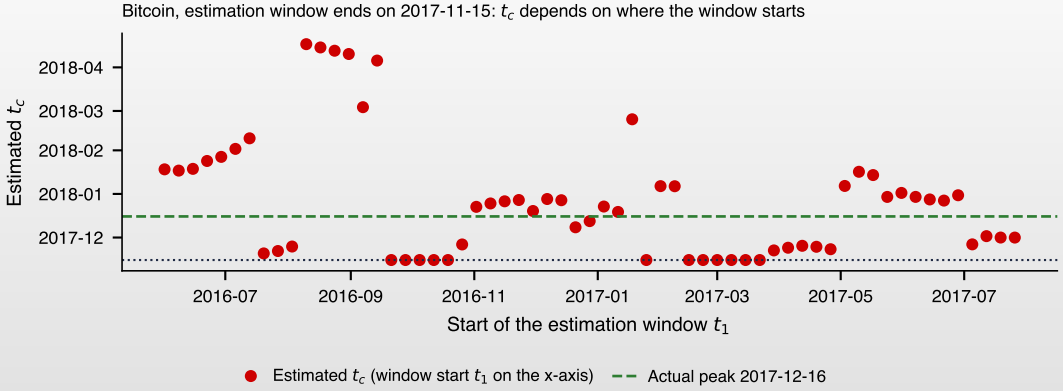
MFM_ch17_seminar

B3 [Solved]: How Much Does the Critical Time Depend on the Window? — Task

- **Question**: how far does the LPPLS critical time t_c move when only the start of the estimation window changes?
- **Data**: Bitcoin daily log closes (LPPLS on Bitcoin: Shu & Zhu (2020b)); 61 windows, all ending on $t_2 = 15$ November 2017, starting every 7 days from 1 June 2016 while the start is no later than 1 August 2017 (last start: 26 July)
  - **Edge fit**: m or ω at the boundary of the search space ($m \leq 0.002$ or ≥ 0.998; $\omega \leq 1.01$ or ≥ 49.9)
- **Tasks**
  1. Fit LPPLS on each window: (A, B, C_1, C_2) by OLS, (t_c, m, ω) by a grid and Nelder–Mead in the search space of Shu & Zhu (2020a, eq. 11); calibration of Filimonov & Sornette (2013).
  2. Report the minimum, median and maximum t_c and the interquartile range of the error, t_c minus the date of the actual peak (highest daily close).
  3. Report the shares of qualified fits (eq. 12), of edge fits and of fits within 30 days of the peak.
  4. Plot t_c against the start of the window, with the actual peak.
- **Report**: the three dates, the interquartile range, the three shares and the chart
- **Interpretation**: would you trust a single t_c reported without its window?
- Notebook: section B3

Back Exercise list

B3 [Solved]: Solution (1/2)



MFM_ch17_seminar

B3 [Solved]: Solution (2/2)

- Steps 1–2: t_c ranges from 15 Nov 2017 to 17 Apr 2018, median 23 Dec 2017; actual peak 16 Dec 2017; interquartile range of the error 53 days, median error 7 days
- Step 3: qualified fits 0%; edge fits 43%; within 30 days of the peak 56%
- Interpretation: the same end date gives critical times months apart, depending only on the start of the window; a single t_c is not a forecast

MFM_ch17_seminar

B4 [Solved]: Markov Switching on the S&P 500 — Task

- **Questions**
  - when was the S&P 500 in its turbulent regime?
  - is the turbulent mean negative?
- **Model** (solved exercises to follow): A2
- **Data**: weekly log returns of the S&P 500 in %, 1990–2026 (1 915 weeks)
- **Tasks**
  1. Estimate the two-regime Markov-switching model by maximum likelihood (Hamilton filter, 20 random starts); report estimates, standard errors and expected durations Hamilton (1989).
  2. Find the first week of 2000, of autumn 2008 and of 2020 with a filtered turbulent probability above 0.5.
  3. Compute the share of weeks in which the filtered and the smoothed regime (probability above 0.5) agree.
  4. Test $\mu_{\textbf{turb}} = 0$ with a t-test at 5% and compute LR against one regime with Normal returns (The Hamilton Filter).
- **Report**: the estimates table, the three dates, the agreement share, t, LR and a chart of the regime probabilities
- **Interpretation**
  - is the turbulent mean significantly negative?
  - why can LR not be compared with a χ^2 critical value?
- Notebook: section B4

Back Exercise list

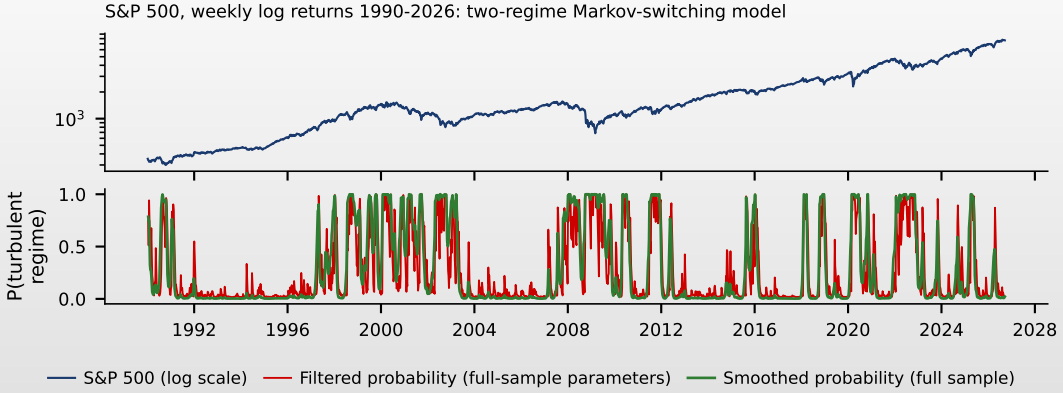
B4 [Solved]: Solution

Regime	$\hat{\mu}$	SE	$\hat{\sigma}$	$\hat{p}_{ii}$	Duration (weeks)
Calm	0.32	0.04	1.49	0.974	38.1
Turbulent	−0.27	0.17	3.67	0.931	14.4

- Step 1: the table; the calm regime covers 73% of the weeks; weekly volatility 3.67% in the turbulent regime vs 1.49% in the calm one
- Step 2: first week above 0.5: 28 Jan 2000 (2000), 3 Oct 2008 (autumn 2008), 28 Feb 2020 (2020)
- Step 3: filtered and smoothed regimes agree in 90% of weeks; the filtered probabilities use full-sample parameters, so they are pseudo real time
- Step 4: $t = -0.269/0.172 = -1.56$, $|t| < 1.96$: do not reject $\mu_{\textbf{turb}} = 0$ at 5%
- LR $= 2(-4\,069.98 - (-4\,326.02)) = 512.1$. Under one regime the transition probabilities are not identified, so χ^2 does not apply; a parametric bootstrap does (B6)
- Interpretation: the turbulent regime marks the stress periods of 2000, 2008 and 2020, but its mean is not significantly negative: regimes describe risk, not predictable losses

MFM_ch17_seminar

B4: Inferred S&P 500 Regimes



MFM_ch17_seminar

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

- **Prompt** sent to an AI assistant: "y is the monthly log price-dividend ratio of the S&P Composite. Write Python code for SADF with the Phillips–Shi–Yu (PSY) minimum window, decide at 5% whether there is a bubble, date-stamp it and tell me when it will burst."
- **Answer**: the code below and the conclusion

```
import numpy as np
from statsmodels.tsa.stattools import adfuller
T = len(y)
w0 = int(np.floor((0.01 + 1.8 / np.sqrt(T)) * T))
adf = lambda x: adfuller(x, maxlag=0, regression='c', autolag=None)[0]
sadf = max(adf(y[:t]) for t in range(w0, T + 1))
cv = adfuller(y, maxlag=0, regression='c', autolag=None)[4]['5%']
print('bubble' if sadf > cv else 'no␣bubble')
flag = [adf(y[t - w0:t + w0]) > cv for t in range(w0, T - w0)]
```

- **Conclusion of the AI**: "An LPPLS fit on the last 36 months gives $\hat{t}_c$ = March 2027: this is when the bubble bursts."

C1 [Proposed]: Critique an AI Answer — Task

- **Model** (solved exercises to follow): B1 and B3; the slide "Where the Right-Tailed Test Rejects"
- Notebook: section C1
- **Data**: the log price-dividend ratio of B1 and simulated random walks of the same length
- **Tasks**
  1. Identify the three main groups of errors (critical values, episode dating, crash prediction); list each separate defect within them, with the line and reason.
  2. For each error, say how you would detect it without knowing the answer.
  3. Write the corrected code: simulated critical values, BSADF on windows that end at t, episodes of at least $\lceil \ln T \rceil$ months.
  4. Log each error you actually found in the format of `AI_ERRORS.md`: request, answer, error, how found, correction.
- **Interpretation**: after the corrections, what can the code say about March 2027?
- AI code usually runs without an error message: the mistakes are in the statistics, not in the syntax

B5 [Proposed]: The Bucharest Stock Exchange — Task

- **Questions**
  - was the 2003–2007 boom explosive in real time?
  - is the current rise of the BET and BET-FI explosive?
- **Model** (solved exercises to follow): B1 and B2
- **Data**: official BET index, monthly log closes 1997–2026; BET-FI (financial investment companies), weekly log closes 2012–2026
- **Tasks**
  1. For each index compute GSADF with Monte Carlo (2 000 paths) and wild bootstrap (1 000 replications) 95% critical values.
  2. Date-stamp the episodes with both critical-value sequences (minimum duration $\lceil \ln T \rceil$) and classify each as a rise or a fall.
  3. Plot BSADF with both sequences for the two indices.
  4. Using the Monte Carlo critical sequence, find the confirmation date of the 2003–2007 episode (the L-th consecutive exceedance).
- **Report**: a table (T, GSADF, both critical values), the episodes with their direction and the chart
- **Interpretation**: is the current rise a bubble just because it is explosive? (On herding and log-periodic patterns before 2007: Pele, Mazurencu & Nijkamp (2013).)
- Notebook: section B5

Back Exercise list

Extension Packs: What Each One Needs

Pack	Exercises	Prerequisites
1. Rational bubbles and limit theory	A3 [Solved], A4, A5	A1; for A3 and A5 the lecture section on testing
2. Regimes	A6, B6	A2, B4
3. LPPLS	A7 [Solved], A8, B7	B2, B3; the lecture section on LPPLS
4. Size under changing volatility	B8 [Solved], B9	B2
5. Choice of windows	B10	B1
6. Project starts	C2, C3	B1–B5, B7, B10; the lecture case study

- Each pack has its own notebook section; within a pack, do the exercises in order

A3 [Solved]: The Right-Tailed Dickey–Fuller Distribution — Task

- **Questions**
  - under a unit root, where does the right-tailed ADF statistic lie?
  - why is 1.645 the wrong critical value?
- **Inputs**: $y_t = y_{t-1} + \varepsilon_t$, ε_t iid (independent, identically distributed) $(0, \sigma^2)$, $y_0 = 0$; regression $\Delta y_t = a + b\, y_{t-1} + e_t$, $t = 1, \ldots, T$
  - FCLT (functional central limit theorem): $T^{-1/2} y_{\lfloor Tr \rfloor} \Rightarrow \sigma W(r)$, $r \in [0,1]$ the fraction of the sample, $\lfloor \cdot \rfloor$ rounded down, $\Rightarrow$ convergence in distribution, W a standard Brownian motion; $\tilde{y}_{t-1}$: y_{t-1} minus its sample mean
- **Tasks**
  1. Write $\hat{b}$ with the demeaned regressor $\tilde{y}_{t-1}$.
  2. Using the FCLT and continuous mapping, find the limits of $T^{-2} \sum \tilde{y}_{t-1}^2$ and $T^{-1} \sum \tilde{y}_{t-1} \varepsilon_t$.
  3. Derive the limit of t_b and show that the limit numerator has mean $-1/2$.
  4. Simulate 20 000 random walks with $T = 1\,000$ and report the 5%, 50% and 95% quantiles of t_b and the share of negative values.
- **Report**: the limit of t_b, the three quantiles, the share of negative values and a histogram
- **Interpretation**: what goes wrong if a bubble test uses 1.645?
- Notebook: section A3

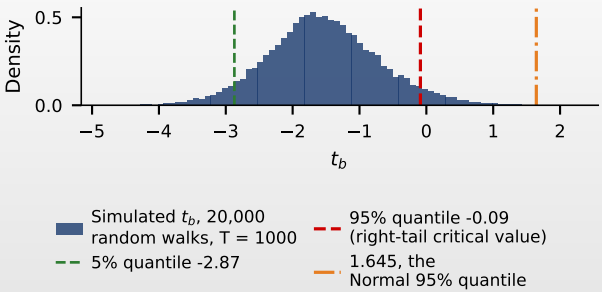
Back Exercise list

A3 [Solved]: Solution (1/2) — Steps 1–3, the Limit

- Step 1: OLS with a constant, $\hat{b} = \sum \tilde{y}_{t-1}\Delta y_t / \sum \tilde{y}_{t-1}^2 = \sum \tilde{y}_{t-1}\varepsilon_t / \sum \tilde{y}_{t-1}^2$ under H_0
- Step 2: $T^{-2}\sum \tilde{y}_{t-1}^2 \Rightarrow \sigma^2 \int_0^1 \tilde{W}^2$, $\tilde{W} = W - \int_0^1 W$; $T^{-1}\sum y_{t-1}\varepsilon_t \Rightarrow \frac{\sigma^2}{2}(W(1)^2 - 1)$, and demeaning subtracts $\sigma^2 W(1)\int_0^1 W$
- Step 3: $\mathrm{SE}(\hat{b}) = s/\sqrt{\sum \tilde{y}_{t-1}^2}$ with $s \to \sigma$, so σ cancels (PSY, eq. 5 with $r_1 = 0$, $r_2 = 1$) Phillips, Shi & Yu (2015a)
  - $t_b \Rightarrow \dfrac{\frac{1}{2}(W(1)^2 - 1) - W(1)\int_0^1 W}{\{\int_0^1 W^2 - (\int_0^1 W)^2\}^{1/2}}$
  - $\mathbb{E}[W(1)^2] = 1$ and $\mathbb{E}[W(1)\int_0^1 W] = \int_0^1 r\,dr = \frac{1}{2}$: the numerator has mean $\frac{1}{2}(1-1) - \frac{1}{2} = -\frac{1}{2}$ (simulated: -0.50)

A3 [Solved]: Solution (2/2) — Step 4, Simulation

- Step 4: 20 000 walks, $T = 1\,000$; quantiles 5%: -2.87, 50%: -1.57, 95%: -0.09
- 96% of the t_b are negative: the whole distribution lies left of zero
- Interpretation (mechanism: Where the Right-Tailed Test Rejects): with 1.645 the size of a right-tailed test is close to zero and its finite-sample power falls; under a consistent explosive alternative the statistic still exceeds any fixed threshold



MFM_ch17_seminar

A4 [Proposed]: The Blanchard–Watson Bubble — Task

- **Questions**
  - who pays for the crash risk of a bursting bubble?
  - why does a full-sample test miss it?
- **Model** (solved exercises to follow): A1
- **Inputs**: $B_{t+\mathbf{1}} = \frac{\mathbf{1}+r}{\pi} B_t + \varepsilon_{t+\mathbf{1}}$ with probability π, $B_{t+\mathbf{1}} = \varepsilon_{t+\mathbf{1}}$ with probability $1-\pi$, with innovations independent of the past and the survival draw; $\mathbb{E}\varepsilon = 0$; $B_{\mathbf{0}} = 15$, $r = 6\%$, $\pi = 0.9$ per year Blanchard & Watson (1982)
  - Evans form (positive restart): collapse to $B_{t+\mathbf{1}} = \delta$, $0 < \delta \leq B_{\mathbf{0}}$; survival $B_{t+\mathbf{1}} = \delta + \frac{\mathbf{1}+r}{\pi}(B_t - \frac{\delta}{\mathbf{1}+r})$; $y = \ln B$, $c = \ln \delta$
- **Tasks**
  1. Derive $\mathbb{E}_t B_{t+\mathbf{1}}$ and compare it with A1: the discounted bubble is a martingale.
  2. For $\varepsilon \equiv 0$, derive the conditional standard deviation $\mathrm{sd}_t(B_{t+\mathbf{1}})$ and evaluate it at $B_t = 15$.
  3. Let N count the transition of the first crash (1 means a crash next year). Give its distribution, then $\mathbb{E}N$, $P(N > 5)$ and the growth rate while the bubble survives.
  4. For the Evans form, check $\mathbb{E}_t B_{t+\mathbf{1}} = (1+r)B_t$ and derive $\mathbb{E}[\Delta y_{t+\mathbf{1}} \mid y_t]$ for $B_t \gg \delta$ and its slope in y_t.
- **Report**: the four derivations and numeric results; leave the Task 4 intercept as a function of c
- **Interpretation**
  - what does a full-sample ADF regression on $\ln B$ see?
  - why is this the Evans pitfall?
- Notebook: section A4

Back Exercise list

A5 [Proposed]: A Mildly Explosive Root — Task

- **Questions**
  - is the right-tailed test consistent against a bubble whose root is only slightly above one?
  - how fast does its statistic grow?
- **Model** (solved exercises to follow): A3
- **Inputs**: $y_t = \rho_T y_{t-1} + \varepsilon_t$, $y_0 = 0$, ε_t iid $(0, \sigma^2)$, $\rho_T = 1 + c/k_T$, $c > 0$, $k_T \to \infty$, $k_T = o(T)$, i.e. $k_T/T \to 0$ (c: how explosive, k_T: how fast the root approaches 1) Phillips & Magdalinos (2007)
- **Tasks**
  1. Write $y_t = \rho_T^t X_t$ with $X_t = \sum_{j \le t} \rho_T^{-j} \varepsilon_j$ and find the limit of $\mathrm{Var}(X_T)(\rho_T^2 - 1)$.
  2. Show that $\sum y_{t-1}^2 \approx X_T^2 \rho_T^{2T}/(\rho_T^2 - 1)$ and that $\frac{\rho_T^T}{\rho_T^{2T-1}}(\hat{\rho} - \rho_T)$ is asymptotically a ratio of two independent Normal variables. Hint: isolate innovations near opposite ends of the sample; $k_T/T \to 0$ separates the two blocks.
  3. Decompose the t-ratio of $b = \rho - 1$ and derive its divergence rate.
  4. Simulate 4 000 paths with $c = 1$, $k_T = T^{0.8}$, $T = 100, 400, 1\,600$ (regression with a constant); report ρ_T^T, the median t-ratio and a histogram.
- **Report**: the three derivations and the simulation table with the histogram
- **Interpretation**: what does the divergence rate say about the power against short, mild bubbles?
- Notebook: section A5

Back Exercise list

A6 [Proposed]: Another Filter Step and One EM Step — Task

- **Questions**
  - why does a small positive return point to the calm regime?
  - why is the EM update a weighted mean?
- **Model** (solved exercises to follow): A2
- **Inputs**: calm $\mu_1 = 0.4$, $\sigma_1 = 2$; turbulent $\mu_2 = -1$, $\sigma_2 = 6$; $p_{11} = 0.95$, $p_{22} = 0.80$; weekly log returns in %; last week $P(\text{turbulent}) = 0.1$, this week $r_t = +1\%$
- **Tasks**
  1. Compute the durations, the stationary probabilities and the unconditional mean and standard deviation.
  2. Do one filter step for $r_t = +1\%$.
  3. Compute the probability of the turbulent regime two weeks ahead, $(1-p, p)P^2$, with p the updated probability of step 2 and P the 2×2 matrix of the p_{ij}.
  4. Given smoothed probabilities ξ_{tj}, maximise $\sum_t \sum_j \xi_{tj} \ln \varphi(r_t; \mu_j, \sigma_j^2)$ over μ_j and σ_j^2 (the M-step of EM) Hamilton (1990).
- **Report**: the numbers of steps 1–3 and the two M-step formulas
- **Interpretation**: how does each week contribute to the estimate of each regime?
- Notebook: section A6

Back Exercise list

B6 [Proposed]: Bitcoin: Two or Three Regimes? — Task

- **Questions**
  - does Bitcoin need a third regime?
  - which regime do its booms belong to?
- **Model** (solved exercises to follow): B4
- **Data**: weekly log returns of Bitcoin in %, 2014–2026 (626 weeks)
- **Tasks**
  1. Estimate two and three regimes with switching mean and variance (30 random starts; check that no regime standard deviation collapses towards 0).
  2. Report the log-likelihood $\hat{\ell}$, AIC and BIC (the slide "Tools for the Extension Exercises") and each regime's mean, standard deviation and duration; plot the smoothed probabilities of the three regimes.
  3. Test two vs three regimes with $\mathrm{LR} = 2(\hat{\ell}_3 - \hat{\ell}_2)$: simulate 199 samples from the fitted two-regime model, refit both models on each, and compute the bootstrap p-value of the slide "Tools for the Extension Exercises" Hansen (1992).
  4. Compare the bootstrap 95% value with the χ^2_6 value.
- **Report**: the table, the chart, LR, the bootstrap p-value and both critical values
- **Interpretation**: why is χ^2 the wrong reference distribution?
- Notebook: section B6

Back Exercise list

A7 [Solved]: LPPLS: Concentrating Out the Linear Parameters — Task

- **Questions**
  - why is LPPLS calibrated over three nonlinear parameters only?
  - which filter conditions can be checked from the parameters alone?
- **Inputs**: $\ln p_t = A + Bf_t + C_1 g_t + C_2 h_t + u_t$ on $[t_1, t_2]$, $f_t = (t_c - t)^m$, $g_t = f_t \cos(\omega \ln(t_c - t))$, $h_t = f_t \sin(\omega \ln(t_c - t))$ Filimonov & Sornette (2013)
  - Filter of Shu & Zhu (2020a, eq. 12): the conditions for a qualified fit, plus $t_c - t_2 \leq (t_2 - t_1)/5$; relative error, Lomb and residual tests (these need data)
- **Tasks**
  1. For fixed (t_c, m, ω), derive (A, B, C_1, C_2) in closed form.
  2. Write the profile objective left for the nonlinear search.
  3. Recover C and ϕ of $C\cos(\omega \ln(t_c - t) - \phi)$ from C_1, C_2.
  4. A fit on $[2025.0, 2026.0]$ gives $t_c = 2026.08$, $m = 0.5$, $\omega = 8$, $B = -1.2$, $C_1 = 0.06$, $C_2 = -0.02$: check every condition that the parameters alone allow and plot the implied trend after subtracting the intercept A (its value is not supplied).
- **Report**: the closed form, the profile, C, ϕ and a pass/fail list
- **Interpretation**: why does concentrating out the linear parameters (A, B, C_1, C_2 as OLS functions of t_c, m, ω) make the calibration stable?
- Notebook: section A7

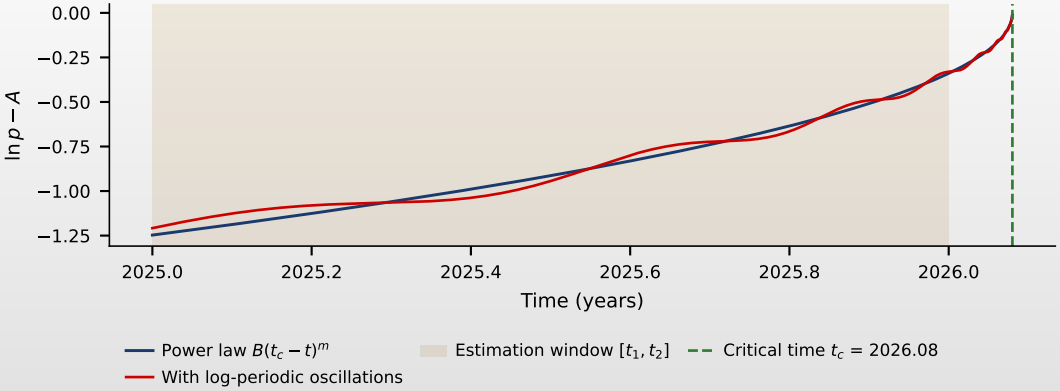
Back Exercise list

A7 [Solved]: Solution

- Step 1: with $X_t = (1, f_t, g_t, h_t)$ and $\beta = (A, B, C_1, C_2)^\top$, minimise $\sum_t (\ln p_t - X_t\beta)^2$: $\hat\beta = (X^\top X)^{-1} X^\top y$, a 4×4 system of sums such as N, $\sum f_t$, $\sum f_t g_t$, $\sum h_t^2$
- Step 2: profile $S(t_c, m, \omega) = y^\top y - y^\top X (X^\top X)^{-1} X^\top y$: only three nonlinear parameters, which removes most local minima of the seven-parameter search
- Step 3: $C\cos(x - \phi) = C\cos\phi\cos x + C\sin\phi\sin x$, so $C_1 = C\cos\phi$, $C_2 = C\sin\phi$: $C = \sqrt{C_1^2 + C_2^2} = 0.0632$, $\phi = \operatorname{atan2}(C_2, C_1)$
- Step 4: $B = -1.2 < 0$; $m = 0.5$ and $\omega = 8$ inside their ranges: pass
- $t_c - t_2 = 0.08 \le 0.2$, i.e. 29 days after the window: pass
- $O = \frac{8}{\pi} \ln \frac{1.08}{0.08} = 6.63 \ge 2.5$; $D = \frac{0.5 \cdot 1.2}{8 \cdot 0.0632} = 1.19 \ge 1$: pass
- The relative error, the Lomb test and the residual test need the price data: a candidate, not yet a qualified fit

MFM_ch17_seminar

A7: Fitted LPPL Trend



MFM_ch17_seminar

A8 [Proposed]: Why the Damping Condition? The Hazard Rate — Task

- **Questions**
  - where does the damping condition of the LPPLS filter come from?
  - does a given fit satisfy it?
- **Model** (solved exercises to follow): A7
- **Inputs**: the Johansen–Ledoit–Sornette (JLS) model Johansen, Ledoit & Sornette (2000), under the risk-neutral measure, with zero interest, no dividends and no diffusion
  - $dp/p = \mu(t)\,dt - \kappa\,dj$; $h(t) \geq 0$: the crash hazard rate, $\kappa > 0$: the crash size; $\mathbb{E}_t[dp] = 0$ before the crash requires the drift $\mu(t) = \kappa\,h(t)$ (with a $\sigma\,dW$ term the drift of $\ln p$ would be $\mu - \sigma^2/2$)
  - Log-price trend: $B\tau^m + C\tau^m \cos(\omega \ln \tau - \phi)$, $\tau = t_c - t$
- **Tasks**
  1. Differentiate the trend with respect to t.
  2. Find the condition on (B, C, m, ω) for $h(t) \geq 0$ at every phase $\theta = \omega \ln \tau - \phi$.
  3. Show that for $\omega \gg m$ it becomes the damping condition $D = \frac{m|B|}{\omega|C|} \geq 1$ (Shu & Zhu, 2020a, eq. 11).
  4. For the fit $t_c = 2026.45$, $m = 0.95$, $\omega = 3.5$, $B = -0.8$, $C_1 = 0.25$, $C_2 = 0$ on $[2025.0, 2026.0]$, check the exact condition, D, O and the t_c range.
- **Report**: the three derivations and the four checks
- **Interpretation**
  - what would a negative hazard rate mean economically?
  - is the fit a bubble signature?
- Notebook: section A8

Back Exercise list

B7: Power with Dependent Signal Dates

- For two equally sized groups with large-fall probabilities p_0 (no signal) and p_1 (signal), let $\bar{p} = (p_0 + p_1)/2$. The Normal planning approximation gives observations per group:

$$n_{\text{iid}} = \frac{[1.9600\sqrt{2\bar{p}(1-\bar{p})} + 0.8416\sqrt{p_0(1-p_0) + p_1(1-p_1)}]^2}{(p_1 - p_0)^2}.$$

- The constants give a two-sided 5% test and 80% power. For $p_0 = 0.30, p_1 = 0.40$, round up to about 356 observations in each group.
- For observed signal-on/off groups use the unpooled independent-data standard error, with their own sample sizes:

$$\text{SE}_{\text{iid}} = \sqrt{\frac{\hat{p}_{\text{on}}(1-\hat{p}_{\text{on}})}{n_{\text{on}}} + \frac{\hat{p}_{\text{off}}(1-\hat{p}_{\text{off}})}{n_{\text{off}}}}.$$

- With dependence, use the planning approximation $n_{\text{dep}} = n_{\text{iid}}\,\text{deff}$, where $\text{deff} = (\text{SE}_{\text{block}}/\text{SE}_{\text{iid}})^2$. A design effect of 2 roughly doubles the sample; this is an approximation assuming similar dependence in the larger sample.

B7 [Proposed]: Does the Confidence Indicator Beat Chance? — Task

- **Question**: is a positive LPPLS confidence indicator for the Nasdaq 100 followed by large falls more often than no signal?
- **Model** (solved exercises to follow): B2 and B3
- **Data**: Nasdaq 100 daily closes and the LPPLS confidence indicator from 1997 Demirer et al. (2019); Shu & Zhu (2020a), precomputed in `ch17_lppls_ci_ndx.csv` (chapter Quantlet folder; recomputing is optional): at every 5th trading day t_2, 141 windows of 750 down to 50 days, in steps of 5
  - **Signal**: indicator > 0; **outcome**: a close at least 20% below the signal-date close within 90 **calendar** days; dates with an incomplete window are dropped
- **Tasks**
  1. Find when the indicator was first positive and its share of positive dates between January 1999 and the March 2000 peak.
  2. Estimate $\Delta = \hat{p}_{\mathbf{on}} - \hat{p}_{\mathbf{off}}$ with a 95% confidence interval (CI) from a circular block bootstrap (90 days = 13 dates, 9 999 replications) and a Newey–West test of $b = 0$ (13 lags) in the linear probability model of the slide "Tools for the Extension Exercises" (both variables 0/1).
  3. Compare the block-bootstrap SE with the independent-data SE and compute the design effect $(\mathrm{SE}_{\mathbf{block}}/\mathrm{SE}_{\mathbf{iid}})^{\mathbf{2}}$.
  4. Compute the sample size for 80% power (equal groups, $p_0 = \hat{p}_{\mathbf{off}}$, $p_1 = p_0 + 10$ points, two-sided 5%), for independent observations and for dependent dates.
- **Report**: the chart of the indicator, Δ with its CI and p-value, the design effect and the sample sizes
- **Interpretation**
  - would the indicator have helped an investor in 2000?
  - are signals followed by more or by fewer large falls than no-signal dates, and why?
- Notebook: section B7

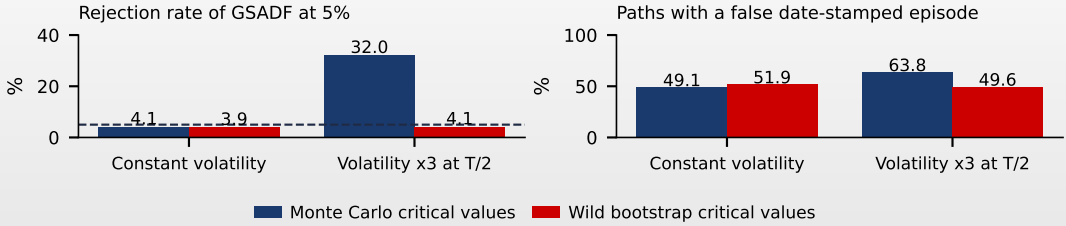
Back Exercise list

B8 [Solved]: Size of GSADF Under a Volatility Shift — Task

- **Questions**
  - how often does GSADF reject a true random walk when volatility triples halfway?
  - does the wild bootstrap fix it?
- **Model** (solved exercises to follow): B2
- **Data**: simulated; 1 000 random walks $y_t = T^{-1} + y_{t-1} + \sigma_t z_t$, $z_t \sim N(0,1)$, $T = 400$, $w_0 = 40$; $\sigma_t = 1$ up to $T/2$, then 3: no bubble Harvey et al. (2016)
- **Tasks**
  1. Compute the rejection rate of GSADF at 5% with the Monte Carlo critical value for constant variance (2.19).
  2. Repeat with a wild bootstrap critical value computed for each path (199 replications, as in B2).
  3. Compute the share of paths with at least one false date-stamped episode (BSADF above its pointwise 95% value for $\lceil \ln T \rceil$ periods).
  4. Do the same for constant volatility as a control and plot the four rates.
- **Report**: a table and a bar chart of the rejection and false-episode rates
- **Interpretation**
  - which critical value would you trust for Bitcoin?
  - does the bootstrap solve the multiplicity of date-stamping?
- Notebook: section B8

Back Exercise list

B8 [Solved]: Solution (1/2) — Chart and Table



	GSADF, MC	GSADF, wild	false ep., MC	false ep., wild
Constant	4.1%	3.9%	49.1%	51.9%
$\times 3$ at $T/2$	32.0%	4.1%	63.8%	49.6%

MFM_ch17_seminar

B8 [Solved]: Solution (2/2) — Steps and Interpretation

- Steps: simulate the path; compute GSADF and BSADF; compare with the fixed Monte Carlo value and with the path's own wild bootstrap values; average over the 1 000 paths (Monte Carlo error ± 1.4 points near 5%)
- The late volatility rise looks like acceleration: 32.0% false rejections with the constant-variance value; the wild bootstrap restores 4.1%
- False episodes in about half of the paths with either value: the bootstrap fixes the size of GSADF, not the multiplicity of date-stamping
- Interpretation: for Bitcoin use the wild bootstrap for the global test and, for the dates, family-wise critical values (which control the probability of a false episode anywhere in the sample)

B9 [Proposed]: Size Under GARCH Volatility — Task

- **Question**: does persistent GARCH volatility (Generalised AutoRegressive Conditional Heteroskedasticity), or a fall in volatility, distort the size of GSADF?
- **Model** (solved exercises to follow): B8
- **Data**: simulated, as in B8, with two new null models: GARCH(1,1) shocks $\sigma_t^2 = 0.02 + 0.10\, e_{t-1}^2 + 0.88\, \sigma_{t-1}^2$ (unconditional variance 1), and a volatility **fall** from 3 to 1 at $T/2$
- **Tasks**
  1. Compute the size of GSADF at 5% with Monte Carlo and with wild bootstrap critical values (1 000 paths, 199 replications each).
  2. Compute the share of paths with a false date-stamped episode and plot the rates as in B8.
- **Report**: a table and a bar chart as in B8
- **Interpretation**
  - does a fall in volatility create false bubbles, as a rise does (B8)? Why?
  - is the wild bootstrap exact under GARCH?
- Notebook: section B9

Back Exercise list

B10 [Proposed]: SADF vs GSADF — Task

- **Question**: which episodes does the recursive ADF (fixed start) miss that BSADF (moving start) finds?
- **Model** (solved exercises to follow): B1
- **Data**: monthly log closes of the Nasdaq 100 and of the S&P 500, 1990–2026 (last close of each month)
- **Tasks**
  1. Compute SADF and GSADF with the minimum-window rule and their 95% critical values from 2 000 simulated random walks.
  2. Date-stamp with the recursive ADF of Phillips–Wu–Yu (PWY) (window start fixed at the first month) and with BSADF of Phillips–Shi–Yu (PSY), minimum duration $\lceil \ln T \rceil$ months Phillips, Wu & Yu (2011).
  3. At each endpoint, take the pointwise 95th percentile across the same 2000 null paths separately for fixed-start ADF and BSADF; do not use the single SADF critical value for dating. List the episodes and plot both statistics with these sequences.
- **Report**: a table of SADF and GSADF, both lists of episodes and the chart
- **Interpretation**: why can a fixed start miss later episodes?
- Notebook: section B10

C2 [Proposed]: Is There an AI Bubble? — Task

- **Question**: is the current rise of NVIDIA (NVIDIA Corporation, a US chip designer) and the Nasdaq 100 more like the dot-com bubble of Cisco, or like explosive fundamentals?
- **Model** (solved exercises to follow): B1, B2, B3; see also the proposed B7
- **Data**: daily, monthly and weekly closes of NVIDIA, Cisco and the Nasdaq 100; quarterly sales and earnings from the quarterly (10-Q) and annual (10-K) reports filed with the Securities and Exchange Commission (SEC), from its EDGAR filings database (Electronic Data Gathering, Analysis, and Retrieval)
- **Tasks**
  1. Run GSADF and BSADF on monthly and weekly data with Monte Carlo and wild bootstrap critical values; date the signals and say whether the current episode is still open.
  2. Compute two-year run-ups, raw and net of the Nasdaq 100, as in Greenwood, Shleifer & You (2019).
  3. LPPLS: NVIDIA and Nasdaq 100 daily log closes from January 2022. Origins every ten observations over the last 130; 141 backward windows (750, 745, ..., 50 observations), with B3 filters.
  4. Build the fundamentals yourselves: filings dated at the filing date (no look-ahead), trailing four-quarter sums, split-consistent per-share values, sales when earnings are ≤ 0.
  5. Test the log price-to-sales (or price-to-earnings) ratio and date-stamp it with family-wise critical values (B8), controlled over a horizon of $\tau_b = 104$ weeks or 24 months Phillips & Shi (2020).
- **Report**: a two-page note with charts and a Quantlet folder (code, a Metainfo.txt file with name, description, keywords and author, charts)
- **Interpretation**: can explosive prices be separated from explosive expected profits Pástor & Veronesi (2009)?
- Notebook: section C2

Back Exercise list

C3 [Proposed]: Does Volume Lead Prices in a Speculative Boom? — Task

- **Question**: does trading volume lead prices in a speculative boom, as in the US housing cycle of 2000–2011 DeFusco, Nathanson & Zwick (2022)?
- **Model** (solved exercises to follow): B2 (Bitcoin episodes); see also the proposed B7 (block bootstrap) and the lecture case study
- **Data**: FRED: monthly Case-Shiller U.S. National Home Price Index (`CSUSHPINSA`) and new one-family houses sold, not seasonally adjusted (`HSN1FNSA`), 2000–2011 and 1990–2011
  - The paper regresses the demeaned log house price on sales lagged k months, $p_t = \beta_k v_{t-k} + \eta_t$ (Eq. 1, η_t: error), in 115 cities; the correlation peaks at $k = 24$ (Figure 2)
- **Tasks**
  1. Build p_t, the log price minus its sample mean, and v_t, sales minus the mean of their calendar month.
  2. For $k = -12, \ldots, 48$ estimate Eq. (1) by OLS without a constant, compute the implied correlation $\hat{\beta}_k \, \mathrm{sd}(v_{t-k})/\mathrm{sd}(p_t)$ and plot it against k.
  3. Give pointwise 95% CIs at the peak lag, at $k = 24$ and at $k = 0$ from a circular block bootstrap with 24-month blocks and 2,000 replications; repeat for 1990–2011.
  4. Date the quiet: from the peak of adjusted volume to the last peak of prices before their decline (Figure 1).
  5. Extension: repeat steps 1–4 for Bitcoin (month-end close in USD, monthly sum of the volume field in BTC-USD.CC.csv) around a BSADF episode (B2); justify any seasonal adjustment Griffin & Shams (2020).
- **Report**: the correlation curve, the peak lag with its CI, the quiet dates, a Quantlet folder and a two-page note
- **Interpretation**: new-home sales exclude resales by short-term buyers: does a national lead of volume confirm speculation, or could credit or construction cycles produce it?
- Notebook: section C3

Back Exercise list

C2 and C3: Assessment Criteria

- C2
  - correct critical values
  - dating without look-ahead information
  - the interpretation
- C3
  - correct critical values and block-bootstrap intervals
  - dating without look-ahead information
  - the interpretation

References (1/2)

- Blanchard, O.J., Watson, M.W. (1982). Bubbles, rational expectations and financial markets. National Bureau of Economic Research (NBER) Working Paper 945.
- Carrasco, M., Hu, L., Ploberger, W. (2014). Optimal test for Markov switching parameters. *Econometrica*, 82(2), 765–784.
- DeFusco, A.A., Nathanson, C.G., Zwick, E. (2022). Speculative dynamics of prices and volume. *Journal of Financial Economics*, 146(1), 205–229.
- Demirer, R., Demos, G., Gupta, R., Sornette, D. (2019). On the predictability of stock market bubbles: evidence from LPPLS confidence multi-scale indicators. *Quantitative Finance*, 19(5), 843–858.
- Diba, B.T., Grossman, H.I. (1988). Explosive rational bubbles in stock prices?. *American Economic Review*, 78(3), 520–530.
- Filimonov, V., Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A*, 392(17), 3698–3707.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). Statistics of Financial Markets: An Introduction, 5th ed. Springer.
- Greenwood, R., Shleifer, A., You, Y. (2019). Bubbles for Fama. *Journal of Financial Economics*, 131(1), 20–43.
- Griffin, J.M., Shams, A. (2020). Is Bitcoin really untethered?. *Journal of Finance*, 75(4), 1913–1964.
- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- Hamilton, J.D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45(1–2), 39–70.

References (2/2)

- Hansen, B.E. (1992). The likelihood ratio test under nonstandard conditions: testing the Markov switching model of GNP. *Journal of Applied Econometrics*, 7(S1), S61–S82.
- Harvey, D.I., Leybourne, S.J., Sollis, R., Taylor, A.M.R. (2016). Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance*, 38, 548–574.
- Johansen, A., Ledoit, O., Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance*, 3(2), 219–255.
- Pástor, Ľ., Veronesi, P. (2009). Technological revolutions and stock prices. *American Economic Review*, 99(4), 1451–1483.
- Pele, D.T., Mazurencu, M., Nijkamp, P. (2013). Herding behaviour, bubbles and log periodic power laws in illiquid stock markets: a case study on the Bucharest Stock Exchange. Social Science Research Network (SSRN) Working Paper.
- Phillips, P.C.B., Magdalinos, T. (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics*, 136(1), 115–130.
- Phillips, P.C.B., Shi, S. (2020). Real time monitoring of asset markets: bubbles and crises. In *Handbook of Statistics: Financial, Macro and Micro Econometrics Using R*, Elsevier, 61–80.
- Phillips, P.C.B., Shi, S., Yu, J. (2015a). Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4), 1043–1078.
- Phillips, P.C.B., Wu, Y., Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: when did exuberance escalate asset values?. *International Economic Review*, 52(1), 201–226.
- Shu, M., Zhu, W. (2020a). Detection of Chinese stock market bubbles with LPPLS confidence indicator. *Physica A*, 557, 124892.
- Shu, M., Zhu, W. (2020b). Real-time prediction of Bitcoin bubble crashes. *Physica A*, 548, 124477.