

The Surge of Bitcoin: Is It Just Another Tulip Mania?

Explosive price dynamics and the limits
of a behavioral interpretation

Woosub Shin

University of Copenhagen
Department of Economics

Seminar Paper for Advanced Finance
Advisors: Frank Hansen and Søren Hovgaard

Original paper: December 20, 2025
Revised edition: October 14, 2026

Abstract

Bitcoin’s repeated price surges invite comparisons with historical speculative episodes, but a price series alone cannot identify why investors trade or whether an asset is overvalued. This paper examines the evidence for explosive dynamics in Bitcoin’s reported daily log closing prices from January 2015 to January 2025 and considers how that evidence relates to behavioral finance. The original analysis reports a generalized supremum augmented Dickey-Fuller (GSADF) statistic of 3.42 and simulated 95% and 99% critical values of 2.72 and 3.15. Taken at face value, these numbers reject the specified unit-root null in favor of some explosive behavior. However, the underlying data and implementation have not been recovered, and inconsistencies between the reported statistic, plotted sequence, minimum window, and return summaries prevent an independently verified claim about the number or exact dates of episodes. The approximate periods discussed in the original analysis are 2017, 2020–2021, and 2024 through the January 2025 sample endpoint. Reported regime means and standard deviations are retained as descriptive quantities; unsupported significance claims are withdrawn. Herding, overconfidence, and narrative transmission remain possible interpretations, alongside changing fundamentals, adoption, liquidity, and macroeconomic conditions. The evidence therefore motivates further investigation of recurrent explosive dynamics without identifying behavioral causality or establishing an equivalence with Tulip Mania.

Keywords: Bitcoin; explosive dynamics; GSADF; BSADF; behavioral finance; reproducibility.

1 Introduction

Bitcoin’s large price increases and subsequent reversals have made it a prominent setting for studying speculative dynamics. The historical analogy in the title raises a useful question: can rapid appreciation reflect feedback trading and investor enthusiasm, even when the underlying technology has economic uses? Answering that question requires separating three claims that are often conflated: a price series can be statistically explosive; a price can exceed a model of fundamental value; and a particular behavioral mechanism can cause that deviation. Evidence for the first claim does not establish the other two.

Historical accounts of financial manias emphasize recurring interactions between innovation, credit, optimism, and reversals (Kindleberger and Aliber, 2011). Behavioral finance adds mechanisms through which beliefs and trading may reinforce price movements. Herding, overconfidence, limited attention, and reference-dependent preferences can affect investment decisions (Bikhchandani and Sharma, 2001; Barberis and Thaler, 2003). These mechanisms offer hypotheses for cryptocurrency markets, but their relevance must be tested against alternative explanations rather than inferred from price increases alone.

Bitcoin also differs from the assets in many historical comparisons. Conventional discounted-dividend valuation is difficult to apply to an asset that does not promise a dividend stream. That difficulty does not imply that its fundamental value is zero or that changes in adoption, network activity, scarcity, settlement services, and access to trading are economically irrelevant. Nor does it make all high returns irrational. A claim that behavioral demand dominates fundamentals needs measurements and a model capable of distinguishing those channels.

The empirical question addressed here is therefore narrower than a causal comparison between psychology and fundamentals: **Does the available evidence support recurrent explosive behavior in Bitcoin prices, and what can that evidence establish about behavioral explanations?** The analysis focuses on the reported January 2015–January 2025 sample. It explains the recursive right-tailed unit-root framework of Phillips, Shi, and Yu (2015a, 2015b),

distinguishes the full-sample GSADF statistic from the backward SADF (BSADF) sequence used for date-stamping, and evaluates the interpretation of the original reported results.

The original paper supplied useful price illustrations and numerical summaries, but did not document enough information to reproduce the experiment. Its time-varying statistic plot is also inconsistent with the reported global statistic. The revised analysis retains the numerical record and explicitly limits the conclusions it can support. It does not substitute a newly downloaded price series, an assumed minimum window, or an unperformed robustness test for missing evidence.

The contribution is consequently an exploratory empirical account with a corrected methodological interpretation. The reported scalar result is evidence conditional on the original implementation; the regime comparisons describe reported associations in the same selected price sample; and the behavioral discussion develops hypotheses rather than causal findings. Section 2 reviews the relevant literature. Section 3 defines the data transformations and econometric framework. Section 4 evaluates the reported results. Section 5 discusses explanations, implications, and a replication design. Section 6 concludes. The appendices record the numerical reconciliation requirements and the provenance of this edition.

2 Related literature

2.1 Speculation and explosive dynamics

A financial bubble is commonly understood as a component of price that cannot be justified by the relevant fundamental valuation model. The empirical challenge is that fundamental value is not directly observed and depends on assumptions about future services, cash flows, risk, and discounting. Historical narratives such as Kindleberger and Aliber (2011) and Shiller (2015) help organize possible mechanisms, but they do not by themselves provide a statistical classification rule for a particular asset.

Recursive right-tailed unit-root tests address a related, more specific question: whether a time series contains intervals of explosive autoregressive behavior. Phillips, Shi, and Yu (2015a, 2015b) develop procedures that allow estimation windows to vary, improving detection and dating when explosive episodes recur. Their framework is particularly relevant when a single full-sample regression could obscure short-lived or multiple episodes. The economic interpretation nevertheless depends on the variable tested, the maintained stochastic model, and the validity of the critical values. An explosive log-price process is not a direct estimate of a price–fundamental-value gap.

This distinction matters for cryptocurrency applications. Bellón and Figuerola-Ferretti (2022) study explosive behavior in Ether prices and examine potential blockchain fundamentals. Their reported interpretation emphasizes growth in demand and fundamentals rather than treating every explosive episode as a speculative bubble. Oldani, Bruno, and Signorelli (2025) apply methods allowing for nonstationary volatility to Bitcoin, Ether, and Ripple over 2018–2024. Their findings differ across currencies and tests. These studies support careful measurement and specification checks, rather than a general conclusion that all cryptocurrency price surges have the same cause.

2.2 Behavioral mechanisms

Herding can arise when investors condition their decisions on other investors’ actions. Bikhchandani and Sharma (2001) distinguish intentional imitation from correlated choices driven by com-

mon information. Observing many investors buying at the same time is therefore insufficient to establish an information cascade. In this paper, the price series does not reveal individual information sets or order-level behavior, so herding is an interpretation to investigate.

Overconfidence is another possible source of excessive trading. Barber and Odean (2001) examine gender, trading, and investment performance among individual common-stock investors. Their findings provide evidence relevant to overconfidence in that setting; they do not establish that overconfidence causes Bitcoin volatility. Applying this mechanism to cryptocurrencies requires additional information on participation, turnover, or beliefs.

Prospect theory describes reference-dependent choice and asymmetric responses to gains and losses (Kahneman and Tversky, 1979). It can motivate hypotheses about holding losses or responding to recent gains, but it does not uniquely predict a Bitcoin boom–bust cycle or directly measure fear of missing out. Narrative transmission and social comparison are related possibilities. Shiller (2015) discusses the social and psychological setting of speculative enthusiasm. In the present analysis, these ideas explain why attention variables might be useful; they are not observed channels.

2.3 Cryptocurrency evidence and fundamentals

Bhambhwani, Delikouras, and Korniotis (2019) examine computing power and network size as cryptocurrency fundamentals and document long-run relationships with prices. Their analysis should not be summarized as demonstrating that speculative sentiment dominates these fundamentals. It instead motivates including economically relevant network variables when assessing competing explanations of price movements.

Haykir and Yagli (2022) study speculative bubbles and herding in cryptocurrency markets. Their results indicate that herding diminishes during bubble periods in their setting, which cautions against treating higher prices or explosive episodes as automatic evidence of stronger herding. Chowdhury and Damianov (2024) investigate relationships between uncertainty measures and cryptocurrency bubbles. Such evidence concerns specific uncertainty measures and empirical specifications; it cannot substitute for an omitted uncertainty or sentiment variable in the present paper.

Taken together, the literature provides competing mechanisms and methodological tools. It does not justify attributing the present sample’s price changes primarily to psychology. A credible behavioral explanation would have to show that measured attention, participation, or trading behavior is associated with the episodes after accounting for relevant common information, market conditions, and fundamental variables. Establishing causality would require a stronger design still.

3 Data and econometric framework

3.1 Reported sample and transformations

The original analysis describes daily Bitcoin prices from January 2015 through January 2025, with open, high, low, and closing prices, volume, and daily percentage changes. Only closing prices enter the reported econometric analysis. The price figures label their units as US dollars. The original paper does not identify the data vendor or exchange, instrument or quote pair, exact first and last observation, closing timezone, sample size, duplicate handling, or missing-day treatment. These omissions matter in a market that trades continuously across venues. A

dollar label alone does not establish whether the series is a particular BTC/USD market or an aggregate converted to dollars.

Let P_t denote the observed closing price, $y_t = \log(P_t)$ its natural logarithm, and $r_t = y_t - y_{t-1}$ the daily log return. All reported return values in this paper are in decimal units. Thus a mean of 0.0015 is a mean log return of 0.15% per observation, not a simple-return average. The original figures use monthly average prices for the long-run illustration and daily prices for the episode panels. Monthly aggregation is a visualization choice; it must not be substituted for the daily input to a test without recalculating the test and its critical values.

The original descriptive statistics and figures are reproduced or summarized below. Their underlying observations were not available for this revision, so the numerical results remain reported quantities rather than independently recomputed estimates. Appendix A identifies checks needed to make them reproducible. No empirical attention, sentiment, participation, or fundamental variable is added in this edition.

3.2 Right-tailed ADF regression

An equivalent and convenient difference-form parameterization of the augmented Dickey-Fuller regression is

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{j=1}^p \phi_j \Delta y_{t-j} + \varepsilon_t. \quad (1)$$

Here α is an intercept, p is the augmentation lag order, and β is the coefficient used to assess persistence. The right-tailed hypotheses are $H_0: \beta = 0$ against $H_1: \beta > 0$. In the original level-form parameterization, the corresponding coefficient is $\delta = 1 + \beta$, so the same hypotheses are $\delta = 1$ against $\delta > 1$. Equation (1) displays an intercept and no deterministic trend. It is a statement of the framework, not verification that the unavailable original code used that exact specification.

For each admissible subsample, the ADF statistic is the t-ratio for β under its null value. Its right-tail critical distribution is nonstandard. The deterministic terms, lag-selection rule, effective regression length, and simulation design must match the implementation used to construct critical values. A conventional left-tailed unit-root critical value is inappropriate for this explosive alternative. An econometric conclusion also concerns the transformed variable actually tested: testing log prices and testing prices are different specifications.

3.3 Global testing and endpoint date-stamping

Let T denote the total number of input observations, r_1 and r_2 the fractional start and end points, and r_0 the minimum admissible window fraction. Write $\text{ADF}(r_1, r_2)$ for the statistic from that subsample. The global GSADF statistic is

$$\text{GSADF}(r_0) = \sup_{r_2 \in [r_0, 1]} \sup_{r_1 \in [0, r_2 - r_0]} \text{ADF}(r_1, r_2). \quad (2)$$

It is one scalar for the specified sample and set of windows. A rejection indicates evidence of some explosive behavior under the maintained model. It does not by itself establish the number of episodes, their dates, a causal mechanism, or an economically measured bubble.

For a fixed endpoint r_2 , the backward SADF statistic varies the starting point:

$$\text{BSADF}_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \text{ADF}(r_1, r_2). \quad (3)$$

The BSADF sequence provides the endpoint-specific evidence used for date-stamping. Its relevant critical-value sequence must be obtained under the corresponding endpoint-specific null experiment. The full-sample GSADF critical value is not automatically a valid threshold for every point of that sequence. Under the same admissible windows and regression settings, GSADF equals the maximum of BSADF across endpoints. This identity provides a simple implementation and reporting check.

An episode rule compares the BSADF sequence with its matched critical-value sequence and specifies how crossings become start and end dates. The empirical report should state the significance level, minimum duration, treatment of short interruptions, lag and endpoint alignment, and the treatment of an episode still active at the sample endpoint. Repeated pointwise crossings also require care: pointwise critical values do not on their own make a set of dated intervals a simultaneous confidence statement. The dating procedure should follow the chosen method rather than an informal visual reading of shaded regions.

3.4 Critical values and the minimum window

The original appendix reports 2,000 Monte Carlo replications and a minimum window of “36 months.” This wording is unresolved. With daily observations, 36 observations and 36 calendar months define very different experiments. The latter would generally exclude test endpoints during roughly the first three years of a sample beginning in January 2015, whereas the original statistic figure starts near the beginning of that sample. No conversion to observations, lag adjustment, or reason for the selected minimum window is given.

A replicable specification should report the minimum window both as an integer observation count and as r_0 , together with T . It should also report the null data-generating process, innovation distribution, initial condition, deterministic terms, lag order or selection criterion, simulation seed, number of replications, and quantile rule. The simulated global critical values and the endpoint-specific dating values should be saved separately. The reported values in Section 4 are not recalibrated here because choosing any of these missing inputs would create a different experiment.

Volatility variation is an additional concern in cryptocurrency data. Under nonstationary volatility, standard homoskedastic simulation critical values need not provide reliable size control. Harvey et al. (2016) develop a wild-bootstrap approach to recursive explosive-bubble testing under nonstationary volatility. Phillips and Shi (2020) discuss bootstrap PSY monitoring that also addresses family-wise error control; related applications include Oldani et al. (2025). These procedures have specific targets and are not interchangeable critical-value generators. A suitable heteroskedasticity-robust procedure is a substantive robustness exercise, not a larger number of replications from an unchanged homoskedastic design. No such new bootstrap is claimed in this paper.

3.5 Comparing returns across selected regimes

Suppose a completed dating procedure assigns a label D_t to each return observation, with $D_t = 1$ for a flagged interval and $D_t = 0$ otherwise. The date alignment must specify whether a label at endpoint t applies to the return ending at t and must avoid treating a retrospective label as

an ex ante trading signal. Group means and standard deviations are then useful descriptions of the selected observations.

For independent groups, an unequal-variance mean comparison uses the Welch statistic

$$t_W = \frac{\bar{r}_1 - \bar{r}_0}{\sqrt{s_1^2/n_1 + s_0^2/n_0}}. \quad (4)$$

Its approximate t reference distribution uses Welch–Satterthwaite degrees of freedom, not an unspecified t distribution. The original paper did not report n_1 , n_0 , the degrees of freedom, exact p-values, or a complete return-to-regime mapping. Financial returns may also exhibit dependence, changing variance, and non-normality. A conventional independent-sample calculation is therefore insufficient to establish robust significance in this setting.

There is a separate problem with the original volatility comparison. One sample standard deviation from each regime is not a sample of volatility observations to which an ordinary two-sample mean t-test can be applied. A variance or standard-deviation comparison needs a stated sampling design and appropriate inference. Possible designs include a dependence-aware bootstrap for $s_1 - s_0$ or a test based on absolute deviations from regime medians. If rolling volatility observations are used instead, their definition and overlap must be reported and their serial dependence addressed. None of these alternatives is silently substituted for the original unexplained statistic.

Most importantly, the labels are selected using the same prices from which returns are computed. Higher returns in intervals selected for explosive price behavior are not independent confirmation of bubbles or evidence of a behavioral cause. HAC standard errors in a return-on-label regression can address certain dependence problems conditional on labels, but they do not eliminate this selection problem. A resampling analysis that seeks to describe uncertainty in the entire procedure must consider repeating detection and classification as well as the regime comparison. The reported comparisons below are consequently descriptive.

4 Reported empirical results

4.1 Distribution of daily returns

Table 1 preserves the return summaries from the original paper. The main text and appendix report the same mean, standard deviation, skewness, and kurtosis; the appendix adds the median and Jarque–Bera statistic.

Statistic	Reported value
Mean daily log return	0.0015
Median daily log return	0.0006
Daily return standard deviation	0.0482
Skewness	1.87
Kurtosis	9.24
Jarque–Bera statistic	6241.38

Table 1. Originally reported descriptive statistics, January 2015–January 2025. Values have not been recomputed. The original discussion compares kurtosis with the Gaussian benchmark of 3, implying ordinary rather than excess kurtosis; the software convention remains to be verified. The original significance marker for Jarque–Bera is omitted because the sample and test implementation are unavailable.

The reported standard deviation corresponds to 4.82% in daily log-return units. Positive skewness describes asymmetry in the third central moment; it does not by itself prove that large positive moves occur more frequently than equally large negative moves. Kurtosis of 9.24, interpreted on the ordinary scale used in the original discussion, is above the Gaussian benchmark. These quantities suggest a non-Gaussian empirical distribution if they were calculated as described. They do not identify speculative motivation, predict a crash probability, or distinguish behavioral effects from changes in economic conditions.

The retained long-run price illustration provides context for the periods discussed in the paper. Its shaded intervals are broad retrospective annotations. They should not be read as exact output from a verified date-stamping procedure.

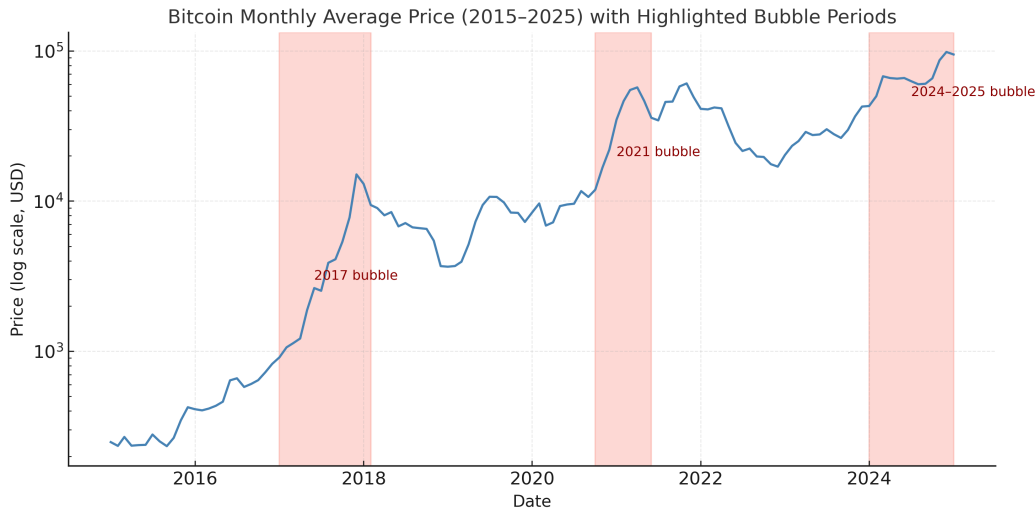


Figure 1. Monthly average Bitcoin price in US dollars, reproduced from the original paper. The vertical scale is logarithmic. Shading and embedded “bubble” labels are original illustrative annotations; they are not independently verified BSADF crossing dates. The underlying daily series and aggregation code were not recovered.

4.2 The reported global statistic

Quantity	Reported value
Observed GSADF	3.42
90% critical value	2.41
95% critical value	2.72
99% critical value	3.15
Monte Carlo replications	2,000

Table 2. Originally reported global test statistic and simulation critical values. The 90%, 95%, and 99% quantiles correspond to upper-tail test sizes of 10%, 5%, and 1%. The original appendix’s minimum-window description is “36 months”; its implementation is unresolved.

The reported statistic exceeds all three reported critical values. If the statistic and critical values were generated under the same correctly specified experiment, this comparison would reject the unit-root null at each listed nominal level, including 1%. It would support the presence of at least some explosive behavior, not prove three separate episodes. The absence of the original experiment prevents a verified significance statement, and no exact p-value can be recovered from these three quantiles.

A further inconsistency prevents use of the original time-varying statistic figure for dating. The plotted path reaches approximately 4.4, above the reported global supremum of 3.42. A BSADF sequence generated from the same regressions could not exceed its GSADF maximum. The plot also uses an approximately constant threshold labeled 2.7 and begins before a literal 36-month minimum window would permit. The figure cannot therefore be repaired by changing “GSADF” to “BSADF” in its caption. It is retained in the unchanged original archive, but is not presented here as a validated statistical result.

4.3 Approximate periods and price illustrations

The original discussion identifies three broad periods. Table 3 records them without converting them into invented start dates, end dates, or durations.

Period discussed in original paper	Available description	Dating status
2017	Broad 2017 phase; text emphasizes late 2017	Exact BSADF crossings and duration unavailable
2020–2021	Broad phase around the 2021 surge; text emphasizes early 2021	Exact BSADF crossings and duration unavailable
2024–January 2025	Text describes mid-2024 through early 2025	Exact crossings unavailable; sample ends in January 2025

Table 3. Approximate historical periods discussed in the original paper. The descriptions are not a verified list of detected episodes. Plotting windows in Figure 2 include observations before and after the first two periods and must not be interpreted as episode durations.

The daily price panels show substantial increases and subsequent declines in the first two plotting windows. The third panel ends near the January 2025 sample boundary, so it does not establish a completed post-episode collapse or an eventual end date. Calling all three panels completed boom-and-collapse cycles would overstate what is visible. The retained figures are descriptive evidence of price paths, not evidence of the motives of traders.

4.4 Descriptive regime comparisons

Reported quantity	Originally labeled bubble observations	Remaining observations
Mean daily log return	0.0048	0.0009
Daily return standard deviation	0.0675	0.0412

Table 4. Original regime summaries retained for transparency. Labels are not independently verified, and group sizes and exact observations are unavailable. The table reports no valid new significance test. The original return t-statistic of 2.57 and volatility statistic of 3.03 are documented in Appendix A rather than treated as verified inferential results.

The reported means correspond to 0.48% and 0.09% in daily log-return units, while the standard deviations correspond to 6.75% and 4.12%. These values describe higher average returns and larger dispersion in the originally flagged observations. They are conditional on an unverified labeling procedure and should not be used as forecasts of future conditional returns or as an independent validation of explosive behavior.

The summaries also require a common-sample reconciliation. If Table 4 partitions exactly the return observations in Table 1, the full-sample mean and variance must equal their pooled

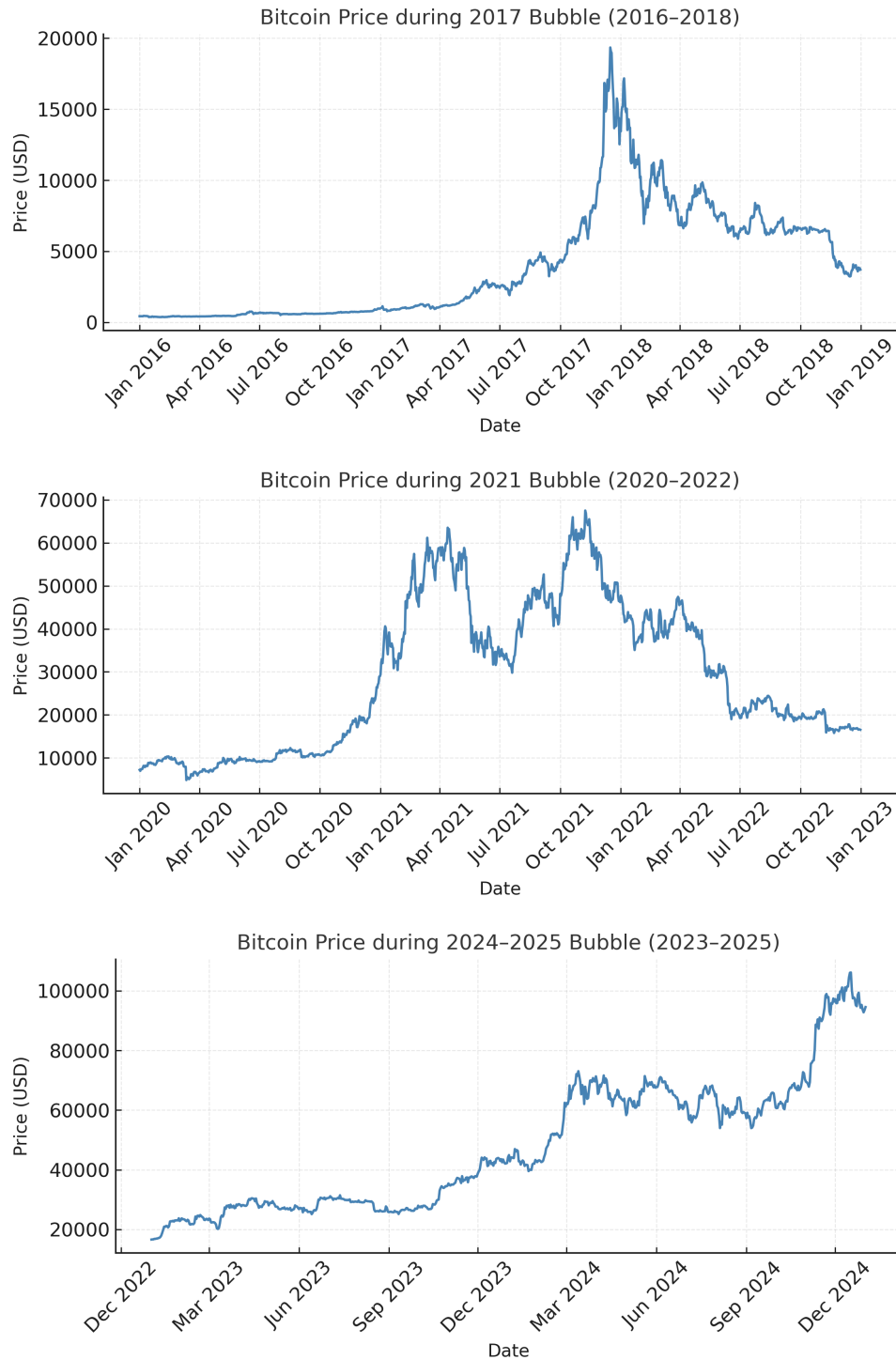


Figure 2. Original daily price illustrations for the periods surrounding the 2017, 2021, and 2024–January 2025 surges. The three panels use different linear vertical scales in US dollars. Original embedded titles use “bubble” as a descriptive label, not a newly verified classification. The third panel is bounded by the sample endpoint.

counterparts. The rounded means imply a flagged-observation share of approximately 15.4%. Combining that share with the reported group standard deviations gives a large-sample pooled standard deviation of approximately 0.04625, rather than 0.0482. This is an arithmetic diagnostic using the printed values, not a replacement empirical estimate. Differences in sample selection, return definition, or computation could explain the discrepancy; ordinary rounding alone is insufficient under a large daily sample. Without counts and observations, the discrepancy remains unresolved.

4.5 What the reported evidence establishes

The numerical record motivates examination of explosive dynamics, but its inferential status must remain conditional. The scalar test comparison is potentially informative if its implementation is recovered and validated. The figures show pronounced price movements, but do not supply valid test dates. The return tables describe the originally selected groups, but lack the information needed for robust inference and common-sample reconciliation. No attention series, investor-level behavior, or fundamentals enter the reported empirical model. The analysis therefore cannot rank behavioral explanations against fundamental or institutional explanations.

5 Interpretation, limitations, and further analysis

5.1 Behavioral explanations as hypotheses

Herding and narrative transmission could generate feedback when visible gains attract attention, new participation supports prices, and further appreciation reinforces beliefs. Such a process is compatible with explosive dynamics. Compatibility is weaker than identification: common news about adoption, access, or liquidity could produce correlated participation and rising prices without an information cascade. The price-only design cannot distinguish those cases.

Overconfidence could encourage turnover or risk-taking, but the present paper does not analyze turnover, leverage, individual beliefs, or trading profits. Higher return dispersion is compatible with many market mechanisms. It is not a measure of overconfidence. Likewise, fear of missing out is a plausible description of some investors' motivations, but neither a price chart nor prospect theory alone establishes its contribution in these episodes.

Attention measures would provide a closer empirical link to the proposed mechanisms. A search-intensity series, a consistently defined volume measure, or an investor-participation measure could be aligned with verified episode dates. Even then, simultaneous movement would remain an association: rising prices may attract attention, attention may affect demand, and both may respond to common information. A causal claim needs a design that addresses this feedback and its timing.

5.2 Alternative explanations and the historical analogy

Changes in network adoption, transaction services, market access, institutional demand, liquidity, and discount rates are possible explanations for changing prices. The analysis does not observe these channels and cannot dismiss them. Results from Bhambhwani et al. (2019), Bel-lón and Figuerola-Ferretti (2022), and Haykir and Yagli (2022) show why a literature review must allow evidence that complicates a simple sentiment-driven account.

The Tulip Mania analogy is consequently a framing device, not an empirical equivalence established by this paper. A rigorous historical comparison would require evidence on the assets' uses,

market institutions, financing, participation, price formation, and the meaning of observed prices in each setting. This study does not make that comparison. Its findings can motivate questions about recurrent speculation without concluding that Bitcoin is “just another” historical mania or that economic usefulness rules speculation out.

The recurrence of pronounced price increases is also not itself a mechanism. Repeated shifts in attention, liquidity, adoption, or the macroeconomic environment could create different episodes for different reasons. Halving events or institutional developments are candidate explanatory variables, but event-specific claims require a dated event record, prespecified windows, and controls. No such event study is performed here.

5.3 Implications within the evidence

The illustrations and reported dispersion show why a reader should distinguish high observed returns from stable or predictable returns. They do not produce a trading strategy, estimate the probability of a crash, or establish how a position should be sized. Retrospective episode labels are especially unsuitable for claims about attainable real-time performance unless the classification is reproduced using only information available at each date.

For public discussion, the methodological implication is that the word “bubble” should be tied to the evidence supporting it. A statistical rejection, an estimated valuation gap, a documented trading mechanism, and an observed collapse are different findings. The present study does not compare regulatory interventions or establish that education, disclosure, or any particular policy is more effective than another. Policy recommendations would require evidence beyond the price-only analysis.

5.4 A reproducible replication design

The first priority is to recover or rebuild the empirical record. A replication should preserve the exact raw price file and its source, instrument, quote units, closing timezone, retrieval date, and sample endpoints. It should retain a cleaning log for missing dates, duplicates, nonpositive values, and changes in the source series. The daily calendar, price-to-return transformation, monthly aggregation, and final observation counts should be explicit.

The second priority is to implement one internally consistent recursive experiment. The regression specification, lag selection, minimum observation window, endpoint alignment, and null simulation must be saved with the results. An automated check should verify that the global GSADF equals the maximum of the BSADF sequence under identical settings. Exact episode dates should be generated from the stated threshold and duration rule, accompanied by the relevant critical-value sequence. An interval extending to the final observation must be reported as unresolved at the endpoint rather than assigned an inferred later termination.

The third priority is sensitivity analysis. Alternative lag specifications, economically defensible minimum windows, and appropriately recalibrated weekly or monthly **price-level** specifications can assess how much the conclusions depend on the original design. A test on returns would address a different process and should not be presented as the same price-explosiveness test at another frequency. A volatility-robust bootstrap should be implemented using a procedure valid for the chosen statistic, and its results should be distinguished from conventional Monte Carlo calibration. More simulations reduce simulation noise; they cannot repair a misspecified null or an incorrect dating statistic.

Finally, regime comparisons should be regenerated from the verified labels and a common return ledger. Report group sizes, means, medians, standard deviations, uncertainty estimates, and the exact sampling rule. Any bootstrap should distinguish inference conditional on fixed labels from

uncertainty that includes the detection step. HAC or dependence-aware methods may improve sampling inference, but neither can turn a price-derived classification into an independent explanation of price behavior. Adding behavioral proxies and fundamental controls would extend the empirical question; it should be reported as new research rather than as a correction to the original numerical results.

5.5 Limits of this edition

This edition corrects the exposition, interpretation, references, and treatment of the available evidence. It does not claim that a new Monte Carlo experiment, robust variance test, attention analysis, or empirical replication has been completed. The original data and executable analysis were not recovered in the targeted source search. Consequently, exact dates, sample counts, minimum-window implementation, lag selection, and critical-value generation remain unresolved. These gaps restrict the strength of the empirical conclusion rather than merely suggesting optional future refinements.

6 Conclusion

The original January 2015–January 2025 Bitcoin analysis reports a GSADF statistic above its simulated critical values and describes several periods of rapid price appreciation. Those results are consistent with a research question about recurrent explosive dynamics. They are not sufficient to demonstrate three precisely dated bubbles, a deviation from a measured fundamental value, or the dominance of behavioral causes.

The distinction between the global GSADF test and the BSADF dating sequence is central. The original plotted sequence cannot be reconciled with the reported global statistic under a common specification, and its early endpoints conflict with a literal interpretation of the stated minimum window. The return comparisons also require a verified regime map and a common-sample reconciliation. Their reported means and standard deviations remain useful parts of the original record, but the unsupported volatility t-test and unverified significance claims should not carry the conclusion.

Behavioral finance supplies plausible hypotheses for further investigation. Herding, overconfidence, and narrative transmission may contribute to speculative episodes, while adoption, fundamentals, liquidity, and macroeconomic conditions offer alternative or complementary explanations. The present evidence does not distinguish their effects. A defensible answer to the title is therefore limited: Bitcoin’s price history motivates comparison with speculative episodes, but this study does not establish that it is another Tulip Mania. Recovering the empirical inputs, validating the dating procedure, and measuring explanatory channels are necessary before making a stronger claim.

References

- Barber, B. M., and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1), 261–292. <https://doi.org/10.1162/003355301556400>
- Barberis, N., and R. Thaler (2003). A survey of behavioral finance. In G. M. Constantinides, M. Harris, and R. M. Stulz (Eds.), *Handbook of the Economics of Finance*, Vol. 1B, pp. 1053–1128. Elsevier. [https://doi.org/10.1016/S1574-0102\(03\)01027-6](https://doi.org/10.1016/S1574-0102(03)01027-6)
- Bellón, C., and I. Figuerola-Ferretti (2022). Bubbles in Ethereum. *Finance Research Letters*, 46, Part B, 102387. <https://doi.org/10.1016/j.frl.2021.102387>
- Bhambhwani, S. M., S. Delikouras, and G. M. Korniotis (2019). *Do fundamentals drive cryptocurrency prices?* Working paper, October 30. https://wpcarey.asu.edu/sites/g/files/litvpz246/files/2021-11/george_korniotis_seminar_paper_november_8_2019.pdf
- Bikhchandani, S., and S. Sharma (2001). Herd behavior in financial markets. *IMF Staff Papers*, 47(3), 279–310. <https://www.imf.org/external/pubs/ft/staffp/2001/01/pdf/bikhchan.pdf>
- Chowdhury, M. S. R., and D. S. Damianov (2024). Uncertainty and bubbles in cryptocurrencies: Evidence from newly developed uncertainty indices. *International Review of Financial Analysis*, 91, 102949. <https://doi.org/10.1016/j.irfa.2023.102949>
- Harvey, D. I., S. J. Leybourne, R. Sollis, and A. M. R. Taylor (2016). Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance*, 38, Part B, 548–574. <https://doi.org/10.1016/j.jempfin.2015.09.002>
- Haykir, O., and I. Yagli (2022). Speculative bubbles and herding in cryptocurrencies. *Financial Innovation*, 8, 78. <https://doi.org/10.1186/s40854-022-00383-0>
- Kahneman, D., and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Kindleberger, C. P., and R. Z. Aliber (2011). *Manias, Panics and Crashes: A History of Financial Crises* (6th ed.). Palgrave Macmillan.
- Oldani, C., G. S. F. Bruno, and M. Signorelli (2025). Collapsing bubbles in the prices of cryptocurrencies. *The Journal of Economic Asymmetries*, 31, e00420. <https://doi.org/10.1016/j.jeca.2025.e00420>
- Phillips, P. C. B., and S. Shi (2020). Real time monitoring of asset markets: Bubbles and crises. In H. D. Vinod and C. R. Rao (Eds.), *Financial, Macro and Micro Econometrics Using R, Handbook of Statistics*, Vol. 42, pp. 61–80. Elsevier. <https://doi.org/10.1016/bs.host.2018.12.002>
- Phillips, P. C. B., S. Shi, and J. Yu (2015a). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4), 1043–1078. <https://doi.org/10.1111/iere.12132>
- Phillips, P. C. B., S. Shi, and J. Yu (2015b). Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review*, 56(4), 1079–1134. <https://doi.org/10.1111/iere.12131>
- Shiller, R. J. (2015). *Irrational Exuberance* (3rd ed.). Princeton University Press.

Appendix A Numerical record and reconciliation

A.1 Original test statistics

The original mean-return comparison reported a statistic of 2.57 with a 5% significance marker. The printed formula corresponds to an unequal-variance mean statistic, but sample sizes, degrees of freedom, label alignment, and dependence adjustments were not supplied. The original volatility row reported 3.03 with the same marker, without specifying a sample of volatility observations or a valid standard-deviation comparison. These two values are preserved here as part of the source record, not as verified tests. They are deliberately excluded from Table 4's inferential presentation.

The original appendix reported a Jarque–Bera statistic of 6241.38 with a 1% significance marker. Its exact sample size, treatment of missing returns, skewness and kurtosis estimators, and reference calibration were not reported. The numerical statistic is retained in Table 1. None of these significance markers has been replaced with a newly calculated p-value.

A.2 The pooling identities

If the reported groups partition the full return sample, let $N = n_1 + n_0$. The mean identity is

$$\bar{r} = \frac{n_1 \bar{r}_1 + n_0 \bar{r}_0}{N}. \quad (\text{A1})$$

For standard deviations computed with the usual sample-variance denominator, the variance identity is

$$(N - 1)s^2 = (n_1 - 1)s_1^2 + (n_0 - 1)s_0^2 + n_1(\bar{r}_1 - \bar{r})^2 + n_0(\bar{r}_0 - \bar{r})^2. \quad (\text{A2})$$

Using the printed means gives $n_1/N \approx (0.0015 - 0.0009)/(0.0048 - 0.0009) \approx 0.15385$. In a large sample, the corresponding pooled standard deviation is approximately 0.04625. Allowing each displayed four-decimal mean and standard deviation to vary by half its rounding unit bounds the large-sample pooled standard deviation below approximately 0.04710, still below the reported full-sample value of 0.0482. This check assumes the same return definition and an exhaustive, nonoverlapping regime partition. The exact sample-variance identity should be checked once n_1 and n_0 are recovered. The present calculation establishes a reporting-reconciliation issue, not which entry is wrong.

A.3 Minimum replication record

A completed empirical record needs the raw data source and immutable file; exact dates, units, timezone, and counts; transformation and cleaning code; the deterministic and lag specification; minimum window in observations and fractional form; null simulation specification and seed; global and endpoint critical values; the full BSADF series; exact regime labels and episode rules; and the common return ledger behind every summary. The implementation should reproduce every table and figure from those saved inputs and verify the pooling and supremum identities above.

Appendix B Provenance and editorial scope

This is a reconstructed, revised edition of the 22-page paper dated December 20, 2025. Its baseline is the PDF published as *bitcoin-bubble-gsadf-seminar-paper.pdf*, identical to the archived submission in the author’s seminar-paper folder. The original PDF is preserved unchanged. An editable original manuscript, the empirical dataset, and the executable GSADF analysis were not recovered in the targeted search; the editable manuscript supplied with this edition is a new reconstruction, not a recovered original source.

The revision retains all numerical summaries from the original paper, with the unsupported test statistics recorded in Appendix A. It consolidates duplicated descriptive and critical-value tables, preserves the original price illustrations, corrects reference identities and characterizations, rewrites the statistical exposition, and narrows conclusions to the evidence available. The original statistic-path image remains in the source archive because its relation to the reported global statistic is unresolved; it is not promoted as a validated BSADF plot. No new empirical simulation, regression, or behavioral-proxy analysis is represented as having been performed.

Acknowledgments

The original paper acknowledged OpenAI’s ChatGPT (GPT-5.1) for assistance with wording, formatting, and figure adjustment. This revised edition also used OpenAI’s Codex with GPT-6 Astra for editorial, mathematical, methodological, and reference review, reconstruction of editable files, and document preparation. The revision adds no newly estimated empirical results. Responsibility for the paper remains with its author.