International Review of Financial Analysis 110 (2026) 104848

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Log-periodicity: Fact or fiction?☆

Klaus Grobys

Finance Research Group, School of Accounting and Finance, University of Vaasa, Wolffintie 34, 65200 Vaasa, Finland

ARTICLE INFO

JEL classification:
C22
C15
C58
C63
G12

Keywords:
False positives
Financial instability
Log-periodic power law model
Bubble detection
S&P 500

ABSTRACT

A common empirical practice in LPPLS applications is to calibrate the model under parameter bounds and then declare an "LPPLS signature" when ADF/PP tests on calibration residuals reject a unit root at conventional tabulated critical values. We show that this procedure exhibits substantial size distortion. Using synthetic series that preserve the roughness and volatility of financial data while excluding log-periodic structure, we compute bootstrap critical values by re-estimating the full two-stage procedure on each synthetic sample. Applied to S&P 500 monthly and daily data, conventional thresholds yield inflated rejection rates. In contrast, the bootstrap restores empirical size to nominal levels and overturns many purported signatures. These findings highlight the need for estimation-aligned inference in LPPLS diagnostics and call for a re-examination of published LPPLS evidence that may reflect size-induced false positives.

1. Introduction

In empirical finance, there is a dangerous temptation to mistake retrospective fit for predictive insight. The Log-Periodic Power Law Singularity (LPPLS) model, introduced by Johansen et al. (2000), claims to reveal hidden regularities—oscillations accelerating toward a critical singularity—that supposedly signal market regime changes. Yet a model that captures the past with precision offers no guarantee of forecasting power. History can be overfit as easily as it can be misunderstood. The LPPLS model's apparent explanatory success demands a more fundamental test: can it distinguish true structure from noise? In scientific modeling, credibility is earned not through the accommodation of known outcomes but through structured attempts at falsification (Edmans, 2024). This paper asks: does the LPPLS model reveal intrinsic market instabilities, or does it merely impose deterministic patterns upon stochastic fluctuations?

To confront this question, we simulate data that keep the familiar features of financial returns—irregularity, heavy tails, volatile bursts—while removing any log-periodic component. This lets us probe directly the LPPLS model's tendency to see patterns in noise. We examine two S&P 500 benchmarks—monthly (1871–2022) and daily (1980–1986)—to test robustness across horizons (e.g., Grobys, 2023; Sornette, 2017). In doing so, we follow common practice in the applied LPPLS literature by enforcing standard parameter bounds (see Sornette, 2017). In line with standard practice, we assess the stationarity of LPPLS residuals using Augmented Dickey–Fuller (ADF) tests (e.g., Grobys, 2023, 2025; Lin et al., 2014). We first compare the residual statistics to conventional tabulated ADF critical values and record the share of samples labeled "LPPLS-consistent," treating such rejections as false positives under the simulation-based null. We then implement an i.i.d. bootstrap, re-estimating the full two-stage procedure on each synthetic sample to obtain bootstrap critical values with correct empirical size. To benchmark these findings within a familiar approach, we also apply a full-sample right-tailed ADF test on log prices (with an intercept and BIC lag selection) and compute both Monte Carlo and bootstrap critical values, thereby contrasting LPPLS-residual evidence with a standard test for mild explosiveness. Finally, to assess external validity, we replicate the estimation–testing procedure on gold futures (using the same data and settings as Grobys, 2025) as an additional robustness check. This design makes clear how tabulated critical values can inflate bubble detections and, in turn, provides a more reliable framework for LPPLS-based diagnostics of speculative bubbles.

Brée et al. (2013) and Brée and Joseph (2013) document an extensive literature applying various LPPLS models to identify financial market bubbles. A phenomenon well-established in the study of financial modeling, particularly in Generalized Autoregressive Conditional

☆ This paper was presented at the Erich-Schneider Seminar in April 2025 at Christian-Albrechts-Universität zu Kiel. The author gratefully acknowledges valuable comments received from Thomas Lux and Daniel Fehrle. Furthermore, the author thanks two anonymous referees for their constructive and helpful comments.

E-mail address: klaus.grobys@uwasa.fi.

https://doi.org/10.1016/j.irfa.2025.104848
Received 20 August 2025; Received in revised form 26 November 2025; Accepted 3 December 2025
Available online 4 December 2025

Heteroskedasticity (GARCH) frameworks, manifests here: when the model fails, it is retrofitted rather than discarded (Mandelbrot, 2008).[1] The tendency of financial theorists to revise parameters post hoc rather than reassess foundational assumptions has profound implications for the empirical validity of LPPLS model applications. Brée and Joseph (2013) tested the standard LPPLS model, originally introduced in Johansen et al. (2000), across 11 stock market crashes in the Hang Seng Index (1970–2008), finding that only seven conformed to the parameter constraints proposed ex post by Johansen and Sornette (2001a). The apparent inconsistency in these calibrations underscores the model's selective alignment with observed bubbles rather than systematic validation. Whereas Brée and Joseph (2013) center their analysis on true positives, Lin et al. (2014) undertake the only extensive examination of false positives, probing whether the LPPLS model falsely signals bubbles in non-speculative data. Their work extends the LPPLS framework by introducing the Volatility-Confined LPPLS (VC-LPPLS) model, incorporating mean-reverting residuals and benchmarking its robustness against synthetic GARCH-driven datasets. Their findings indicate an exceptionally low false positive rate (0.2 %) when tested on these artificially generated series. Yet, this inference rests upon the premise that GARCH processes adequately approximate financial market dynamics—a premise contested by empirical studies revealing that real-world asset price movements frequently diverge from the volatility clustering intrinsic to GARCH frameworks. Thus, while Lin et al. (2014) assert a methodological improvement, their conclusions remain circumscribed by their reliance on simulated GARCH-driven benchmarks. This limitation highlights the need for a broader validation framework that systematically evaluates the LPPLS model's reliability in detecting speculative bubbles in financial markets.

Notably, the introduction of volatility constraints, as proposed by Lin et al. (2014), presents several methodological concerns. The standard LPPLS model is parsimonious, permitting robust calibration without additional constraints. Given that financial markets inherently exhibit heteroskedasticity, the log-periodic component of LPPLS model already accommodates volatility clustering. Imposing an explicit volatility function risks overfitting, introducing parameter uncertainty, and diminishes predictive reliability. Empirical studies further suggest that log-periodic oscillations alone suffice for bubble detection, with volatility restrictions offering only marginal improvements, if any (Gustavsson et al., 2016; Shu & Song, 2024). Moreover, the predominant body of research employs the standard LPPLS model, reinforcing its centrality in empirical investigations. Prioritizing this formulation ensures direct assessment of the model's efficacy without confounding influences from additional constraints. This choice preserves analytic tractability while avoiding excessive parameterization that could obscure fundamental log-periodic structures. Thus, the LPPLS model, in its original form, remains a fundamental and pragmatic tool for financial market analysis, providing a reliable framework for identifying speculative regimes while sidestepping unnecessary complexities that may cloud empirical findings.

Moreover, using a GARCH model to simulate returns specifies conditional variance dynamics for returns: when $\alpha + \beta < 1$, returns are covariance-stationary and the conditional variance is mean-reverting. However, for markets such as the S&P 500, empirical work often points to highly persistent volatility (near-IGARCH), so a single low-order parametric GARCH—even with Student-t innovations—may underrepresent salient features of the data-generating process for the for the purposes of LPPLS residual ADF testing. Consequently, false-positive rates may be understated when LPPLS residual tests are judged against conventional tabulated critical values. To avoid embedding a particular volatility law in the null, we employ an i.i.d. bootstrap of returns—as opposed to using GARCH models—compounding to prices to preserve the empirical marginal distribution while excluding any imposed log-periodic structure. We then re-estimate the LPPLS models and re-test on each resample, constructing residual-ADF critical values that are aligned with the exact two-stage procedure (same bounds, starting values, and BIC lag selection).

Our study reframes LPPLS diagnostics around size-controlled inference. We follow Sornette's (2017) practice of enforcing parameter bounds in LPPLS model calibration and then examine the common next step in the literature: applying ADF tests to the resulting residuals and consulting tabulated critical values (e.g., Lin et al., 2014). The premise is simple but consequential: ADF tables for raw data series do not generally apply to residual data from a nonlinear model with estimated and constrained regressors. Uncritically applying tabulated critical values may induce size distortion and encourage seeing structure in noise. Our contribution is to align inference with estimation. We construct the residual-ADF null distribution under the exact two-stage procedure—using the same bounds, starting values, and BIC-based lag selection—via an i.i.d. bootstrap that re-estimates the LPPLS model and re-tests on every resample. This estimation-aligned bootstrap yields critical values that restore empirical size to nominal levels and, in turn, reduce spurious detections of log-periodicity.

We further contextualize the residual evidence against a standard ADF benchmark on log prices (BADF) and extend the design to gold futures to gauge external validity. The guiding message for empirical practice is straightforward: inference should be calibrated to the way the statistic is produced; when it is, apparent regularities may recede toward chance.

Applied to the original S&P 500 data, the LPPLS fit suggests moderate acceleration toward a critical point in the monthly series and a more pronounced acceleration in the daily series. However, using tabulated ADF critical values, the residual ADF indicates the residuals are stationary at conventional cutoffs. Yet the synthetic-data validation tells a different story. At the nominal 1 % level, the empirical rejection rate under the bootstrap null is 84.8 % (monthly) and 82.8 % (daily)—a severe oversizing. Using bootstrap-based critical values, the LPPLS residual ADF statistics for the original datasets ($\widehat{\lambda}_{\mathrm{ADF}} = -3.2277$ monthly; $\widehat{\lambda}_{\mathrm{ADF}} = -2.6194$ daily) are not significant (bootstrap $p = 0.5960$ and $p = 0.8130$, respectively). As a levels-based benchmark, we compute a full-sample BADF with an intercept and BIC-selected lags. Evaluated against the tabulated 1 % critical value, the empirical size under the bootstrap-simulated null is 13 %—still oversized, but well below the residual-ADF case. Using bootstrap inference, the BADF indicates mild explosiveness for the monthly S&P 500 sample (bootstrap $p = 0.0370$) and no such evidence for the daily sample (bootstrap $p = 0.5930$). A robustness check on gold futures points the same way. At the nominal 1 % level, the bootstrap size is 81 % for the residual ADF. Evaluating the original residual statistic against bootstrap cutoffs ($\widehat{\lambda}_{\mathrm{ADF}} = -4.6215$) yields $p = 0.0580$: reject at 10 % but not at 5 %—a marginal signal consistent with Grobys (2025). Taken together, the evidence is plain: empirical ADF quantiles relevant for LPPLS residuals sit far below tabulated critical values, and relying on standard tables induces substantial size distortion. Calibrating inference to the way the statistic is produced changes

[1] The LPPLS model, first formulated by Johansen et al. (2000), has not remained static; rather, it has undergone successive modifications, each seeking to refine its predictive capabilities and extend its applicability within financial markets. A notable revision is the approach of Shu and Song (2024), which sharpens critical-time estimation by incorporating singularity corrections. Another substantive refinement is the Volatility-Confined LPPL model, embedding a mean-reverting volatility process to fortify robustness against market fluctuations (Lin et al., 2014). The LPPL framework has further been expanded to encompass negative bubbles (anti-bubbles)—capturing self-reinforcing downward spirals and subsequent rebounds (Yan et al., 2010). Beyond structural enhancements, researchers have sought to integrate fundamental economic factors, aiming to bridge the gap between speculative dynamics and macroeconomic underpinnings (Zhou & Sornette, 2006). Additional developments include genetic algorithm optimization for parameter calibration (Filimonov & Sornette, 2013) and second-order LPPL models, designed to capture more intricate price fluctuations (Sornette & Zhou, 2002).

the picture; many apparent regularities recede toward chance.

This paper is organized as follows: The next section provides an overview of recent literature on the application of the LPPLS model in various financial markets. Section three presents the datasets, while section four describes the methodology. Section five presents the results, followed by a discussion in section six. Finally, the last section concludes.

2. Literature review

The LPPLS model, initially introduced in the seminal work of Johansen et al. (2000), has since been subjected to extensive empirical scrutiny across diverse financial domains. This approach has been particularly employed to characterize speculative phenomena and detect instability in cryptocurrency, stock, energy, and carbon credit markets. Sornette (2017) provides a comprehensive survey of studies employing the LPPLS model to examine financial bubbles, whereas the present discussion focuses on a selective yet representative set of recent contributions to the literature.

One recent line of inquiry concerns the application of the LPPLS framework to cryptocurrency markets, where it serves as an instrument to identify speculative phases and abrupt shifts in valuation. The work of Ahn et al. (2024), Grobys (2024), and Zhang et al., 2024 exemplifies this effort. These studies reveal that super-exponential price escalations in Bitcoin and Ethereum exhibit distinct LPPLS signatures prior to substantial corrections. However, methodological variations exist. While Ahn et al. (2024), Van Eyden et al. (2023), Johansen and Sornette (2001b), and Grobys (2024) adopt the conventional LPPLS framework to analyze historical data, Zhang et al., 2024 incorporates wavelet analysis to enhance granularity in bubble detection. Further distinctions arise in the temporal scope of inquiry: whereas Grobys (2024) spans more than a decade of Bitcoin price movements, Ahn et al. (2024) centers on discrete speculative intervals. Despite these variations, the studies collectively underscore the utility of the LPPLS model in diagnosing instability within digital asset markets.

A second avenue of recent research concerns equity markets, where the LPPLS model has been deployed to detect speculative surges and anticipate crises in various national stock indices (e.g., Grobys, 2023; Gupta et al., 2023, 2025; Johansen & Sornette, 2001b). Investigations such as those of Cepni et al. (2025) and Zhao and Sornette (2021) share the common objective of refining bubble detection by incorporating LPPLS confidence metrics. Some studies, such as Ji and Zhang (2024), advance the methodology by introducing a Sequential Quadratic Programming algorithm to optimize parameter estimation, while others, such as Zhao and Sornette (2021), depart from traditional implementations by integrating event-study techniques to analyze post-bubble dynamics. Furthermore, the contributions of Shu et al. (2021) and Song et al. (2022) differentiate between endogenous and exogenous crashes, illustrating the adaptability of the LPPLS methodology to distinct modeling paradigms. Collectively, this body of work affirms the relevance of the LPPLS model in equity market research, while also demonstrating the necessity of methodological refinement to enhance reliability.

Parallel efforts have been directed toward energy and commodity markets, where the LPPLS model is employed to discern speculative fluctuations in crude oil, gold, and agricultural commodities. Empirical studies, including those by Gupta et al. (2024), Chang (2024), Cifarelli and Paesani (2021), and Grobys (2025), affirm the applicability of LPPLS in this domain. Nonetheless, the specific methodological approaches diverge: whereas Chang (2024) and Xu et al., 2025 supplement the LPPLS model with econometric techniques such as GARCH and Markov regime-switching models, Cifarelli and Paesani (2021) introduce a heterogeneous-agent framework to distinguish between fundamentalist and speculative market participants. Additionally, Grobys (2025) applies a traditional power-law framework to gold futures, diverging from oil-focused analyses. Yang et al. (2024), meanwhile, integrates machine-learning classification to assess the interaction between monetary policy uncertainty and energy price bubbles. These methodological variations highlight both the versatility of the LPPLS model framework and the necessity of auxiliary statistical tools in specific financial contexts.

A final stream of recent literature explores speculative episodes in carbon credit and environmental markets, with studies such as those by Ghosh et al. (2021) and Huang and Wang (2024) investigating the influence of regulatory interventions on bubble formation. While both studies confirm that the LPPLS model is a viable instrument for detecting speculative surges in carbon credit pricing, their approaches differ. Huang and Wang (2024) employ Supremum Augmented Dickey-Fuller Test (SADF) and Generalized Supremum Augmented Dickey-Fuller Test (GSADF) unit root tests to validate bubble episodes, whereas Ghosh et al. (2021) relies exclusively on LPPLS confidence indicators. Additionally, their scopes diverge: Huang and Wang (2024) examine multiple emissions trading schemes across different jurisdictions, whereas Ghosh et al. (2021) focuses specifically on carbon credit ETFs. Despite these methodological distinctions, both studies converge on the conclusion that speculation in carbon markets is significantly influenced by policy regimes and macroeconomic conditions, reinforcing the importance of regulatory oversight.

Across all domains of financial research, the LPPLS model remains a central tool for characterizing speculative bubbles, yet considerable variation exists in its empirical implementation. A primary point of differentiation concerns validation strategies: whereas studies in equity markets frequently benchmark LPPLS model forecasts against historical market downturns, as in Shu et al. (2021) and Song et al. (2022), others, such as Zhao and Sornette (2021), emphasize post-bubble event analysis. Additionally, robustness assessments vary; LPPLS confidence indicators are prevalent in stock market research but less frequently utilized in cryptocurrency and energy market studies. Furthermore, while econometric supplements such as GARCH and Markov-switching models may enhance LPPLS model applications in commodity markets, alternative enhancements—such as machine-learning algorithms and wavelet decomposition—have emerged in cryptocurrency research. In sum, while the LPPLS model appears to consistently detect speculative behavior across diverse markets, its predictive efficacy is contingent upon methodological refinements tailored to specific financial contexts.

3. Data

We investigate log-periodicity using monthly and daily financial data in the main analysis. Consistent with Grobys (2023), the monthly dataset for the S&P 500 index was sourced from Robert Shiller's publicly accessible data library (www.econ.yale.edu/~shiller/data.htm). To maintain comparability with prior research, the present study utilizes the same data sample as that of Grobys (2023), encompassing the period from January 1871 to November 2022. In addition to the monthly data, daily observations of the S&P 500 were obtained for the period spanning January 2, 1980, to December 31, 1986, from www.investing.com. This dataset aligns with the sample employed by Sornette (2017) and Grobys (2023) in the calibration of the LPPLS model. Notably, this period concludes 202 trading days prior to the widely recognized stock market crash of October 19, 1987.

Grobys (2023) identifies the October 1987 crash as an enduring intellectual enigma for at least three principal reasons. First, the economic significance of a single-day decline exceeding 20 % in equity market capitalization was extraordinary. Second, the probabilistic occurrence of such an event was anomalous relative to conventional financial models, an observation echoing Mandelbrot's assertion that it constituted "a number outside the scale of nature" (Mandelbrot, 2008, p. 4). Third, the crash transpired in the absence of discernible premonitory signals, a phenomenon that remains unexplained within the framework of traditional financial theories. Similarly, Sornette (2017) underscores that extensive research has sought to elucidate the underlying causes of the

October 1987 crash, particularly by examining trading behaviors and market structures. Despite these efforts, no definitive cause has been identified. Notably, the sharp market decline observed in October 1987 was preceded by an exceptional market surge over the first nine months of the year, a pattern evident across multiple economies. In the United States, for instance, stock prices recorded a substantial increase of 31.4 % during this period. Some analysts contend that the downturn in October was a consequence of the preceding speculative bubble, driven by excessively inflated asset prices.

The October 1987 stock market crash is widely regarded as among the most significant financial collapses in modern history to date. If the dependency structures present in the monthly and daily datasets manifest as periodic oscillations—oscillations that, according to theoretical models, culminate in finite-time singularities—then employing i.i.d. bootstrapped data, as formulated by Efron (1992), would effectively eliminate these dependencies. Therefore, for each data set $j = (1, 2)$, we compute $b = 1, \ldots, B$ synthetic data sets as follows:

(a) First, for each given data set, we compute the log-returns of the data: $r_{j,t} = \ln\left(\frac{P_{j,t}}{P_{j,t-1}}\right)$, where $P_{j,t}$ denotes the price of the S&P 500 at time unit t for a given data set j.

(b) Second, we bootstrap each data vector j of log-returns using random sampling with replacement giving us B bootstrapped log-return vectors for each data set:

$$\left(\boldsymbol{r}_{j,1} \quad \boldsymbol{r}_{j,2} \quad \ldots \quad \boldsymbol{r}_{j,B}\right) = \begin{pmatrix} r_{j,1,2} & r_{j,2,2} & \ldots & r_{j,B,2} \\ r_{j,1,3} & r_{j,2,3} & \vdots & r_{j,B,3} \\ \vdots & \vdots & & \vdots \\ r_{j,1,T} & r_{j,2,T} & \ldots & r_{j,B,T} \end{pmatrix}_{(T-1),B}$$

(c) Third, for each data set, we compute B synthetic S&P 500 indices by means of compounding as: $P_{j,b,t} = P_0 \exp\left(\sum_{s=1}^{t} r_{j,b,s}\right)$, with P_0 denoting the initial price of the corresponding original data set j:

$$\left(\boldsymbol{P}_{j,1} \quad \boldsymbol{P}_{j,2} \quad \ldots \quad \boldsymbol{P}_{j,B}\right) = \begin{pmatrix} P_{j,1} & P_{j,1} & \ldots & P_{j,1} \\ P_{j,1,2} & P_{j,2,2} & \vdots & P_{j,B,2} \\ \vdots & \vdots & & \vdots \\ P_{j,1,T} & P_{j,2,T} & \ldots & P_{j,B,T} \end{pmatrix}_{T,B}$$

(d) Finally, the synthetic S&P 500 indices $\left(\boldsymbol{P}_{j,1} \quad \boldsymbol{P}_{j,2} \quad \ldots \quad \boldsymbol{P}_{j,B}\right)$ are transformed in terms of their natural logarithms.

This procedure is employed to generate $B = 1000$ synthetic log-indices. Tables A.1 and A.2 present the descriptive statistics for the original datasets, whereas Tables A.3 and A.4 summarize the corresponding statistics for the sample means obtained from i.i.d. bootstrapped samples, $\left(\bar{\boldsymbol{r}}_{j,1} \quad \bar{\boldsymbol{r}}_{j,2} \quad \ldots \quad \bar{\boldsymbol{r}}_{j,B}\right)$. A comparative analysis between Tables A.1 and A.3 reveals that the mean of the bootstrapped sample means (0.3710) closely approximates the sample mean of the original data (0.3696), a pattern similarly observed when comparing Table A.2 with Table A.4. This outcome is anticipated, as the bootstrapping methodology inherently preserves statistical properties, ensuring consistency in sample means. Figs. 1–2 display alternative historical paths of the S&P 500 generated from bootstrapped log-compounded returns for the monthly and daily samples, respectively. Because the monthly panel spans a long horizon, compounding magnifies small return differences into large dispersion in price levels. The upper tail of the bootstrap (e.g., the 97.5 % quantile) therefore explodes and dominates the vertical scale, obscuring the bulk of the distribution. To preserve readability, Fig. 1 is plotted in log prices. In both figures, the realized series and the bootstrap median lie very close at the terminal date T—as expected under our resampling scheme: starting from the same initial value and drawing returns from the empirical distribution, the median compounded path tracks the data's unconditional drift and thus aligns with the observed endpoint up to sampling variation.

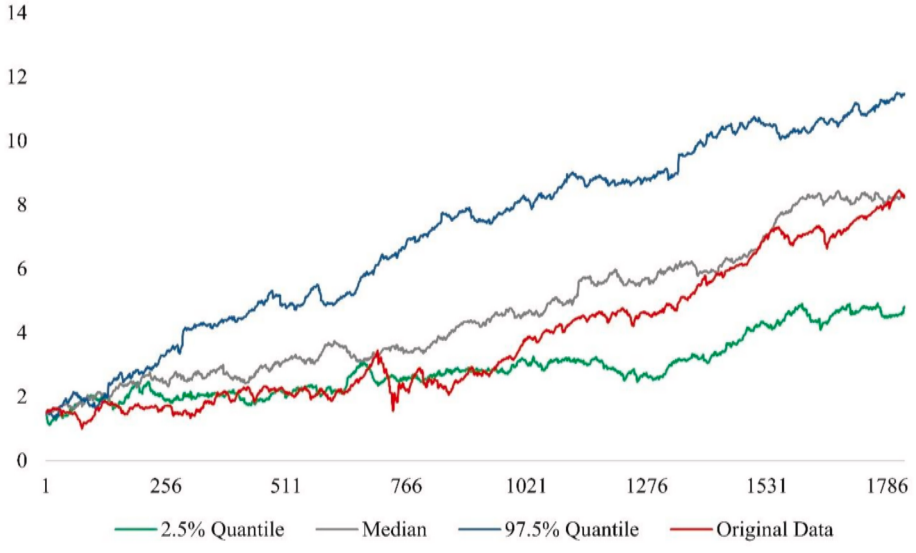

Fig. 1. Alternative historical trajectories for the S&P 500 based on log-compounded returns for January 1871–November 2022.
This figure illustrates 1000 synthetic trajectories of the S&P 500 index generated by compounding monthly log-returns resampled with replacement from the historical data (January 1871–November 2022, source: Shiller database). Prices are plotted in natural logarithms to preserve readability. The red solid line denotes the realized historical path, while the other lines depict the empirical dispersion of the bootstrap simulations. The proximity of the realized and median synthetic paths at the terminal point reflects that the bootstrap preserves unconditional drift but removes log-periodic dependencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

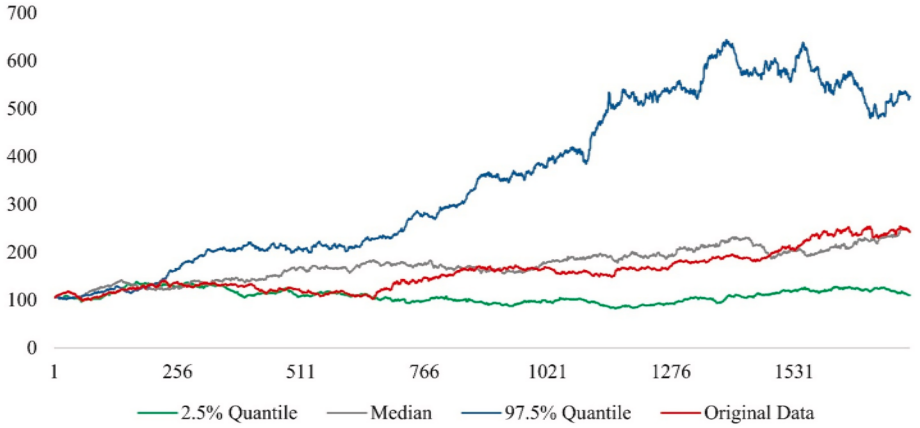

Fig. 2. Alternative historical trajectories for the S&P 500 based on log-compounded returns for January 2, 1980–December 31, 1986.
This figure presents 1000 synthetic trajectories of the S&P 500 index obtained from daily i.i.d. bootstrap resampling of returns from January 2, 1980, to December 31, 1986 (source: Investing.com). The red line indicates the observed price series, and the other lines show the range of bootstrapped paths. The median bootstrap path tracks the realized series closely, demonstrating that the procedure preserves first-moment properties while eliminating periodic structure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Methodology

4.1. Main analysis

4.1.1. Implementing the LPPLS model using log-prices of the S&P 500

A plain power law model for financial log-prices is given by the following:

$$\ln[P_t] = A + B(t_c - t)^{\beta}, \tag{1}$$

where $\ln[P_{j,t}]$ denotes the logarithm of the value of the S&P 500 index at time t, t_c is the critical time, A is the expected value of the logarithmic S&P 500 when approaching t_c, B defines the exposure to faster-than-exponential growth, and β is the power law exponent controlling faster-than-exponential price growth (Sornette, 2017). The critical time t_c indicates the end of the accelerating oscillations, which results in a finite-time singularity manifested in a regime change (Zhang et al., 2016). According to Sornette (2017), the simple power law model of Eq. (1) needs to be extended by accounting for periodic oscillations:

$$\ln[P_t] = A + B(t_c - t)^{\beta}[1 + C\cos(\omega\ln(t_c - t) + \phi)], \tag{2}$$

where C denotes the exposure of the log-periodic oscillations around the power law singular growth, ω denotes the angular log-frequency of oscillations during the formation of the bubble, ϕ is the phase parameter, and all other notations are as previously defined. The LPPLS model of Eq. (2) is first implemented for the original data sets on the S&P 500 using the following set of constraints (Sornette, 2017):

$$-\infty < A < \infty,$$

$$-10 < B < 10,$$

$$0.1 \le \beta \le 0.9,$$

$$|C| < 1,$$

$$5 \le \omega \le 15,$$

$$T \le t_c \le 2T,$$

$$-\pi < \phi < \pi.$$

Note that $-10 < B < 10$ allows growth/decay flexibility without excessive divergence, $0.1 \le \beta \le 0.9$ ensures valid power-law behavior (prevents infinite variance), $|C| < 1$ ensures oscillations remain moderate, $5 \le \omega \le 15$ prevents extreme oscillations while allowing flexibility, $T \le t_c \le 2T$ ensures the critical time is in the future but within a reasonable range, and $-\pi < \phi < \pi$ constrains the phase shift to avoid numerical instability.

Furthermore, the LPPLS model is calibrated using the following initial starting values:

$$A = \ln[P_T] + 1.00,$$

$$B = -0.10,$$

$$\beta = 0.50,$$

$$C = 0.01,$$

$$\omega = 6.00,$$

$$t_c = T + 10,$$

$$\phi = 0.00.$$

Note that, $B = -0.10$ is consistent with the expected super-exponential growth, $\beta = 0.50$ is chosen within a valid theoretical range corresponding to $0.1 \le \beta \le 0.9$, $C = 0.01$ initiates small oscillatory contribution to start, $\omega = 6.00$ corresponds to a common log-periodic oscillation frequency, $t_c = T + 10$ ensures the predicted event arrives beyond the dataset, and $\phi = 0.00$ is initially set to zero, but optimized during the calibration.

After calibrating the LPPLS model using the original data sets, the model is calibrated for the synthetic datasets of S&P 500 log-prices. To ensure comparability, each LPPLS model calibration uses the same sample constraints and initial parameter values.

4.1.2. Testing the LPPLS signatures for statistical significance

Lin et al. (2014) test the LPPLS hypothesis within a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) framework. They first generate synthetic financial time series from GARCH(1,1) processes and calibrate LPPLS on rolling windows. Crucially, they then *condition* on LPPLS fits whose parameters fall within the canonical bounds (e.g., on β, ω, C, t_c). For those parameter-admissible fits, they test residuals,

$$\ln P_t - \widehat{\ln P_t} = \widehat{u}_t, \tag{3}$$

for stationarity using Augmented Dickey–Fuller (ADF) regressions,

$$\Delta\widehat{u}_t = \delta\widehat{u}_{t-1} + \gamma_1\Delta\widehat{u}_{t-1} + \cdots + \gamma_p\Delta\widehat{u}_{t-p} + \varepsilon_t, \quad (4)$$

where ε_t is assumed i.i.d., and rejection of a unit root at conventional tabulated levels (e.g., 5 %, 1 %, 0.1 %) is interpreted as evidence of a statistically significant LPPLS "signature." Under this conditioned procedure, Lin et al. (2014) report a very low false-positive rate in GARCH-generated series (about 0.2 % of samples satisfy the bounds *and* yield stationary residuals). In their study, unit-root tests on residuals reject stationarity in non-bubble periods, whereas parameter-admissible, LPPLS-identified episodes exhibit stationary residuals—suggesting, at least preliminarily, that the LPPLS model can distinguish bubble regimes from heteroskedastic but otherwise standard dynamics.

Note that in the study of Lin et al. (2014), the GARCH model is used to simulate returns, which are then compounded to prices before calibrating the LPPLS model under parameter bounds; residuals from admissible LPPLS model fits are subsequently tested with ADF tests using tabulated critical values. This design treats GARCH as a generator of conditional variance dynamics for returns—under $\alpha+\beta<1$, returns are covariance-stationary and conditional variance is mean-reverting—while the LPPLS framework specifies a deterministic structure in (log) price levels with a power-law acceleration and log-periodic oscillations. However, for asset markets like the S&P 500, empirical work often points to highly persistent volatility (near-IGARCH), so a single low-order parametric GARCH model (or even GARCH models with *t*-distributed innovations) may underrepresent salient features of the data-generating process for the purposes of LPPLS residual ADF testing. To avoid embedding a particular volatility law in the null, our study uses an i.i.d. bootstrap of returns, compounding to prices to preserve the empirical marginal distribution while excluding any imposed log-periodic structure. We then re-estimate LPPLS and re-test on each resample, constructing residual-ADF critical values that are aligned with the exact two-stage procedure (same bounds, starting values, and BIC lag selection).

As outlined in Section 3, the first dataset comprises bootstrapped monthly observations spanning January 1871 to November 2022, while the second dataset consists of bootstrapped daily observations covering the period from January 2, 1980, to December 31, 1986. As noted earlier, our i.i.d. resampling removes serial dependence by construction, enabling a more rigorous evaluation of the LPPLS signature (log-periodic structure) versus noise. Consequently, the ADF test statistics are estimated for the residuals of LPPLS models calibrated to synthetic data and compared with those derived from the original data.

4.2. Additional analysis

4.2.1. Benchmarking LPPLS-residual ADF false-positive rates against full-sample BADF tests

To benchmark the false-positive rates from the LPPLS-residual ADF test we perform additional analysis by implementing a full-sample ADF test for mild explosiveness on log prices. Specifically, we assess whether a plain ADF test on levels—implemented with a comparable sample length T and the same BIC lag-selection rule—yields higher, lower, or similar false-positive rates relative to the ADF test applied to LPPLS residuals. For both tests, we report two sets of critical values: (i) Monte-Carlo "tabulated" values computed under a generic unit-root null (a model-agnostic baseline), and (ii) bootstrap-based values obtained by i.i.d. resampling from the *same* synthetic data-generating process used earlier to calibrate the LPPLS results, re-estimating the full estimation–test sequence on each resample. Using an identical synthetic generator ensures a methodologically aligned comparison across testing procedures. This design directly contextualizes the estimated false positive rates for the LPPLS residual tests against a widely used ADF-type benchmark under matched sample and lagging conditions.

It is important to note that standard ADF response-surface tables (e.g., MacKinnon) report left-tailed critical values for tests of stationarity under various deterministic specifications, but they do not provide right-tailed cutoffs for the explosive alternative $H_1:\rho>0$. In addition, our implementation uses finite-sample BIC lag selection, which alters the null distribution relative to fixed-lag tables. For these reasons, and to supply a standard, model-agnostic benchmark, we compute right-tailed ADF critical values by Monte Carlo under a generic unit-root null calibrated to our sample length T, deterministic specification, and lag-selection rule. This follows common practice in the explosive-root literature, where inference for right-tailed ADF-type tests is obtained by simulation rather than pre-tabulated values. We then apply the full-sample right-tailed ADF to log prices and evaluate significance using these Monte-Carlo critical values. In parallel, we also report bootstrap-based critical values obtained by i.i.d. resampling (re-estimating the procedure on each resample) for both the LPPLS-residual ADF and the ADF benchmark, to examine whether size is restored when inference matches the finite-sample procedure.

Furthermore, our empirical design chosen for our main analysis estimates a single LPPLS model specification on a fixed window (monthly or daily) and then makes one test decision based on the residual ADF test—there is no scanning over sub-samples. The most appropriate benchmark is the full-sample right-tailed ADF on log prices (equivalently, the backward ADF (BADF) evaluated at the terminal date T in the Phillips–Shi–Yu framework): it applies the *same decision logic* (one test on one sample), uses the *same finite-sample ingredients* (BIC lag selection, given T), and targets the *same alternative* (mild explosiveness, $H_1:\rho>0$). By contrast, SADF and especially GSADF are search procedures: they maximize ADF test statistics over expanding or moving windows and reject if *any* window is explosive. That supremum structure introduces a built-in multiple-testing component and a different null distribution; using SADF/GSADF as the benchmark would conflate differences in window-search multiplicity with differences in model performance, and would not reflect how LPPLS is applied here. Finally, among alternatives, ADF-type tests are the most widely used and transparent baseline for (non-)stationarity and mild explosiveness in financial time series, whereas Markov-switching models require stronger parametric assumptions (regime number, transition structure), are sensitive to starting values and identification, and yield results that are less directly comparable to a single-window LPPLS decision rule. For these reasons, we use the full-sample ADF (BADF) as the benchmark and report both Monte-Carlo "tabulated" and bootstrap-based critical values to place our LPPLS findings in a standard, methodologically aligned context.

We design the Monte Carlo experiment for the null data-generating process as follows: For each replication $b=1,\ldots,B$ we generate a unit-root series of the length $T=2000$ similar as in the empirical application[2]:

$$y_0^{(b)} = 0,$$

$$y_t^{(b)} = y_{t-1}^{(b)} + u_t^{(b)},$$

$$u_t^{(b)} \sim \mathcal{N}(0,1), t=1,\ldots,T.$$

The innovation variance is without loss because the ADF t-statistic is scale-invariant under H_0. On each simulated path $y_{1:T}^{(b)}$ we estimate the standard ADF regression

$$\Delta y_t = \alpha + \gamma t + \rho y_{t-1} + \sum_{i=1}^{p}\phi_i\Delta y_{t-i} + \varepsilon_t,$$

with the deterministic specification fixed ex ante to a constant term that is $\alpha\neq 0$ and $\gamma=0$. It is noteworthy that the ADF test regression used

[2] Using the exact sample lengths $T=1770$and $T=1823$ leads to the same critical values, up to negligible Monte Carlo error.

here deviates from the main analysis and this is for the following reason: The null model used to obtain critical values determines their distribution. Our Monte Carlo "tabulated" critical values are generated under a generic Gaussian unit-root process with no drift and the same BIC lag rule. By contrast, the bootstrap critical values are generated from i.i.d. resampled S&P 500 returns, which are chained back to price levels, thereby inheriting the marginal features of the data (e.g., non-zero mean, heavy tails) but not serial dependence. Under a right-tailed ADF test without an intercept (AR specification), even a modest positive drift and fat-tailed innovations shift mass into the right tail of the statistic, producing much larger critical values—this reflects the chosen null, not an error. Moreover, with long samples and BIC typically selecting few lags, standard errors are small, so slight positive deviations of $\widehat{\rho}$ from zero can yield large t-statistics, further thickening the right tail under the bootstrap null. These considerations motivate us to employ the benchmark ADF test on log prices with an intercept to account for drift in equity indices, providing a specification that is methodologically aligned with the data-generating features.

The lag length p is selected by BIC over $p = 0,\ldots,K$, where we use a Schwert-style cap $K = \lfloor 12(T/100)^{1/4} \rfloor$. For each replication we record the ADF t-statistic on y_{t-1} (the coefficient ρ). Using the B simulated t-statistics, we estimate the left- and right-tailed critical thresholds via the empirical quantiles of the simulated distribution.

- Left-tail (stationarity alternative $H_1 : \rho < 0$): the 10th, 5th, and 1st percentiles $(Q_{0.10}, Q_{0.05}, Q_{0.01})$.
- Right-tail (mild explosiveness $H_1 : \rho > 0$): the 90th, 95th, and 99th percentiles $(Q_{0.90}, Q_{0.95}, Q_{0.99})$.

The Monte-Carlo "tabulated" critical values are calibrated to our empirical design—they match the sample length, impose the presence of a constant term, and implement the identical BIC lag-selection rule—yet they remain model-agnostic, as they are generated under a generic unit-root null rather than via the LPPLS model estimation procedure. Note again that all inference is conducted on a single, prespecified full sample per dataset. No rolling-window search is performed; consequently, no multiple-testing adjustment is required. Our benchmark BADF refers to the full-sample ADF with lag length selected by BIC, against which we report empirical size and rejection behavior.

4.2.2. Gold futures: cross-market robustness and replication

To assess whether our conclusions extend beyond U.S. equities, we replicate the entire estimation–testing procedure on gold futures. Gold is an instructive comparison for three reasons. First, it is a large and systemically relevant market that attracts sustained attention from investors and policymakers. Second, it has been the focus of renewed interest in recent years, with research documenting structural shifts in demand and pricing dynamics. Third, exchange-traded futures provide liquid, high-frequency price series with transparent contract specifications, which are well suited to the bootstrap resampling and re-estimation design used here. Using the same daily gold futures data as in the study of Grobys (2025)—December 2, 2015 to November 6, 2024—and applying the same LPPLS model estimation settings as in our S&P 500 analysis (Section 4.1.1), we ensure results that are directly comparable across asset classes and, at the same time, conduct a scientific replication of the study of Grobys (2025). Scientific replication is particularly important in empirical finance (e.g., Hou et al., 2020), where evidence is primarily observational and reliability is best assessed against specifications that are similar but not identical. The objective here is a robustness check of our main findings; a broader examination of emerging markets or cryptocurrencies is left for future research.

4.2.3. Dependence-aware resampling via the stationary (geometric) block bootstrap

To assess robustness to serial dependence and volatility clustering in returns, we complement the i.i.d. bootstrap with the stationary (geometric) block bootstrap applied to log-returns. The procedure generates resamples of length T by concatenating contiguous blocks whose lengths are geometrically distributed with expected value m. Operationally, at each step $t > 1$ a new block is initiated with probability $p = 1/m$ at a randomly drawn start index (sampling start indices with replacement); otherwise the resample advances to the next observation in the current block, wrapping around at T. This design preserves generic short-range dependence without imposing a specific parametric structure.

We set the expected block length to $m = \lceil T^{1/3} \rceil$, a conventional choice in the block-bootstrap literature, and keep the total length of synthetic samples equal to the original sample size T. For each resampled return path, we compound to a synthetic price series, re-estimate LPPLS under the same parameter bounds and initialization as in the baseline specification, compute residuals, and apply the residual ADF test with lag order selected by BIC. For illustration, we implement this dependence-aware resampling for the monthly S&P 500 series using $B = 1,000$ bootstrap replications.

4.2.4. Start–value robustness: randomizing starting values

Nonlinear least squares estimation of the LPPLS model can exhibit sensitivity to initial conditions because the objective function is non-convex and the nonlinear parameters $(\beta, \omega, \phi, t_c)$ may admit multiple local minima. To ensure that the reported calibration is not an artefact of a particular initialization, we conduct a start–value robustness check on the original sample. The aim is twofold: (i) to verify that the solution we report is representative within the admissible parameter region defined in the present study, and (ii) to document basic convergence diagnostics under a transparent and replicable initialization scheme.

Methodologically, we carry out a multi–start calibration in which only the nonlinear parameters are randomized, while the linear coefficients are set by conditional least squares. Specifically, for each run $k = 1,\ldots,K$ we draw

$$\beta^{(0)} \sim \mathrm{Unif}\left[\underline{\beta}, \overline{\beta}\right],$$

$$\omega^{(0)} \sim \mathrm{Unif}\left[\underline{\omega}, \overline{\omega}\right],$$

$$\phi^{(0)} \sim \mathrm{Unif}[-\pi, \pi],$$

$$t_c^{(0)} \sim \mathrm{Unif}(t_{\max} + 1, t_{\max} + T],$$

using the same bounds $(\underline{\beta}, \overline{\beta})$ and $(\underline{\omega}, \overline{\omega})$ as in the main analysis (Section 4.1.1), where $t_{\max}$ is the last in–sample time point and T the sample length (so that $\tau = t_c - t > 0$ for all observations). Conditional on the draw $\left(\beta^{(0)}, \omega^{(0)}, \phi^{(0)}, t_c^{(0)}\right)$, we construct the LPPLS regressors and obtain the linear coefficients—denoted $A^{(0)}, B^{(0)}, C^{(0)}$—by ordinary least squares. The resulting vector

$$\theta^{(0)} = \left(A^{(0)}, B^{(0)}, \beta^{(0)}, C^{(0)}, \omega^{(0)}, t_c^{(0)}, \phi^{(0)}\right)$$

is then passed to the constrained nonlinear least squares estimation procedure under the same parameter bounds defined in section 4.1.1. This "nonlinear–randomized, linear–OLS" initialization concentrates randomization where it is empirically consequential (the nonlinear block) while exploiting separability of the model to set the linear block efficiently. We repeat this procedure for K independently generated starts. For each run we examine the optimizer's exit flag, objective value (sum of squared residuals), and the estimated parameter vector. Convergence is defined ex ante by a positive solver exit flag together with finite, in–bounds estimates.

4.2.5. Profiling the objective in the critical time t_c

LPPLS model estimation is known to exhibit weak curvature in the

critical time t_c, which can translate into wide dispersion of $\widehat{t}_c$ across admissible initializations without materially changing fit quality. To examine this directly—and to document identification in t_c without altering our study's inferential focus—we construct a profile of the least-squares criterion in t_c. The idea is standard: for each fixed value of t_c, we re-estimate all remaining parameters to (locally) minimize the sum of squared residuals (SSE), and we then view the minimized SSE as a function of t_c. A broad, shallow trough in this profile indicates weak identification of t_c.

Let $t_{\max}$ denote the last in-sample time point and T the sample length. We evaluate the profile on a uniform grid of 200 points over the admissible interval

$$t_c \in (t_{\max}+1, t_{\max}+T],$$

which ensures $\tau = t_c - t > 0$ for all observations. For each grid value $t_c = g_i$, we fix t_c by setting its lower and upper bounds equal to g_i and then constrained-optimize the remaining parameters under the same bounds and numerical tolerances defined in Section 4.1.1. The starting vector for the first grid point is the unconstrained optimum $\widehat{\theta}$ from the original monthly S&P 500 data sample; for subsequent points we retain the same starting vector. We record, for each grid point, the solver exit flag, the estimated parameter vector, and the minimized SSE. A grid evaluation is classified as successful if the solver returns EF > 0 (success flag), the estimates are finite and within bounds, and SSE is finite. Otherwise it is treated as a failure and omitted from the profile. This treatment isolates the geometry of the criterion from occasional local non-convergence when t_c is exogenously fixed. To further aid interpretation of the profile, we plot a 0.5 % tolerance line as a horizontal reference at $1.005 \times \min_u \mathrm{SSE}(u)$, where $\min_u \mathrm{SSE}(u)$ is the minimum SSE attained over the t_c grid. Formally, if $\mathrm{SSE}_{\min} = \min_u \mathrm{SSE}(u)$, the threshold is $\mathrm{SSE}_{\mathrm{thr}} = 1.005\,\mathrm{SSE}_{\min}$. Grid points with $\mathrm{SSE}(t_c) \le \mathrm{SSE}_{\mathrm{thr}}$ deliver fits that are within 0.5 % of the best attainable on the grid and can be regarded as practically indistinguishable in fit quality.

5. Results

5.1. Main results

5.1.1. Results from original data sets on the S&P 500

First, we fit the original monthly S&P 500 log-price series (January 1871–November 2022) to the LPPLS specification in Eq. (2), using the constraints and starting values in Section 4.1.1. The estimates in Table 1 indicate a power-law exponent $\widehat{\beta} = 0.4884$, consistent with moderate super-exponential acceleration as the critical time is approached. The critical time is estimated at $\widehat{t}_c = 1859.6166$, i.e., about 36.62 months beyond the sample endpoint $T = 1823$. Next, we fit the daily S&P 500 log-price series (January 2, 1980–December 31, 1986) to the same LPPLS model with identical constraints and initial values. As reported in Table 1, the estimated exponent $\widehat{\beta} = 0.2195$ implies a pronounced super-exponential acceleration, with $\widehat{t}_c = 1879.4155$, approximately 109.42 days beyond the daily sample endpoint $T = 1770$. In both frequencies, $\widehat{\beta} \in (0,1)$ and $\widehat{t}_c > T$, delivering qualitatively similar LPPLS signatures; we therefore turn to residual diagnostics in terms of unit-root testing to assess their statistical credibility.

Whereas Figs. 3 and 4 show the evolutions of the log-prices for the original data as well as the LPPLS model, Figs. 5 and 6 plot the estimated model residuals over the sample. Visual inspection of the estimated model residuals suggests stationarity. Using the residuals as defined in Eq. (3), we carry out ADF tests for the original data sets in line with Eq. (4). Consistent with Grobys (2023&2025), we choose the lag-order in line with the Schwarz-Criterion. From Table 2 we observe that the residuals from both data sets are stationary even on a 1 % level. Since Lin et al. (2014) assess LPPLS consistency by applying ADF tests to residuals at conventional significance levels (e.g., 5 %, 1 %, 0.1 %), our 1 % rejections—under tabulated critical values—would likewise be classified as statistically significant LPPLS signatures.

5.1.2. Results from synthetic monthly S&P 500 data

We begin by calibrating the LPPLS model in Eqs. (1–2) to the synthetic monthly datasets. Table 3 reports descriptive statistics for the estimated parameters. As expected under hard bounds, several estimates lie near the admissible limits. In particular, the power-law exponent $\widehat{\beta}$ spans most of the constrained interval: recall that smaller β (closer to 0) implies more pronounced super-exponential acceleration as $t \to t_c$, whereas larger β (closer to 1) approaches exponential growth and indicates weaker acceleration. The median estimated critical time is $\widehat{t}_c = 2053.9337$, about 19.24 years beyond the sample endpoint, and

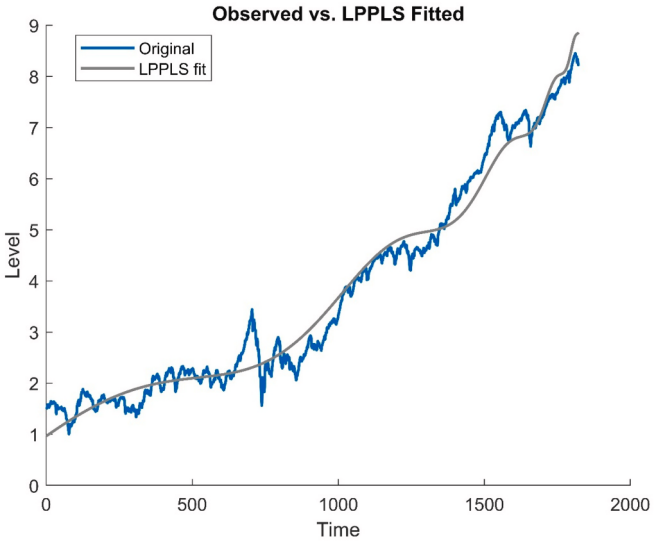

Fig. 3. Calibration of the LPPLS model on monthly S&P 500 data (January 1871–November 2022).
The figure plots the observed log S&P 500 prices (blue solid line) alongside the fitted Log-Periodic Power Law Singularity (LPPLS) curve (grey line). The calibration applies the constrained parameter set described in Section 4.1.1. The estimated trajectory exhibits moderate super-exponential acceleration toward the critical time, consistent with the model's expected behavior under bounded parameters. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Calibration of the LPPLS model on original S&P 500 datasets.

Data set	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$
Monthly	10.3330	−0.2454	0.4884	−0.0534	7.2116	1859.6166	−3.1416
Daily	6.6236	−0.3575	0.2195	0.0292	5.6505	1879.4155	−3.1416

This table reports estimated parameters of the LPPLS model applied to the original monthly (1871–2022) and daily (1980–1986) log-price series. Parameters are defined in Eq. (2); estimation follows bounded nonlinear least squares with standard parameter constraints. The critical time ($\widehat{t_c}$) denotes the inferred termination point of super-exponential growth, measured relative to the sample endpoint.

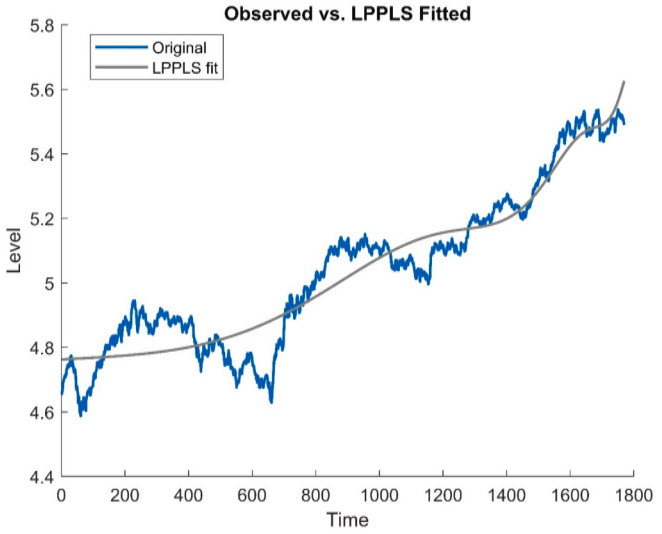

Fig. 4. Calibration of the LPPLS model on daily S&P 500 data (January 2, 1980–December 31, 1986).
This figure compares daily log-prices of the S&P 500 with the fitted LPPLS curve under the standard parameter bounds. The estimation captures stronger acceleration toward the inferred critical point relative to the monthly sample, suggesting that short-horizon data amplify the LPPLS curvature.

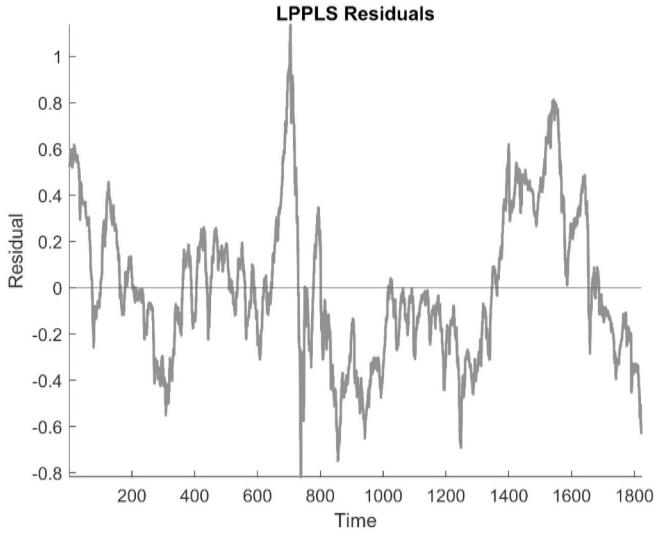

Fig. 5. Residuals of the LPPLS model for monthly S&P 500 data (1871–2022).
The figure shows the residuals from the LPPLS fit in Fig. 3. Visual inspection suggests approximate stationarity, a property formally assessed via Augmented Dickey–Fuller tests reported in Table 2. These residuals form the basis for evaluating the statistical significance of LPPLS "signatures."

substantially later than the $\widehat{t}_c$ obtained from the original data. Moreover, the earliest critical time estimate is essentially at the lower bound—the month when the sample ends ($\widehat{t}_c = T = 1823$)—while the latest sits at the upper bound ($\widehat{t}_c = 2T = 3646$). Thus, the synthetic-data estimates effectively span the imposed window $T < t_c < 2T$, underscoring the role of the constraints and the limited identification of t_c in this setting.

Table 4 reports the correlation matrix of the LPPLS parameters estimated from the synthetic monthly data. Nineteen of the twenty-one pairwise correlations are statistically significant at the 5 % level. For example, the correlation between $\widehat{A}$ and $\widehat{t}_c$ is 0.54 with a t-statistic of 20.13, indicating significance at any conventional level; a larger estimated critical time naturally coincides with a higher fitted terminal log-

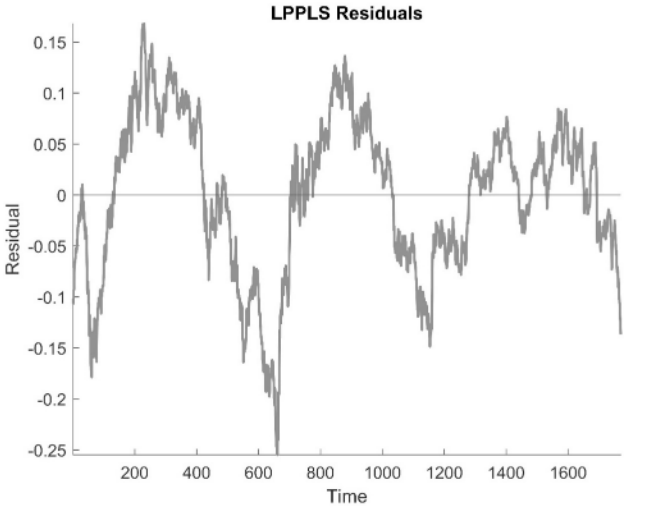

Fig. 6. Residuals of the LPPLS model for daily S&P 500 data (1980–1986).
This figure presents the residuals from the LPPLS calibration in Fig. 4. The residual series appears mean-reverting, consistent with the stationarity results in Table 2, though later bootstrap analyses demonstrate that conventional ADF thresholds overstate significance.

Table 2
Augmented Dickey–Fuller (ADF) tests of LPPLS residuals from original datasets.

Monthly data		Daily data	
ADF test statistic	Lags	ADF test statistic	Lags
−3.2277***	2	−2.6194***	1

*** Statistically significant on a 1 % level.
ADF statistics test the null of a unit root in residuals from the LPPLS fits in Table 1. Critical values correspond to MacKinnon one-sided thresholds (10 %, 5 %, 1 %). The critical values for statistical significance at the 10 %, 5 %, and 1 % levels are −1.62, −1.94, and − 2.57, respectively. Lag order is selected by Schwarz criterion. Significant rejections under tabulated values imply apparent LPPLS "signatures," though later bootstrap analysis reveals size distortion.

level, yielding a positive association between the level parameter and $\widehat{t}_c$. By contrast, the phase $\widehat{\phi}$ and the angular log-frequency $\widehat{\omega}$are negatively correlated (-0.45, $t = -15.81$), consistent with the view that phase shifts and frequency adjustments can partially substitute in fitting oscillatory structure. While many remaining correlations are modest in magnitude, relatively strong negative associations—such as $(\widehat{\phi}, \widehat{\omega})$, $(\widehat{A}, \widehat{B})$, and $(\widehat{A}, \widehat{\beta})$—point to parameter trade-offs that merit further investigation, particularly in relation to identification strength and the role of hard bounds.

How often do "significant" LPPLS signatures arise by chance? If the 1 % level used in practice (e.g., Lin et al., 2014) were appropriate for residual ADF testing, then under synthetic monthly log-price paths the empirical rejection rate should be about 1 %—i.e., only ~1 % of left-tail ADF statistics should fall below the tabulated 1 % critical value (≈ −2.57 for the model specification used in the ADF test regression). Table 5 summarizes the distribution of ADF statistics from 1000 LPPLS calibrations on synthetic monthly data. The median is $\widehat{\lambda}_{\mathrm{ADF}} = -3.4158$; since the tabulated 1 % cutoff is −2.57, at least half of the samples would be labeled "LPPLS-consistent" under conventional thresholds. In fact, 848/1000 statistics satisfy $\widehat{\lambda}_{\mathrm{ADF}} < -2.57$, implying a false-positive rate of 84.8 % at the 1 % level. Consistently, the bootstrap left tail critical values at the 10 %/5 %/1 % significance levels are −4.3753/−4.6774/−5.2444, all much more negative than the tabulated values. At the 5 % level, 969/1000 statistics lie below the tabulated cutoff (≈ −1.94), implying a 96.9 % false-positive rate under table thresholds. Applying

Table 3
Calibration of the LPPLS model on synthetic monthly data.

Statistic	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t}_c$	$\widehat{\phi}$
Minimum	3.4434	−10.0000	0.1000	−0.9990	5.0000	1823.0000	−3.1416
1 % Qnt.	5.2267	−10.0000	0.1055	−0.2990	5.0000	1823.0000	−3.1416
2.5 % Qnt.	5.6948	−10.0000	0.1332	−0.2219	5.0000	1823.0000	−3.1416
5 % Qnt.	6.4211	−10.0000	0.1527	−0.1609	5.0000	1823.0000	−3.1416
10 % Qnt.	7.2450	−5.9002	0.1945	−0.1049	5.0000	1823.0020	−3.1416
Median	10.8462	−0.0180	0.8070	0.0516	5.8752	2053.9337	−1.2299
90 % Qnt.	27.3032	−0.0063	0.9000	0.1534	6.6062	3223.0592	3.1416
95 % Qnt.	37.1718	−0.0053	0.9000	0.2060	7.0268	3646.0000	3.1416
97.5 % Qnt.	41.4985	−0.0044	0.9000	0.2547	7.5551	3646.0000	3.1416
99 % Qnt.	48.3577	−0.0032	0.9000	0.3527	8.3848	3646.0000	3.1416
Maximum	53.8685	−0.0016	0.9000	0.7852	15.0000	3646.0000	3.1416
Mean	13.9631	−1.1505	0.6955	0.0420	5.8897	2276.5283	−0.2223
Standard Deviation	9.2136	2.9747	0.2528	0.1255	0.7787	536.2921	2.9485
Excess Kurtosis	4.3123	4.5739	−0.1012	12.6518	22.6485	1.1311	−1.8925
Skewness	2.1609	−2.5225	−1.0894	−0.9446	2.8434	1.4725	0.1252
T	1000	1000	1000	1000	1000	1000	1000

Table 3 summarizes empirical distributions of LPPLS parameters estimated from 1000 bootstrapped monthly datasets. We report extreme values (minimum/maximum), selected quantiles (1 %, 2.5 %, 5 %, 10 %, 50 %, 90 %, 95 %, 97.5 %, 99 %), the mean and sample standard deviation, skewness, excess kurtosis (kurtosis relative to the normal distribution), and the sample size T. Quantiles are computed from the empirical distribution.

Table 4
Correlation matrix of estimated LPPLS parameters derived from monthly synthetic data.

Panel A. Estimated correlation matrix of LPPLS parameters.							
	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t}_c$	$\widehat{\phi}$
$\widehat{A}$	1.00	−0.87***	−0.69***	−0.06**	−0.31***	0.54***	0.25***
$\widehat{B}$		1.00	0.79***	0.06*	0.23***	−0.29***	−0.24***
$\widehat{\beta}$			1.00	0.06**	0.12***	−0.17***	−0.21***
$\widehat{C}$				1.00	−0.09***	0.04	0.08***
$\widehat{\omega}$					1.00	−0.38***	−0.45***
$\widehat{t}_c$						1.00	0.19***
$\widehat{\phi}$							1.00
Panel B. Estimated t-statistics.							
$\widehat{A}$	–	−55.40	−30.29	−2.03	−10.48	20.13	8.25
$\widehat{B}$		–	40.71	1.86	7.32	−9.63	−7.96
$\widehat{\beta}$			–	1.96	3.93	−5.54	−6.86
$\widehat{C}$				–	−2.74	1.40	2.63
$\widehat{\omega}$					–	−13.06	−15.81
$\widehat{t}_c$						–	6.07
$\widehat{\phi}$							–

Panel A of this table presents the correlation matrix of estimated LPPLS parameters derived from monthly synthetic data. For each correlation r_{ij} in Panel B, we report the corresponding t-statistic for testing $H_0 : \rho_{ij} = 0$ against $H_1 : \rho_{ij} \neq 0$, based on the usual small-sample approximation under joint normality:

$$t_{ij} = r_{ij}\sqrt{\frac{N_{ij}-2}{1-r_{ij}^2}}, \text{with degrees of freedom df} = N_{ij} - 2,$$

where N_{ij} is the number of paired observations for series $i$and j. Two-sided p-values can be obtained from the Student-t distribution with $N_{ij} - 2$ degrees of freedom.

the size-correct bootstrap critical values to the original monthly residual statistic, $\widehat{\lambda}_{\mathrm{ADF}} = -3.2277$, we do not reject even at 10 % (bootstrap $p = 0.5960$). These results document a large gap between empirical ADF quantiles relevant for LPPLS residuals and tabulated thresholds, with table values inducing substantial size distortion.

5.1.3. Results from synthetic daily S&P 500 data

Table 6 reports descriptive statistics for the LPPLS parameters estimated from the synthetic daily data. As with the monthly simulations, the power-law exponent spans the admissible range, $0.1 \leq \widehat{\beta} \leq 0.9$. The median critical time is $\widehat{t}_c = 1958.4728$, which is 79.0573 trading days later than the $\widehat{t}_c$ obtained from the original daily sample. Consistent with the imposed bounds, the parameters $\widehat{\beta}$, $\widehat{C}$, $\widehat{\omega}$, $\widehat{\phi}$, and $\widehat{t}_c$ collectively cover the full constrained ranges. In particular, the earliest $\widehat{t}_c$coincides with the sample endpoint ($\widehat{t}_c = T = 1770$), while the latest sits at the upper boundary ($\widehat{t}_c = 2T = 3540$), indicating that the synthetic-data estimates effectively span the identification window $T < t_c < 2T$.

A key distinction between the two samples is informational: for the daily pre-crash window, the subsequent collapse on 19 October 1987 is an ex-post fact, whereas for the monthly long-horizon series no comparable terminal event is observed within the sample. The question, therefore, is whether the LPPLS—when paired with residual ADF tests—distinguishes fact from fiction in resampled daily data. Table 7 reports ADF statistics from 1000 LPPLS calibrations on synthetic daily series. Mirroring the monthly results, the median is $\widehat{\lambda}_{\mathrm{ADF}} = -3.3332$, well below the tabulated 1 % left-tail cutoff (≈ -2.57), and 828/1000 statistics satisfy $\widehat{\lambda}_{\mathrm{ADF}} < -2.57$, implying an 82.8 % false-positive rate at the nominal 1 % level under table values. The bootstrap left-tail critical values for the daily sample are −4.3808/−4.6898/−5.1854 at the 10

Table 5
Distribution of ADF statistics for residuals from synthetic monthly data.

Statistic	ADF statistic	Lag Length
Minimum	−6.0201	0
1 % Qnt.	−5.2444	0
2.5 % Qnt.	−4.9290	0
5 % Qnt.	−4.6774	0
10 % Qnt.	−4.3753	0
Median	−3.4158	0
90 % Qnt.	−2.4143	0
95 % Qnt.	−2.1398	0
97.5 % Qnt.	−1.8772	1
99 % Qnt.	−1.4904	4
Maximum	−0.1412	13
Mean	−3.4052	0.1090
Standard Deviation	0.7880	0.7885
Excess Kurtosis	0.3515	129.1491
Skewness	0.0360	10.5582
T	1000	1000

This table presents the descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic monthly data. The critical values for 10 %, 5 %, and 1 % statistical significance levels for the standard ADF test are −1.62, −1.94, and − 2.57. The lag-order is chosen in line with the Schwarz-Criterion. The large fraction of statistics below tabulated critical values indicates inflated false-positive rates when conventional thresholds are used.

%/5 %/1 % levels, respectively—very close to those from the synthetic monthly data (−4.3753/−4.6774/−5.2444), underscoring that the empirical quantiles relevant for LPPLS residuals are much more negative than the tabulated thresholds. At the 5 % level, 959/1000 statistics lie below the tabulated cutoff (≈ −1.94), yielding a 95.9 % table-based false-positive rate. Finally, applying the size-correct bootstrap critical values to the original daily residual statistic ($\widehat{\lambda}_{\mathrm{ADF}} = -2.6194$), we do not reject even at 10 % (bootstrap $p = 0.8130$). Taken together, these daily results reinforce the monthly evidence: there is a large gap between the empirical ADF quantiles relevant for LPPLS residuals and tabulated values, and reliance on table thresholds induces substantial size distortion.

5.2. Additional results

5.2.1. Results from benchmarking with the BADF test

The preceding sections established that ADF tests applied to LPPLS residuals are subject to size distortion when evaluated against conventional (tabulated) critical values, and that bootstrap critical values aligned with the two-stage procedure restore size. To place those findings in context, we now report a benchmark based on log prices: a full-sample right-tailed ADF (BADF) implemented on the same monthly and daily S&P 500 samples (with comparable sample lengths). This benchmark does not search over windows and therefore mirrors our single-window LPPLS decision rule, while targeting a different alternative (mild explosiveness in levels rather than residual stationarity). For completeness and comparability, we provide both sets of thresholds for BADF—Monte-Carlo "tabulated" critical values matched to T and specification, and bootstrap critical values obtained by re-estimating the BADF on i.i.d. resamples. The goal is not episode dating, but to assess whether a standard, widely used ADF-type test on levels flags mild explosiveness in the same samples and under the same finite-sample choices.

Using $T = 2,000$ and $B = 10,000$ in the Monte-Carlo design, the right-tailed ADF critical values—i.e., the 90th, 95th, and 99th percentiles of the null distribution—are − 0.4150, − 0.0755, and 0.5909, respectively. Applying the full-sample ADF (BADF) to the log S&P 500 price series (model specification including constant terms, BIC lag selection) yields test statistics of 1.1775 (monthly) and − 0.5560 (daily). That is, using the Monte-Carlo "tabulated" right-tail cutoffs, we reject

Table 7
Descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data.

Statistic	ADF Statistic	Lag Length
Minimum	−6.0721	0
1 % Qnt.	−5.1854	0
2.5 % Qnt.	−4.9596	0
5 % Qnt.	−4.6898	0
10 % Qnt.	−4.3808	0
Median	−3.3332	0
90 % Qnt.	−2.2983	0
95 % Qnt.	−2.0368	0
97.5 % Qnt.	−1.6306	1
99 % Qnt.	−1.2951	1
Maximum	−0.5915	4
Mean	−3.3261	0.0360
Standard Deviation	0.8224	0.2622
Excess Kurtosis	0.1678	128.7633
Skewness	0.0783	10.2625
T	1000	1000

This table presents the descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data. The critical values for 10 %, 5 %, and 1 % statistical significance levels for the standard ADF test are −1.62, −1.94, and − 2.57. The lag-order is chosen in line with the Schwarz-Criterion. The lag-order is chosen in line with the Schwarz-Criterion.

Table 6
Calibration of the LPPLS model on synthetic daily data.

Statistic	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$
Minimum	2.3015	−6.7295	0.1000	−0.9990	5.0000	1770.0000	−3.1416
1 % Qnt.	4.4415	−5.0184	0.1000	−0.9990	5.0000	1770.0000	−3.1416
2.5 % Qnt.	4.6750	−4.0701	0.1000	−0.6280	5.0000	1770.0000	−3.1416
5 % Qnt.	4.9269	−2.8071	0.1000	−0.2940	5.0000	1770.0000	−3.1416
10 % Qnt.	5.1661	−1.5280	0.1000	−0.1518	5.0000	1770.0026	−3.1416
Median	5.9403	−0.0057	0.6647	0.0541	6.0280	1958.4728	−1.7520
90 % Qnt.	8.7485	−0.0006	0.9000	0.3484	7.5760	2761.7118	3.1416
95 % Qnt.	11.0187	−0.0004	0.9000	0.6301	8.3212	3495.5588	3.1416
97.5 % Qnt.	13.9107	0.0002	0.9000	0.9990	9.4924	3540.0000	3.1416
99 % Qnt.	15.9048	0.0583	0.9000	0.9990	11.7616	3540.0000	3.1416
Maximum	20.1571	1.2081	0.9000	0.9990	15.0000	3540.0000	3.1416
Mean	6.5924	−0.4068	0.5915	0.0736	6.2174	2147.6659	−0.3259
Standard Deviation	2.1519	1.0428	0.3085	0.3054	1.2607	470.5718	2.9139
Excess Kurtosis	10.1349	11.0132	−1.3007	4.2051	12.1262	2.1382	−1.8510
Skewness	2.9543	−3.2527	−0.4848	0.0701	2.7723	1.6771	0.1898
T	1000	1000	1000	1000	1000	1000	1000

Table 6 summarizes empirical distributions of LPPLS parameters estimated from 1000 bootstrapped daily datasets. We report extreme values (minimum/maximum), selected quantiles (1 %, 2.5 %, 5 %, 10 %, 50 %, 90 %, 95 %, 97.5 %, 99 %), the mean and sample standard deviation, skewness, excess kurtosis (kurtosis relative to the normal distribution), and the sample size T. Quantiles are computed from the empirical distribution.

$H_0 : \rho = 0$ for monthly data by a substantial margin, including at the 1 % level, whereas for daily data, we fail to reject $H_0 : \rho = 0$ even on a 10 % level (see Table 8). These findings indicate that the BADF test detects mild explosiveness only for the monthly sample.

Table 9 reports descriptive statistics for the BADF test statistics from $B = 1,000$ i.i.d. bootstrap replications based on the monthly S&P 500 log-price series used in the main analysis. That is, in each replication, we resample returns i.i.d., reconstruct the price path, take log prices, and re-estimate the BADF statistic under the bootstrap null using the same specification (intercept included; BIC lag selection). The bootstrap distribution is centered to the left of the tabulated 10 % right-tail cutoff: its median $\widehat{Q}_{0.50} = -0.4787$ is below $\lambda_{\mathrm{ADF},0.90} = -0.4150$. The bootstrap right-tailed critical values are 0.7566, 1.0305, and 1.7350 at the 10 %, 5 %, and 1 % levels, respectively. At the nominal 1 % level (critical value 0.5909), the empirical rejection rate under the bootstrap null is 13 %, i.e., the test is oversized. When evaluated against the bootstrap critical values (Table 9), the BADF statistic for the original monthly sample, $\widehat{\lambda}_{\mathrm{ADF}} = 1.1775$, exceeds the 5 % cutoff (1.0305) but not the 1 % cutoff (1.7350); thus, we reject $H_0 : \rho = 0$ at the 5 % level but not at 1 % (the bootstrap p-value is 0.037).

Table 10 reports descriptive statistics for the BADF test statistics from $B = 1,000$ i.i.d. bootstrap replications based on the daily S&P 500 data. In each replication, we resample returns i.i.d., reconstruct the price path, take log prices, and re-estimate the BADF statistic under the bootstrap null using the same specification (intercept included; BIC lag selection). As in the monthly case, the bootstrap distribution is shifted left relative to the tabulated cutoff: its median, $\widehat{Q}_{0.50} = -0.8711$, lies below the Monte-Carlo 'tabulated' 10 % right-tail critical value, $\lambda_{\mathrm{ADF},0.90} = -0.4150$. The bootstrap right-tailed critical values are 0.4935, 0.8216, and 1.5796 at the 10 %, 5 %, and 1 % levels, respectively. Consistent with the monthly results, at the nominal 1 % level (tabulated critical value 0.5909), the empirical rejection rate under the bootstrap null is 8.7 %, i.e., the test is oversized. By contrast, when evaluated against the bootstrap critical values (Table 13), the BADF statistic from the daily sample, $\widehat{\lambda}_{\mathrm{ADF}} = -0.5560$, does not exceed the 5 % cutoff (0.8216), so we fail to reject $H_0 : \rho = 0$ at the 5 % level (p-value 0.3920).

Knowing that stock market crash as of October 1987 arrived only a few months in the ex-post sample of our daily data sample, a reader might wonder how come that the BADF does not assert a bubble formation in the period preceding the crash. Two possible aspects of the testing framework could explain why the full-sample BADF statistic on the daily pre-crash window does not exceed the simulated critical values:

(i) Tested alternative versus LPPLS dynamics. The BADF is a linear unit-root test targeting mild explosiveness in the sense of $\rho > 0$ in $\Delta y_t = \alpha + \rho y_{t-1} + \sum_{i=1}^{p} \phi_i \Delta y_{t-i} + \varepsilon_t$. By construction, it does not target the log-periodic (oscillatory) structure of the LPPLS specification. A time series can exhibit LPPLS-type curvature and accelerating oscillations without generating a sufficiently large positive ADF coefficient to trigger a right-tailed rejection. Hence, failure to reject with BADF does not contradict evidence consistent with LPPLS; it reflects that the two procedures test different alternatives.

(ii) Size–power trade-off with an intercept. For the benchmark on log prices we include an intercept to absorb deterministic drift, which may improve finite-sample size but generally reduces power to detect short explosive episodes unless they dominate the sample. At daily frequency, the explosive run-up preceding October 1987

Table 8
Right-tailed Augmented Dickey-Fuller tests on the original data sets using a constant in the test regression.

Monthly data		Daily data	
ADF test statistic	Lags	ADF test statistic	Lags
1.1775***	2	−0.5560	0

*** Statistically significant on a 1 % level.
This table presents the results of Augmented Dickey-Fuller (ADF) tests conducted on the log-prices of the original datasets. The model used for test regression accounts for a constant term. The critical values for the right-tailed test derived from Monte Carlo for statistical significance at the 10 %, 5 %, and 1 % levels are −0.4150, −0.0755, and 0.5909, respectively. The lag order is selected according to the Schwarz criterion. The monthly series exhibits mild explosiveness under Monte Carlo derived inference, whereas the daily series does not.

Table 9
Descriptive statistics of BADF (ADF with intercept) test statistics for synthetic monthly log-price series of the S&P 500.

Statistic	ADF Statistic	Lag Length
Minimum	−4.1310	0
1 % Qnt.	−2.8071	0
2.5 % Qnt.	−2.4146	0
5 % Qnt.	−2.1145	0
10 % Qnt.	−1.7631	0
Median	−0.4787	0
90 % Qnt.	0.7566	0
95 % Qnt.	1.0305	0
97.5 % Qnt.	1.2808	1
99 % Qnt.	1.7350	4
Maximum	2.7851	13
Mean	−0.5066	0.1130
Standard Deviation	0.9795	0.8005
Excess Kurtosis	0.0532	121.7901
Skewness	−0.1027	10.2100
T	1000	1000

The table reports summary statistics of the full-sample right-tailed ADF statistic estimated with a constant term on synthetic monthly log-price paths for the S&P 500. For reference, Monte-Carlo "tabulated" critical values matched to the sample length and specification are −0.4150, −0.0755, and 0.5909 at the 10 %, 5 %, and 1 % significance levels, respectively. In each estimation, the lag order p is selected by the Schwarz (BIC) criterion. At the 5 % significance level, the monthly series exhibits mild explosiveness under size-corrected (bootstrap) inference.

Table 10
Descriptive statistics of BADF (ADF with intercept) test statistics for synthetic daily log-price series of the S&P 500.

Statistic	ADF Statistic	Lag Length
Minimum	−3.6289	0
1 % Qnt.	−3.0044	0
2.5 % Qnt.	−2.6905	0
5 % Qnt.	−2.5222	0
10 % Qnt.	−2.2173	0
Median	−0.8711	0
90 % Qnt.	0.4935	0
95 % Qnt.	0.8216	0
97.5 % Qnt.	1.1271	0
99 % Qnt.	1.5796	1
Maximum	2.3144	4
Mean	−0.8491	0.0320
Standard Deviation	1.0339	0.2550
Excess Kurtosis	−0.3894	144.1863
Skewness	0.0816	10.9887
T	1000	1000

The table reports summary statistics of the full-sample right-tailed ADF statistic estimated with a constant term on synthetic daily log-price paths for the S&P 500. For reference, Monte-Carlo "tabulated" critical values matched to the sample length and specification are −0.4150, −0.0755, and 0.5909 at the 10 %, 5 %, and 1 % significance levels, respectively. In each estimation, the lag order p is selected by the Schwarz (BIC) criterion. The daily series does not exhibit mild explosiveness under size-corrected inference, reinforcing evidence derived from Monte Carlo "tabulated" critical values.

may be relatively short compared with the length of the pre-crash window, and BIC lag selection further attenuates persistence by absorbing short-run dynamics. These features lower the BADF statistic relative to its critical values, even in periods widely regarded as bubbly.

5.2.2. Results from daily data on gold futures

Consistent with Grobys (2025), we obtain daily gold futures data from 2 December 2015 to 6 November 2024 (2147 observations). Tables A.5–A.6 report descriptive statistics for daily returns and for the bootstrapped sample means, and Fig. 7 plots the original price series alongside alternative historical trajectories generated by resampling. As with the S&P 500 simulations, the bootstrap median is very close to the realized series at the terminal date T, as expected under our resampling design. We next fit the original daily log-price series of gold futures to the LPPLS specification in Eq. (2), using the constraints and starting values described in Section 4.1. The estimates in Table 11 yield a power-law exponent $\widehat{\beta} = 0.7081$, consistent with less pronounced super-exponential acceleration (relative to the S&P 500 fits) as the critical time is approached. The estimated critical time is $\widehat{t}_c = 2177.9168$, i.e., about 30.92 trading days beyond the sample endpoint $T = 2147$.

Fig. 8 overlays the fitted LPPLS curve on the log-price data, while Fig. 9 plots the corresponding residuals over the sample. Visual inspection suggests stationarity, which we formally evaluate using the ADF regression in Eq. (4) applied to residuals defined in Eq. (3). The lag order is chosen by the Schwarz (BIC) criterion. As reported in Table 12, the ADF statistic for the original residuals is -4.6215, which—under tabulated left-tailed critical values—would be deemed statistically significant at the 1 % level, consistent with the convention in Lin et al. (2014). Turning to the resampling evidence, Table 13 reports descriptive statistics for LPPLS parameters estimated from synthetic daily gold series. As in the S&P 500 simulations, $\widehat{\beta}$ spans the admissible range $0.1 \le \widehat{\beta} \le 0.9$. The median critical time is $\widehat{t}_c = 2394.2255$, which is 247.23 trading days later than the $\widehat{t}_c$ from the original fit. Consistent with the imposed bounds, $\widehat{\beta}$, $\widehat{C}$, $\widehat{\omega}$, $\widehat{\phi}$, and $\widehat{t}_c$ collectively cover the full constrained ranges. In particular, the earliest $\widehat{t}_c$ coincides with the sample endpoint ($\widehat{t}_c = T = 2147$), while the latest occurs at the upper boundary ($\widehat{t}_c = 2T = 4294$), indicating that the synthetic estimates span the identification window $T < t_c < 2T$.

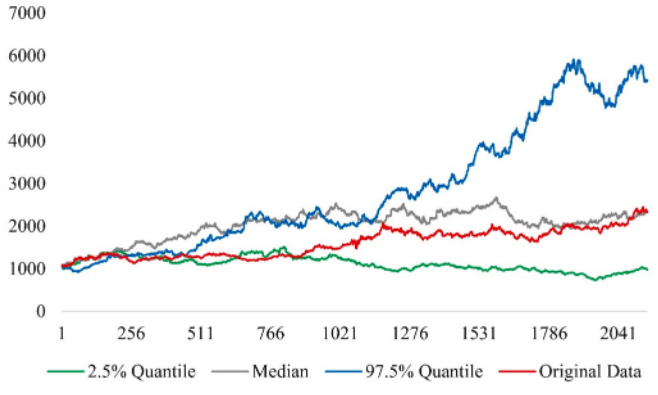

Fig. 7. Alternative historical trajectories for gold futures (December 2, 2015–November 6, 2024).
This figure displays 1000 synthetic trajectories of gold futures prices, generated by compounding daily log-returns resampled with replacement from the original dataset (source: Grobys, 2025; Investing.com). The realized price path (red line) and the median bootstrap trajectory (grey line) align at the sample endpoint, confirming that the bootstrap reproduces unconditional drift while excluding deterministic oscillations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 14 reports ADF statistics from 1000 LPPLS calibrations on synthetic daily gold series. Mirroring the S&P 500 results, the median is $\widehat{\lambda}_{\mathrm{ADF}} = -3.3047$, well below the tabulated 1 % left-tail cutoff (≈ -2.57), and 810/1000 statistics satisfy $\widehat{\lambda}_{\mathrm{ADF}} < -2.57$, implying an 81.0 % false-positive rate at the nominal 1 % level under table values. The bootstrap left-tail critical values for the daily gold sample are -4.3515 / -4.6634 / -5.0516 at the 10 % / 5 % / 1 % levels, respectively—very close to those obtained for the S&P 500 synthetic monthly and daily data—reinforcing that the empirical quantiles relevant for LPPLS residuals are far more negative than tabulated thresholds. At the 5 % level, 947/1000 statistics lie below the tabulated cutoff (≈ -1.94), yielding a 94.7 % table-based false-positive rate.

Finally, when we evaluate the original daily residual statistic against the bootstrap critical values, $\widehat{\lambda}_{\mathrm{ADF}} = -4.6215$ produces a bootstrap p-value of 0.0580. Thus, we reject at the 10 % level but not at 5 %. Interpreted cautiously, this is marginal evidence—at the 10 % level—of an LPPLS-consistent episode in gold, consistent with the narrative in Grobys (2025). Taken together, these results corroborate the evidence from the S&P 500: there is a pronounced gap between empirical ADF quantiles for LPPLS residuals and the tabulated values, and reliance on table thresholds induces substantial size distortion.

5.2.3. Dependence-aware resampling: results and comparison with i.i.d. bootstrap (monthly S&P 500)

Appendix Tables A.7–A.8 report, respectively, the empirical distributions of the LPPLS parameter estimates and the residual ADF statistics obtained under the stationary (geometric) block bootstrap applied to monthly S&P 500 returns. These results are intended to assess the sensitivity of our inference to serial dependence and volatility clustering. For comparability, the resampling design mirrors the baseline specification except for the block structure (expected block length $m = \lceil T^{1/3} \rceil$); all estimation bounds, initialization, and lag-order selection (BIC) are held fixed.

Relative to the i.i.d. bootstrap results in Tables 3 and 5, the block-bootstrap distributions of the LPPLS parameters exhibit very similar location and overall shape. Any observed differences are of second order and consistent with the introduction of weak dependence—e.g., modest changes in dispersion for some parameters—without altering the qualitative features of the empirical distributions. This indicates that the calibration step is not materially affected by the choice between i.i.d. and dependence-aware resampling at the monthly frequency.

The residual ADF statistics (Appendix Table A.8) likewise remain broadly comparable to those in Table 5. While dependence-aware resampling can admit slight shifts in quantiles and a marginally different distribution of BIC-selected lag orders—consistent with preserving serial correlation—the substantive conclusions are unchanged. In particular, conventional tabulated ADF critical values continue to over-reject in this two-stage setting, whereas the estimation-aligned bootstrap yields empirical rejection rates in line with nominal size. Hence, the central finding—that size distortion arises when applying tabulated ADF cutoffs to residuals from the constrained nonlinear first stage—is robust to the adoption of a block-bootstrap design on monthly data.

Overall, the dependence-aware analysis corroborates the baseline i.i.d. evidence. Allowing for generic short-range dependence through stationary blocks does not modify the interpretation of our results, but it strengthens their credibility by demonstrating that the conclusions do not hinge on the i.i.d. resampling assumption.

5.2.4. Multi-start robustness on the original sample (monthly S&P 500)

This subsection documents the start-value robustness check—for illustration—on the original monthly S&P 500 sample, using $K = 500$ independently generated initializations for the nonlinear block $(\beta, \omega, \phi, t_c)$, with the linear coefficients (A, B, C) set by conditional OLS before constrained NLS re-estimation. Table A.9 reports distributional

Table 11
Calibration of the LPPLS model on original data set for Gold futures.

Dataset	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$
Daily	7.7315	−0.0026	0.7081	0.2718	5.0000	2177.9168	0.9473

This table reports estimated parameters of the LPPLS model applied to daily gold futures (2015—2024) log-price series. Parameters are defined in Eq. (2); estimation follows bounded nonlinear least squares with standard parameter constraints. The critical time ($\widehat{t_c}$) denotes the inferred termination point of super-exponential growth, measured relative to the sample endpoint.

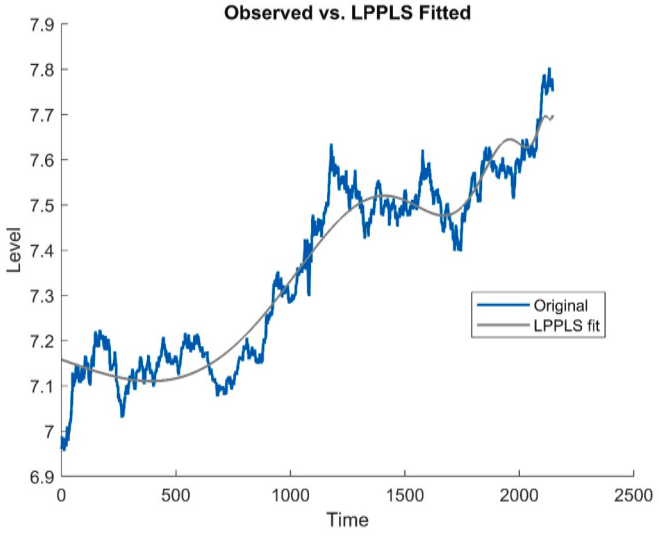

Fig. 8. Calibration of the LPPLS model on daily gold futures (2015–2024). This figure plots observed daily gold futures log-prices together with the fitted LPPLS curve. The estimated model yields a critical time approximately 30 days beyond the sample end, indicating moderate super-exponential growth consistent with transient speculative dynamics. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

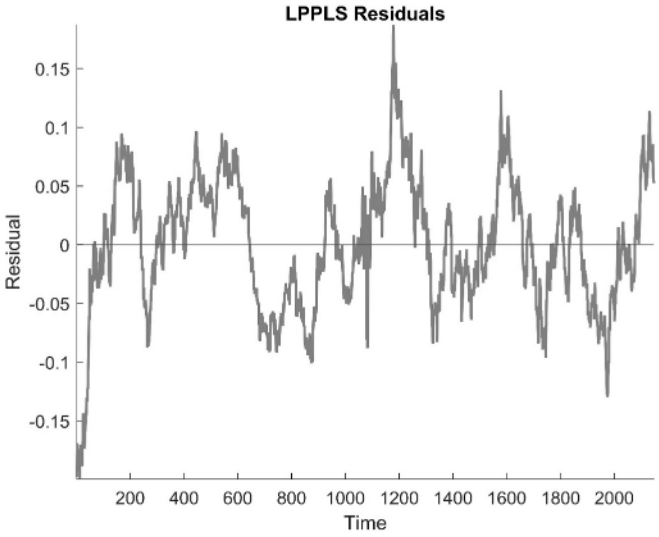

Fig. 9. Residuals of the LPPLS model for daily gold futures (2015–2024). This figure shows residuals from the LPPLS calibration in Fig. 8. The pattern appears weakly mean-reverting; formal ADF results in Table 12 confirm statistical stationarity at the 1 % level under conventional thresholds. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 12
Augmented Dickey-Fuller tests of the residuals from original data set on Gold futures.

ADF test statistic	Lags
−4.6215***	0

*** Statistically significant on a 1 % level.
ADF statistics test the null of a unit root in residuals from the LPPLS fits in Table 11. Critical values correspond to MacKinnon one-sided thresholds (10 %, 5 %, 1 %). The critical values for statistical significance at the 10 %, 5 %, and 1 % levels are −1.62, −1.94, and − 2.57, respectively. Lag order is selected by Schwarz criterion. Significant rejections under tabulated values imply apparent LPPLS "signatures."

summaries for the optimized parameters in the order $[A, B, \beta, C, \omega, t_c, \phi]$ together with the objective value (SSE).

Two features emerge clearly. First, the optimization landscape is flat in economically relevant regions. The best objective value is virtually identical to the median across converged runs—SSE(best)/SSE(median) = 0.999. Consistent with this, the minimum SSE coincides with its 10 % quantile (both 17.742), whereas the median SSE is 17.762. These diagnostics indicate a benign, flat basin around the selected solution rather than a sharp, start-specific optimum; consequently, the reported fit is not an artefact of a particular initialization.

Second, the dispersion of the critical time is wide but well within the admissible window, as expected for LPPLS. The optimized t_c spans 1852.3 to 3646.0 (i.e., from just beyond the sample end to one sample length ahead, $t_{\max} < t_c \leq t_{\max} + T$), with a median of 2967.1. This confirms that the admissible search region is effectively explored and that t_c is weakly pinned down within it, a property already noted in the literature and consistent with the flat objective.

Regarding convergence, the SSE is finite for every run, implying that non-convergence is captured by the solver's exit flag (EF) alone under our rule. In our implementation, 263 of 500 starts returned EF > 0 (converged), while 237 of 500 were classified as failures by the solver's criterion (EF ≤ 0). Thus, Converged/K = 263/500 (~52.6 %): roughly half of the randomized initializations reach an acceptable solution (EF > 0 with finite, in-bounds estimates). The remaining ~47 % either enter infeasible regions, violate admissibility checks, or terminate without a success flag—behavior that is common in nonconvex calibrations with broad start bounds. Importantly, the flatness jdocumented here (SSE(best) ≈ SSE(median)) shows that the multiplicity of starts is not generating materially different solutions of comparable quality; rather, many initializations land in essentially the same trough of the objective. Across the other parameters, the summaries are in line with the imposed bounds and standard calibrations: β concentrates in the upper half of $[0.1, 0.9]$ (median 0.7197), ω covers $[5, 15]$ with a median of 11.306, and ϕ spans $[-\pi, \pi]$ as designed.

5.2.5. *Objective profile in* t_c*: evidence of weak curvature*

To evaulate the profile in t_c, we use the starting vector for the first grid point is the unconstrained optimum $\widehat{\theta}$ from the original sample (see Section 5.1.1):

Table 13
Calibration of the LPPLS model on synthetic daily Gold futures data.

Statistic	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$
Minimum	5.0551	−6.6377	0.1000	−0.9990	5.0000	2147.0000	−3.1416
1 % Qnt.	6.4257	−4.1983	0.1000	−0.9990	5.0000	2147.0000	−3.1416
2.5 % Qnt.	6.6721	−3.1396	0.1000	−0.9990	5.0000	2147.0000	−3.1416
5 % Qnt.	6.9696	−2.0450	0.1000	−0.6215	5.0000	2147.0000	−3.1416
10 % Qnt.	7.2823	−0.9996	0.1000	−0.2944	5.0000	2147.0268	−3.1416
Median	8.1531	−0.0025	0.7179	0.0069	5.9557	2394.2255	−1.6752
90 % Qnt.	10.0117	−0.0004	0.9000	0.4087	7.8196	3205.8131	3.1416
95 % Qnt.	12.1669	0.0001	0.9000	0.8294	9.2329	3833.1623	3.1416
97.5 % Qnt.	14.2763	0.0046	0.9000	0.9990	10.7265	4294.0000	3.1416
99 % Qnt.	16.5522	0.0627	0.9000	0.9990	13.1727	4294.0000	3.1416
Maximum	22.0894	0.6664	0.9000	0.9990	15.0000	4294.0000	3.1416
Mean	8.5730	−0.2745	0.6179	0.0249	6.3030	2570.8822	−0.3887
Standard Deviation	1.7841	0.8060	0.3042	0.3707	1.4991	516.2677	2.8635
Excess Kurtosis	13.6120	17.2837	−1.1853	2.1652	9.2866	3.0398	−1.7961
Skewness	3.1505	−3.9168	−0.5910	0.0480	2.6599	1.8186	0.2443
T	1000	1000	1000	1000	1000	1000	1000

Table 13 summarizes empirical distributions of LPPLS parameters estimated from 1000 bootstrapped daily gold-futures datasets. We report extreme values (minimum/maximum), selected quantiles (1 %, 2.5 %, 5 %, 10 %, 50 %, 90 %, 95 %, 97.5 %, 99 %), the mean and sample standard deviation, skewness, excess kurtosis (kurtosis relative to the normal distribution), and the sample size T. Quantiles are computed from the empirical distribution.

Table 14
Descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data on Gold futures.

Statistic	ADF Statistic	Lag Length
Minimum	−6.2405	0
1 % Qnt.	−5.0516	0
2.5 % Qnt.	−4.8230	0
5 % Qnt.	−4.6634	0
10 % Qnt.	−4.3515	0
Median	−3.3047	0
90 % Qnt.	−2.2518	0
95 % Qnt.	−1.9218	0
97.5 % Qnt.	−1.4935	1
99 % Qnt.	−1.0906	2
Maximum	−0.2698	4
Mean	−3.2864	0.0490
Standard Deviation	0.8407	0.3077
Excess Kurtosis	0.3577	66.2071
Skewness	0.1429	7.6062
T	1000	1000

This table presents the descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data on Gold futures. The critical values for 10 %, 5 %, and 1 % statistical significance levels for the standard ADF test are −1.62, −1.94, and − 2.57. The lag-order is chosen in line with the Schwarz-Criterion. The lag-order is chosen in line with the Schwarz-Criterion. The bootstrap-based inference restores empirical size, confirming that tabulated thresholds overstate significance in LPPLS diagnostics.

$$\theta^{(0)} = \left(A^{(0)}, B^{(0)}, \beta^{(0)}, C_1^{(0)}, \omega^{(0)}, t_c^{(0)}, \phi^{(0)}\right)$$

$$= (10.3330, -0.2454, 0.4884, -0.0534, 7.2116, 1859.6166, -3.1416);$$

Fig. A.1 displays the profile of the least-squares criterion as a function of the critical time t_c, obtained by fixing t_c on a grid and re-estimating all remaining parameters at each grid point. The plotted curve is $t_c \mapsto \mathrm{SSE}(t_c)$, i.e., the minimized criterion conditional on t_c. From Fig. A.1 we observe that the profile exhibits a broad, shallow trough: a substantial fraction of the t_c grid points lie below the 0.5 % tolerance line, indicating that many distinct t_c values yield virtually identical fit quality. This directly evidences weak curvature in t_c and explains the dispersion of $\widehat{t}_c$ observed in the multi-start robustness check (see Section 5.2.4). The vertical line at the unconstrained estimate $\widehat{t}_c$ (≈ 1859.6) falls in a region with slightly higher SSE than the lowest portion of the trough; however, the difference is well within the 0.5 % tolerance, so it is economically negligible. Occasional spikes reflect isolated non-convergence when t_c is fixed and do not alter the conclusion that the criterion is flat over a wide set of t_c values. Taken together with the multi-start robustness check in Section 5.2.4 (which showed $\mathrm{SSE(best)/SSE(median)} \approx 1$ across converged runs), the profile confirms that dispersion in $\widehat{t}_c$ reflects the flatness of the criterion rather than instability of the estimation. Importantly, none of these features affect our main conclusion on estimation-aligned residual-ADF size control: the bootstrap calibration remains robust regardless of the precise t_c chosen within the flat region.

6. Discussion

6.1. Synthesis of results and comparison with earlier research

The results offer three overarching messages. First, point estimates of the critical time $\widehat{t}_c$ differ from values reported in prior studies (e.g., Grobys, 2023; Sornette, 2017). Using the same daily S&P 500 sample as in the present study—while noting that the LPPLS specifications may differ slightly—Grobys (2023) and Sornette (2017) estimate the critical time (finite-time singularity) to be 294 and 158 trading days ahead, respectively. The actual crash occurred 202 trading days after December 31, 1986 (October 19, 1987). By contrast, the present study places the crash time at approximately 109 trading days after the sample endpoint. This is not anomalous; it is consistent with Brée et al.'s (2013) argument that t_c is a "sloppy" parameter—highly sensitive to seemingly minor changes in specification, sample endpoints, and constraints. In our setting, $\widehat{t}_c$ often lies well beyond the sample endpoint T, which suggests that within the admissible window $T < t_c < 2T$ the objective is relatively flat in the t_c direction. That geometry helps explain why different samples and bound sets can deliver materially different $\widehat{t}_c$ without implying instability in the rest of the fit. In other words, disagreement in $\widehat{t}_c$ across studies is expected once one acknowledges its weak curvature and the role of hard bounds.

Second, the bootstrap critical values for the residual ADF tests are strikingly similar across the monthly, daily, and gold applications at the 10 %, 5 %, and 1 % levels. This clustering has a methodological origin. First, we hold the empirical design fixed: the same constrained LPPLS calibration, the same residual ADF specification, and BIC for lag selection. Second, the large sample lengths in all cases mean BIC typically selects low lag orders, yielding comparable finite-sample distributions for the test statistic. Third, the i.i.d. resampling of returns preserves the empirical marginal (including tail thickness); once compounded to levels and residualized, the resulting null features are broadly similar. In short, the near alignment of bootstrap quantiles reflects design

invariance rather than market-specific peculiarities.

Third, when inference is aligned with the estimation procedure (i.e., using bootstrap critical values from the full two-stage pipeline), the LPPLS-residual tests reject far less often than when judged against conventional tables. In our main samples, only gold exhibits marginal significance (10 % level). The daily S&P 500 pre-1987 window does not reject under bootstrap thresholds despite the ex-post crash that follows shortly after T. This pattern is consistent with limited power of a single full-sample unit-root test to detect short, late-sample explosive episodes, especially with an intercept absorbing drift and BIC absorbing short-run dynamics. As a result, failure to reject in that daily window does not contradict the historical "bubble" narrative; it reflects the test's power properties given the design we match to the LPPLS pipeline.

A potential concern is weak identification in the calibrated LPPLS models. Two clarifications are in order. (i) In our re-estimation, the critical time $\widehat{t}_c$ typically lies well beyond T, and the fits are stable under the chosen bounds; this pattern does not indicate pervasive weak identification of the full parameter vector. (ii) We do observe boundary hits for the phase parameter, $\widehat{\phi} \approx -\pi$. This is a standard outcome when the oscillatory amplitude $|\widehat{C}|$ is small: the log-periodic term carries little weight, the objective function is nearly flat in ϕ, and a constrained optimizer tends to park ϕ at the boundary. Importantly, our inference does not rely on ϕ. It rests on residual ADF tests, which summarize whether the residuals are stationary and mean-reverting in the original fits. Hence, boundary behavior in ϕreflects an identification nuance of a weakly weighted component, not a failure of the residual-based diagnostic used to assess statistical significance.

6.2. Implications

The implication is methodological. A common practice calibrates LPPLS under hard bounds and then declares a "signature" when residual ADF tests reject at tabulated cutoffs. Our results show that this practice is oversized in the constrained, two-stage setting: table values do not reflect the finite-sample distribution of an ADF statistic computed on estimated residuals from a nonlinear, bounded first stage. In effect, the calibration can superimpose oscillatory structure that is then "certified" by inappropriate critical values, inflating false positives.

When inference is aligned with estimation—by recomputing critical values via a bootstrap that mirrors the full procedure—many purported "signatures" recede. As a complementary benchmark, the full-sample BADF on log prices is less oversized under our design and therefore produces fewer false positives than the LPPLS-residual ADF. This does not elevate BADF to a universal standard; rather, it shows that benchmarking against a familiar ADF-type test helps contextualize LPPLS findings and separate size issues from claims about detection power.

For the literature, the message is direct: results obtained from LPPLS residual tests against tabulated ADF thresholds warrant re-examination. Studies that enforce bounds and then rely on conventional tables for residual stationarity may reflect size distortion. Reassessing those findings with estimation-aligned critical values (bootstrap or Monte Carlo, as appropriate) is a natural and necessary next step.

6.3. Limitations and avenues for future research

Our approach prioritizes size-corrected inference for the LPPLS residual tests, but it does so at a cost: like any unit-root test, the bootstrap-calibrated ADF inherits the size–power trade-off. In particular, a single full-sample test has limited power for short or late-sample explosive episodes. Exploring window-based procedures (e.g., SADF/GSADF) under an analogous estimation-aligned calibration would be informative, though it introduces multiple-testing adjustments and departs from the single-window logic of the LPPLS application we analyze.

An additional avenue is to refine identification via theory-guided bounds or a Bayesian specification with informative priors. While a Bayesian treatment (e.g., shrinkage on m, ω or priors on t_c) could further formalize parameter uncertainty, such extensions are deferred to future work.

Our cross-market check is limited to gold futures. While this replication is informative, a broader assessment across emerging equity markets, cryptocurrencies, and other commodities would help map where LPPLS residual-ADF inference is most susceptible to size distortion and where levels-based benchmarks such as BADF are most informative. We view this broader mapping as an important agenda for future research.

7. Conclusion

This study revisits the empirical practice of diagnosing bubbles with the LPPLS model by aligning inference with the way statistics are produced. The central message is straightforward. When the LPPLS model is calibrated under parameter bounds and its residuals are judged against conventional ADF test tables, the resulting tests are substantially oversized for the constrained, two-stage setting. Using synthetic series that preserve the roughness and tail behavior of financial returns while excluding log-periodic structure, we show that tabulated critical values can label as "significant" a large share of episodes that are, by construction, devoid of log-periodicity. When critical values are recomputed by re-estimating the full procedure on each resample, the empirical size is restored and many apparent "signatures" recede.

Applied to S&P 500 monthly and daily samples, tabulated critical values yield inflated rejection rates for LPPLS residual-ADF tests; bootstrap critical values calibrated to the exact estimation design overturn these findings. A benchmark based on the full-sample ADF on log prices (with an intercept and BIC lag selection) provides additional context: while that benchmark is not a universal standard, it is less oversized under our design and produces far fewer false positives than LPPLS residual-ADF tests. The cross-market application to gold futures points in the same direction: empirical quantiles relevant for residual tests lie far from tabulated critical values, and size-correct inference changes the inference of apparent regularities.

The results also clarify what should—and should not—be inferred from LPPLS model calibrations. Estimated critical times can diverge across studies without implying instability of the entire fit. In our setting, $\widehat{t}_c$ often lies well beyond the sample endpoint, consistent with the view that $\widehat{t}_c$ is a "sloppy" parameter whose objective surface is relatively flat within admissible bounds. Boundary hits for the phase parameter when the oscillatory amplitude is small are likewise unsurprising and immaterial for our inference, which is conducted on residual stationarity rather than on the phase itself.

Methodologically, the contribution is to replace table look-ups with estimation-aligned inference. The i.i.d. bootstrap on returns avoids embedding a parametric volatility law in the null, preserves the empirical marginal distribution, and removes any imposed log-periodic structure once compounded to levels. Re-estimating the LPPLS model and re-testing on every synthetic data sample then delivers critical values that reflect the finite-sample distribution of the statistic actually used in practice. In this sense, the paper speaks to a broader point in empirical finance: credibility rests on falsification under a null that mirrors the full procedure, not under a null chosen for convenience.

For applied work, two practical implications follow. First, claims of "LPPLS signatures" established by residual ADF tests at tabulated cutoffs warrant re-examination with estimation-aligned critical values (bootstrap or Monte Carlo, as appropriate). Second, reporting both the tabulated critical values and the aligned thresholds—along with empirical sizes from simulations—offers readers a transparent view of how much of an apparent signal survives once size is controlled.

In sum, once inference is calibrated to estimation, many of the striking patterns attributed to log-periodicity lose their statistical force. This is not a rejection of disciplined modeling; it is a reminder that in

financial data—irregular, heavy-tailed, and restless—pattern and chance often rhyme. A disciplined path forward is clear: match the test to the statistic, report size alongside power, and let regularities that survive those checks earn their name.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Declaration of competing interest

The author reports there are no competing interests to declare.

Appendix A. Appendix

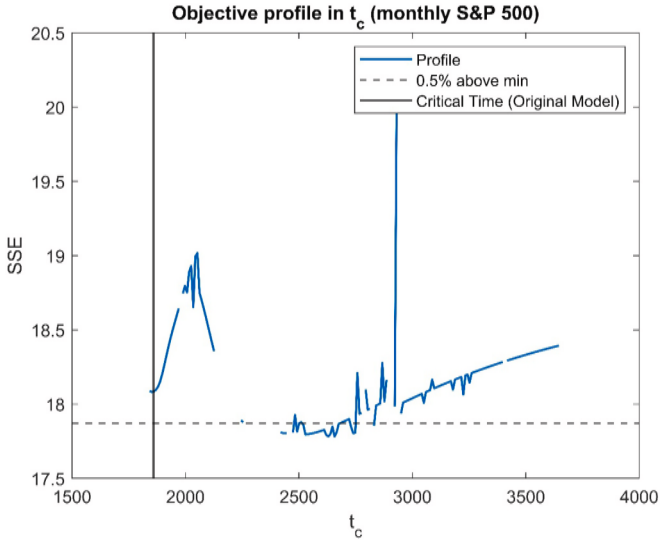

Fig. A.1. Profile of the least-squares criterion in the critical time t_c.
We use the unconstrained optimum as starting values, $\theta^{(0)} = \left(A^{(0)}, B^{(0)}, \beta^{(0)}, C_1^{(0)}, \omega^{(0)}, t_c^{(0)}, \phi^{(0)}\right) = (10.3330, -0.2454, 0.4884, -0.0534, 7.2116, 1859.6166, -3.1416)$; at each grid value, t_c is held fixed while $(A, B, \beta, C_1, \omega, \phi)$ are re-estimated subject to the same bounds as in the main specification. The vertical line marks $\widehat{t}_c$ from the unconstrained fit, and the dashed line indicates a 0.5 % tolerance above the minimum SSE, i.e., the set $\{t_c : \mathrm{SSE}(t_c) \leq 1.005 \times \min_u \mathrm{SSE}(u)\}$. The broad, flat trough around $\widehat{t}_c$ evidences weak curvature in t_c; equivalently, many t_c values fall within this 0.5 % band and deliver indistinguishably good fit quality.

Table A.1
Descriptive statistics for monthly returns on the S&P 500 for the period January 1871 to November 2022.
This table reports the descriptive statistics for monthly data on the S&P 500. The data were downloaded from Robert Shiller's data library (www.econ.yale.edu/~shiller/data.htm). The data sample is from January 1871 to November 2022 corresponding to 1822 observations. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	−30.7528
1 % Qnt.	−12.5734
2.5 % Qnt.	−8.5558
5 % Qnt.	−6.0530
10 % Qnt.	−3.9267
50 % Qnt.	0.6501
90 % Qnt.	4.3535
95 % Qnt.	5.7762
97.5 % Qnt.	6.7737
99 % Qnt.	8.6403
Maximum	40.7459
Mean	0.3696
Standard Deviation	4.0641
Excess Kurtosis	11.2988
Skewness	−0.5124
T	1822

Table A.2
Descriptive statistics for daily returns on the S&P 500 for the period January 2, 1980, until December 31, 1986.
This table reports the descriptive statistics for daily data on the S&P 500. The data were downloaded from investing.com. The data sample is from January 2, 1980, until December 31, 1986, corresponding to 1769 observations. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	−4.9280
1 % Qnt.	−2.2052
2.5 % Qnt.	−1.7171
5 % Qnt.	−1.4181
10 % Qnt.	−1.0062
50 % Qnt.	0.0301
90 % Qnt.	1.1440
95 % Qnt.	1.5611
97.5 % Qnt.	1.8859
99 % Qnt.	2.3603
Maximum	4.6459
Mean	0.0468
Standard Deviation	0.9027
Excess Kurtosis	1.7455
Skewness	0.0704
T	1769

Table A.3
Descriptive statistics for the bootstrapped sample means for monthly returns on the S&P 500 covering the period January 1871 to November 2022.
This table reports the descriptive statistics for the bootstrapped sample means for monthly returns on the S&P 500. The original data were downloaded from Robert Shiller's data library (www.econ.yale.edu/~shiller/data.htm). The data sample for the original data is from January 1871 to November 2022 corresponding to 1822 observations. The sample means are estimated for $B = 1000$ synthetic samples derived from bootstrapping. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	0.0555
1 % Qnt.	0.1518
2.5 % Qnt.	0.1823
5 % Qnt.	0.2160
10 % Qnt.	0.2514
50 % Qnt.	0.3736
90 % Qnt.	0.4821
95 % Qnt.	0.5135
97.5 % Qnt.	0.5478
99 % Qnt.	0.5835
Maximum	0.6796
Mean	0.3710
Standard Deviation	0.0918
Excess Kurtosis	0.0752
Skewness	−0.1260
T	1000

Table A.4
Descriptive statistics for the bootstrapped sample means for daily returns on the S&P 500 covering the period January 2, 1980, until December 31, 1986.
This table reports the descriptive statistics for the bootstrapped sample means for daily returns on the S&P 500. The original data were downloaded from investing.com. The data sample for the original data is from January 2, 1980, until December 31, 1986, corresponding to 1769 observations. The sample means are estimated for $B = 1000$ synthetic samples derived from bootstrapping. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	−0.0174
1 % Qnt.	−0.0069
2.5 % Qnt.	0.0037
5 % Qnt.	0.0109
10 % Qnt.	0.0209
50 % Qnt.	0.0474
90 % Qnt.	0.0745
95 % Qnt.	0.0809
97.5 % Qnt.	0.0906
99 % Qnt.	0.1014
Maximum	0.1165
Mean	0.0476
Standard Deviation	0.0217
Kurtosis	0.1595
Skewness	−0.0411
T	1000

Table A.5
Descriptive statistics for daily returns on gold futures for the period December 2, 2015, until November 6, 2024.
This table reports the descriptive statistics for daily data on gold futures. The data were downloaded from investing.com. The data sample is from December 2, 2015, until November 6, 2024, corresponding to 2146 observations. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	−5.1140
1 % Qnt.	−2.5490
2.5 % Qnt.	−1.9244
5 % Qnt.	−1.5275
10 % Qnt.	−0.9994
50 % Qnt.	0.0380
90 % Qnt.	1.0772
95 % Qnt.	1.4596
97.5 % Qnt.	1.7699
99 % Qnt.	2.5998
Maximum	5.7754
Mean	0.0368
Standard Deviation	0.9278
Excess Kurtosis	4.1862
Skewness	−0.1002
T	2146

Table A.6
Descriptive statistics for the bootstrapped sample means for daily returns on Gold futures covering the period December 2, 2015, until November 6, 2024.
This table reports the descriptive statistics for the bootstrapped sample means for daily returns on Gold futures. The original data were downloaded from investing.com. The data sample for the original data is from December 2, 2015, until November 6, 2024, corresponding to 2146 observations. The sample means are estimated for $B = 1000$ synthetic samples derived from bootstrapping. Quintiles (Qnt) range from 1 % to 99 %.

Statistic	Value
Minimum	−0.0316
1 % Qnt.	−0.0117
2.5 % Qnt.	−0.0036
5 % Qnt.	0.0033
10 % Qnt.	0.0105
50 % Qnt.	0.0373
90 % Qnt.	0.0628
95 % Qnt.	0.0716
97.5 % Qnt.	0.0764
99 % Qnt.	0.0844
Maximum	0.1100
Mean	0.0370
Standard Deviation	0.0207
Kurtosis	0.1185
Skewness	−0.0646
T	1000

Table A.7
Calibration of the LPPLS model on synthetic monthly data on the S&P 500 derived from blocks bootstrap.
Table A.7 summarizes empirical distributions of LPPLS parameter estimates from 1000 bootstrap resamples of the monthly S&P 500. We report extreme values (minimum/maximum), selected quantiles (1 %, 2.5 %, 5 %, 10 %, 50 %, 90 %, 95 %, 97.5 %, 99 %), the mean and sample standard deviation, skewness, and excess kurtosis (kurtosis relative to the normal distribution). T denotes the number of bootstrap draws ($T = 1,000$). Quantiles are computed from the bootstrap empirical distribution.

Statistic	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$
Minimum	1.2890	−10.0000	0.1000	−0.9990	5.0000	1823.0000	−3.1416
1 % Qnt.	3.9029	−10.0000	0.1000	−0.6802	5.0000	1823.0000	−3.1416
2.5 % Qnt.	4.6018	−10.0000	0.1013	−0.3771	5.0000	1823.0000	−3.1416
5 % Qnt.	5.6585	−10.0000	0.1292	−0.2491	5.0000	1823.0000	−3.1416
10 % Qnt.	6.6099	−10.0000	0.1641	−0.1582	5.0000	1823.0000	−3.1416
Median	11.0734	−0.0177	0.8064	0.0494	5.9342	2014.8585	0.0187
90 % Qnt.	27.4746	−0.0060	0.9000	0.1833	6.9718	2990.7193	3.1416
95 % Qnt.	36.0314	−0.0045	0.9000	0.2455	7.3983	3645.9999	3.1416
97.5 % Qnt.	45.0740	−0.0036	0.9000	0.3535	8.2928	3646.0000	3.1416
99 % Qnt.	54.0501	−0.0021	0.9000	0.5600	8.8448	3646.0000	3.1416
Maximum	81.2357	−0.0008	0.9000	0.9990	10.7550	3646.0000	3.1416
Mean	14.2384	−1.3932	0.6673	0.0293	5.9849	2213.2124	−0.0929
Standard Deviation	10.0383	3.2013	0.2785	0.1938	0.8337	497.3024	2.9512
Excess Kurtosis	7.0997	3.0124	−0.7334	9.1850	3.5021	2.1430	−1.9070
Skewness	2.4068	−2.1934	−0.8747	−0.5422	1.3985	1.7298	0.0325
T	1000	1000	1000	1000	1000	1000	1000

Table A.8
Descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data on monthly S&P 500 data derived from blocks bootstrap. This table presents the descriptive statistics for estimated ADF-test statistics for residuals derived from synthetic daily data on monthly S&P 500 data derived from blocks bootstrap. The critical values for 10 %, 5 %, and 1 % statistical significance levels for the standard ADF test are −1.62, −1.94, and − 2.57. The lag-order is chosen in line with the Schwarz-Criterion. The lag-order is chosen in line with the Schwarz-Criterion. The bootstrap-based inference restores empirical size, confirming that tabulated thresholds overstate significance in LPPLS diagnostics.

Statistic	ADF Statistic	Lag Length
Minimum	−5.9418	1
1 % Qnt.	−5.3167	1
2.5 % Qnt.	−4.8916	1
5 % Qnt.	−4.6730	1
10 % Qnt.	−4.3079	1
Median	−3.2531	1
90 % Qnt.	−2.1984	2
95 % Qnt.	−1.8217	4
97.5 % Qnt.	−1.4922	5
99 % Qnt.	−1.0507	7.0100
Maximum	0.6040	24.0000
Mean	−3.2455	1.7320
Standard Deviation	0.8622	1.6324
Excess Kurtosis	1.0622	77.7525
Skewness	0.1990	7.4943
T	1000	1000

Table A.9
Descriptive statistics for multi-start LPPLS calibrations on the monthly S&P 500 data.
Table A.9 reports empirical distributions of the optimized LPPLS parameters and the objective value (SSE) obtained from $K = 500$ randomized initializations on the original sample. For each quantity we report: minimum, selected quantiles (1 %, 2.5 %, 5 %, 10 %, 50 %, 90 %, 95 %, 97.5 %, 99 %), maximum, mean, sample standard deviation, skewness, excess kurtosis, and T, the number of converged calibrations. Random starting values are drawn for the nonlinear parameters $(\beta, \omega, \phi, t_c)$ i.i.d. Uniform within the bounds as detailed in section 4; the linear coefficients (A, B, C) are initialized by conditional OLS. Convergence is defined ex ante as a positive solver exit flag together with finite, in-bounds estimates; all summaries and quantiles are computed over the resulting set of converged solutions.

Statistic	$\widehat{A}$	$\widehat{B}$	$\widehat{\beta}$	$\widehat{C}$	$\widehat{\omega}$	$\widehat{t_c}$	$\widehat{\phi}$	SSE
Minimum	2.2614	−0.0875	0.5080	−0.0631	5.2251	1852.3000	−3.1416	17.7420
1 % Qnt.	2.3125	−0.0771	0.5080	−0.0587	5.6321	1893.7000	−3.1416	17.7420
2.5 % Qnt.	2.4837	−0.0735	0.5086	−0.0473	5.6623	2013.5000	−3.1416	17.7420
5 % Qnt.	2.7432	−0.0682	0.5365	−0.0450	6.0339	2163.7000	−3.1416	17.7420
10 % Qnt.	2.8153	−0.0455	0.5893	−0.0378	6.8212	2264.5000	−3.1416	17.7420
Median	4.0469	−0.0129	0.7197	0.0139	11.3060	2967.1000	−3.1416	17.7620
90 % Qnt.	6.0199	−0.0025	0.9000	0.0366	13.8370	3646.0000	3.1416	19.2410
95 % Qnt.	6.0745	−0.0024	0.9000	0.0449	14.6670	3646.0000	3.1416	19.3290
97.5 % Qnt.	6.1239	−0.0024	0.9000	0.0500	15.0000	3646.0000	3.1416	19.3580
99 % Qnt.	6.5200	−0.0024	0.9000	0.0500	15.0000	3646.0000	3.1416	20.1960
Maximum	6.6651	−0.0024	0.9000	0.0611	15.0000	3646.0000	3.1416	20.9370
Mean	4.1732	−0.0188	0.7355	0.0015	10.8270	2925.6000	−1.7650	18.0280
Standard Deviation	1.1226	0.0193	0.1184	0.0292	2.6283	516.8600	2.4969	0.5994
Excess Kurtosis	−0.9639	1.6784	−1.1524	−1.1791	−0.9622	−1.2659	−0.0426	3.9179
Skewness	0.3732	−1.5062	−0.0302	−0.1227	−0.3574	−0.1430	1.3652	2.1851
T	263	263	263	263	263	263	263	263

Data availability

Datasets analyzed in this study, together with the MATLAB code and the exact result matrices underlying reported tables and figures, are publicly available on Zenodo (concept DOI: https://doi.org/10.5281/zenodo.17404552). This manuscript corresponds to Version 2 of the deposit. The posted result matrices enable deterministic reproduction of all numbers without rerunning stochastic procedures; scripts for optional re-runs include documented parameter bounds and the ADF specification.

References

Ahn, K., Jang, H., Kim, J., & Ryu, I. (2024). COVID-19 and REITs crash: Predictability and market conditions. *Computational Economics, 63*(3), 1159–1172.
Brée, D. S., Challet, D., & Peirano, P. P. (2013). Prediction accuracy and sloppiness of log-periodic functions. *Quantitative Finance, 13*(2), 275–280.
Brée, D. S., & Joseph, N. L. (2013). Testing for financial crashes using the log periodic power law model. *International Review of Financial Analysis, 30*, 287–297.
Cepni, O., Gupta, R., Nel, J., & Nielsen, J. (2025). Monetary policy shocks and multi-scale positive and negative bubbles in an emerging country: The case of India. *Financial Innovation, 11*(1), 1–25.
Chang, C. L. (2024). Extreme events, economic uncertainty and speculation on occurrences of price bubbles in crude oil futures. *Energy Economics, 130*, Article 107318.

Cifarelli, G., & Paesani, P. (2021). Navigating the oil bubble: A non-linear heterogeneous-agent dynamic model of futures oil pricing. *The Energy Journal, 42*(5), 101–122.
Edmans, A. (2024). *May contain lies: How stories, statistics, and studies exploit our biases—And what we can do about it*. University of California Press.
Efron, B. (1992). Bootstrap methods: Another look at the jackknife. In *Breakthroughs in statistics: Methodology and distribution* (pp. 569–593). New York, NY: Springer New York.
Filimonov, V., & Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A: Statistical Mechanics and its Applications, 392* (17), 3698–3707.
Ghosh, B., Kenourgios, D., Francis, A., & Bhattacharyya, S. (2021). How well the log periodic power law works in an emerging stock market? *Applied Economics Letters, 28* (14), 1174–1180.
Grobys, K. (2023). A finite-time singularity in the dynamics of the US equity market: Will the US equity market eventually collapse? *International Review of Financial Analysis, 89*, Article 102787.
Grobys, K. (2024). No reward—No effort: Will bitcoin collapse near to the year 2140? *Finance Research Letters, 63*, Article 105294.
Grobys, K. (2025). Is gold in the process of a bubble formation? New evidence from the ex-post global financial crisis period. *Research in International Business and Finance, 75*, Article 102727.
Gupta, R., Nel, J., & Nielsen, J. (2023). US monetary policy and BRICS stock market bubbles. *Finance Research Letters, 51*, Article 103435.
Gupta, R., Nel, J., Nielsen, J., & Pierdzioch, C. (2025). Stock market volatility and multi-scale positive and negative bubbles. *The North American Journal of Economics and Finance, 75*, Article 102300.
Gupta, R., Nielsen, J., & Pierdzioch, C. (2024). Stock market bubbles and the realized volatility of oil price returns. *Energy Economics, 132*, Article 107432.
Gustavsson, M., Levén, D., & Sjögren, H. (2016). The timing of the popping: Using the log-periodic power law model to predict the bursting of bubbles on financial markets. *Financial History Review, 23*(2), 193–217.
Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies, 33*(5), 2019–2133.
Huang, W., & Wang, Y. (2024). Identifying price bubbles in global carbon markets: Evidence from the SADF test, GSADF test and LPPLS method. *Energy Economics, 134*, Article 107626.
Ji, H., & Zhang, H. (2024). Application of the LPPL model in the identification and measurement of structural bubbles in the Chinese stock market. *The North American Journal of Economics and Finance, 70*, Article 102060.
Johansen, A., Ledoit, O., & Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance, 3*(02), 219–255.
Johansen, A., & Sornette, D. (2001a). Bubbles and anti-bubbles in Latin-American, Asian and Western stock markets: An empirical study. *International Journal of Theoretical and Applied Finance, 4*, 853–920.
Johansen, A., & Sornette, D. (2001b). Finite-time singularity in the dynamics of the world population, economic and financial indices. *Physica A: Statistical Mechanics and its Applications, 294*(3–4), 465–502.
Lin, L., Ren, R. E., & Sornette, D. (2014). The volatility-confined LPPL model: A consistent model of 'explosive'financial bubbles with mean-reverting residuals. *International Review of Financial Analysis, 33*, 210–225.
Mandelbrot, B. B. (2008). *The (mis) behaviour of markets: A fractal view of risk, ruin and reward*. Profile books.
Shu, M., & Song, R. (2024). Detection of financial bubbles using a log-periodic power law singularity (LPPLS) model. *Wiley Interdisciplinary Reviews: Computational Statistics, 16* (2), Article e1649.
Shu, M., Song, R., & Zhu, W. (2021). The 'COVID'crash of the 2020 US stock market. *The North American Journal of Economics and Finance, 58*, Article 101497.
Song, R., Shu, M., & Zhu, W. (2022). The 2020 global stock market crash: Endogenous or exogenous? *Physica A: Statistical Mechanics and its Applications, 585*, Article 126425.
Sornette, D. (2017). *Why stock markets crash: Critical events in complex financial systems*. Princeton university press.
Sornette, D., & Zhou, W. X. (2002). The US 2000-2002 market descent: How much longer and deeper? *Quantitative Finance, 2*(6), 468.
Van Eyden, R., Gupta, R., Nielsen, J., & Bouri, E. (2023). Investor sentiment and multi-scale positive and negative stock market bubbles in a panel of G7 countries. *Journal of Behavioral and Experimental Finance, 38*, Article 100804.
Xu, H. C., Tan, Y. Z., Fan, H. X., & Zhou, W. X. (2025). Early warning of bubbles in the agricultural commodity market: Evidence from LPPLS confidence indicators. *Journal of Management Science and Engineering, 10*(2), 245–261.
Yan, W., Woodard, R., & Sornette, D. (2010). Diagnosis and prediction of tipping points in financial markets: Crashes and rebounds. *Physics Procedia, 3*(5), 1641–1657.
Yang, J., Dong, D., Liang, C., & Cao, Y. (2024). Monetary policy uncertainty and the price bubbles in energy markets. *Energy Economics, 133*, Article 107503.
Zhang, J., Wang, H., Chen, J., & Liu, A. (2024). Cryptocurrency price bubble detection using log-periodic power law model and wavelet analysis. *IEEE Transactions on Engineering Management, 71*, 11796–11812.
Zhang, Q., Zhang, Q., & Sornette, D. (2016). Early warning signals of financial crises with multi-scale quantile regressions of log-periodic power law singularities. *PLoS One, 11* (11), Article e0165819.
Zhao, D., & Sornette, D. (2021). Bubbles for Fama from Sornette. In *Swiss finance institute research paper* (pp. 21–94).
Zhou, W. X., & Sornette, D. (2006). Fundamental factors versus herding in the 2000–2005 US stock market and prediction. *Physica A: Statistical Mechanics and its Applications, 360*(2), 459–482.