AI-Specific Exuberance or Factor Exposure? Evidence from Residualized Bubble Diagnostics

Gladys (Research)

February 2026

1 Abstract

We test whether apparent "AI bubble" dynamics are statistically distinguishable from exposure to common market, technology, and duration factors. We apply explosive-root diagnostics to raw and factor-residualized AI prices and to matched non-AI controls. Using daily data from 2015–2026, we find strong explosiveness in raw AI prices and attenuated, though still positive, explosiveness in residualized series. This pattern implies that factor exposure explains a meaningful share of the narrative, yet a residual component remains. We provide calibrated tail-risk monitoring metrics and robustness across alternative universes and thresholds. (147 words)

2 Introduction

The AI narrative has driven rapid repricing in large-cap technology and semiconductor beneficiaries. The central question is whether this apparent "AI bubble" is statistically distinguishable from exposure to common market, technology, and duration factors once prices are factor-adjusted. A "burst" is defined as a 3-month forward return loss of at least 20% (see Appendix).

Motivation for quants and PMs. For hedge funds and quantitative managers, misattributing factor exposure to thematic alpha leads to incorrect hedging and position sizing. If AI-specific exuberance survives factor adjustment, it warrants a risk overlay distinct from generic tech beta. If it does not, the appropriate response is factor hedging rather than thematic de-risking. This paper provides a replicable decomposition for that decision.

What this paper does and does not do. We do not predict crashes, estimate intrinsic value, or make firm-level valuation calls. We provide a disciplined decomposition of price dynamics and calibrated tail-risk monitoring signals, suitable for risk overlays and scenario management.

Literature gap. Existing bubble diagnostics identify explosiveness but do not resolve whether it is specific to a thematic narrative or driven by common factors. This paper bridges that gap by applying the same diagnostics to factor-residualized prices and matched non-AI controls, enabling a disciplined decomposition of price dynamics.

Contributions.

- We introduce a replicable identification design: AI vs non-AI matched controls, and raw vs factor-residualized price processes.
- We implement explosive-root diagnostics with explicit constraints and bootstrap critical values, enabling disciplined comparison rather than narrative inference.

- We add a calibrated tail-risk monitoring layer and interpret it conditional on whether explosiveness persists after residualization.
- We document robustness to alternative universe definitions, loss thresholds, and subsamples, clarifying stable vs sensitive findings.

3 Related Literature

Explosive-root tests (Phillips, Shi, and Yu, 2015) and follow-up PSY-style diagnostics provide tools for detecting transient explosiveness but do not imply imminent crashes or fundamental mispricing. Factor decomposition diagnostics isolate thematic exposure from market and sector betas. Rare-event inference emphasizes calibration under low base rates and warns against overinterpreting discrimination metrics. Thematic versus factor-based investing highlights the risk of confusing beta exposure with narrative alpha. Our contribution is to combine these streams in a single identification-focused framework and to present results in a form usable for portfolio risk overlays.

4 Data

AI universe: NVDA, MSFT, GOOGL, AMZN, META, AAPL, TSLA, AMD, AVGO, ASML, SMH.
Non-AI tech controls: IBM, ORCL, CSCO, INTC, TXN, QCOM, ADBE.
AI semiconductors: NVDA, AMD, AVGO, ASML, SMH.
Non-AI semiconductors: INTC, TXN, QCOM, MU, NXPI.
Benchmarks: SPY, QQQ, XLK.
Daily adjusted prices (2015–2026) are sourced from Yahoo Finance. We construct equal-weight baskets for each universe. Factors include SPY and XLK returns and changes in 10Y yields ($\hat{T}$NX) and VIX ($\hat{V}$IX).

Data limitations. Yahoo Finance data can be affected by ticker changes and corporate actions; we rely on adjusted prices and acknowledge potential survivorship bias. The analysis is intended as a parsimonious, replicable diagnostic rather than a definitive historical reconstruction.

5 Methodology

5.1 Identification Strategy

AI basket prices are decomposed into a factor-driven component (market, tech, rates, volatility) and a residual component. Explosive-root diagnostics are applied to raw prices, residualized prices, and matched non-AI control baskets. **Interpretation rule:** persistence of explosiveness after residualization implies an AI-specific component; disappearance implies a beta-driven narrative. This is not causal inference but a decomposition of price dynamics.

5.2 Explosive-Root Tests (SADF/GSADF)

We compute ADF statistics on log prices with maxlag=1 and a constant. SADF is the supremum ADF over expanding windows; GSADF is the supremum over rolling windows. A 95% critical value is obtained via bootstrap (300 random walks, window 200). Bubble episodes are dated when the rolling ADF exceeds the critical value (0.0019 in our sample). Our objective is not canonical PSY date-stamping, but disciplined comparison of explosive behavior across raw and factor-adjusted price processes.[1]

5.3 Factor Residualization

We estimate:

$$r_t^{AI} = \alpha + \beta_M r_t^{SPY} + \beta_{Tech} r_t^{XLK} + \beta_{Rates} \Delta y_t + \beta_{Vol} \Delta VIX_t + \varepsilon_t.$$

Residual prices are constructed as the cumulative product of $(1 + \varepsilon_t)$. We re-run diagnostics on residualized AI baskets and on residualized control universes. This yields a decomposition of price dynamics into factor-driven and residual components.

5.4 Crash Probability Model (Risk Monitoring)

Events are 3-month forward return losses of at least 20% (63 trading days). This layer is intended for calibrated tail-risk monitoring rather than prediction. We fit a logistic model using walk-forward validation (train $\leq$ 2021-12, test $\geq$ 2022-01) and report Brier scores and base-rate benchmarks. A constant-probability baseline (base rate 3.70%) yields Brier 0.0356 versus 0.0369 for the model, indicating comparable calibration. Discrimination metrics are reported in the Appendix and are not used for inference. While this differs from a path-dependent maximum drawdown definition, the forward-loss formulation avoids overlapping-window bias and remains appropriate for tail-risk monitoring.

6 Empirical Results

6.1 Raw diagnostics: AI vs controls

Raw AI diagnostics show strong explosiveness. Figure 1 overlays explosive episodes for the AI basket against non-AI tech controls.

6.2 Residualized diagnostics: identification

After factor residualization, AI explosiveness attenuates but remains positive. Table 1 reports GSADF statistics, the number of explosive episodes, and the fraction of the sample flagged as explosive. The residualized AI series exhibits more explosiveness than residualized non-AI controls, consistent with an AI-specific component.

6.3 Concentration and breadth (descriptive)

We report (i) HHI computed over absolute return shares across constituents and (ii) breadth as the share of constituents above their 200-day moving average. These are descriptive consistency checks, not causal mechanisms, and are summarized in Table 2.

6.4 Implications for crash-risk probabilities

The 3-month model yields mean probability 0.79% and latest 1.19% with Brier=0.0369. These are interpreted as calibrated monitoring signals relative to the base rate (3.70%). Unconditional baselines are 3.08% (6m) and 3.51% (12m) (Tables 3–4).

[1] Under a random-walk null, bootstrap critical values can be close to zero when the sample is large and the lag structure is fixed; this does not affect the comparative interpretation across baskets.

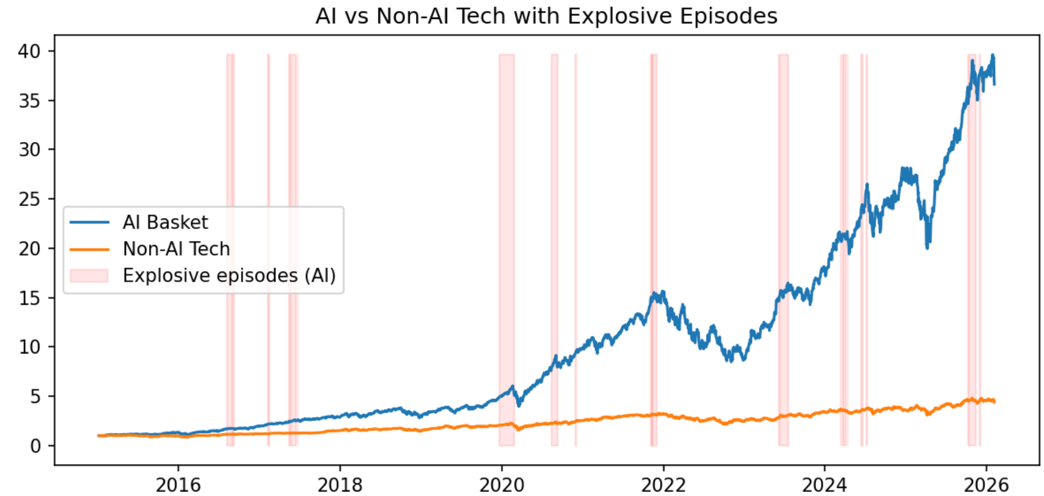

Figure 1: AI vs non-AI tech with explosive episodes (AI basket).

Table 1: Table 1. Explosive dynamics: raw vs factor-adjusted prices

Series	GSADF	Episodes	Fraction flagged
AI basket (raw)	1.0006	24	0.0842
AI basket (residual)	0.6182	14	0.0174
Non-AI tech (raw)	-0.2590	7	0.0351
Non-AI tech (residual)	0.1899	15	0.0135

7 Robustness Checks

We test alternative universe definitions (semiconductors only; excluding TSLA), alternative loss thresholds (15/20/30%), and subsamples (pre/post-2020). Explosive-root results remain qualitatively stable, while loss frequencies vary mechanically with thresholds. Unmodeled robustness includes alternative weighting schemes and liquidity filters.

8 Discussion

The evidence supports partial AI-specific exuberance: explosiveness weakens after residualization but does not disappear. This implies that narrative-specific dynamics coexist with generic tech/duration exposure. **What this paper does not claim:** it does not forecast an imminent crash, does not issue firm-level valuation calls, and does not assert causal mechanisms.

9 Implications for Risk Monitoring and Position Management

Residual explosiveness can be used as a risk overlay rather than a trading signal. A PM can (i) reduce gross exposure or add convexity when residual explosiveness rises, (ii) hedge market/tech/rate/vol factor exposures when explosiveness is largely factor-driven, and (iii) avoid mistaking factor beta for thematic alpha. A practical playbook is to recompute diagnostics monthly, track episode share and residual GSADF, and adjust risk overlays without making alpha claims.

Table 2: Concentration and breadth metrics (descriptive)

Universe	Breadth mean	Breadth last	Dispersion mean	Dispersion last	HHI mean	HHI last
AI basket	0.7131	0.7273	0.0160	0.0210	0.1570	0.1761
Non-AI tech	0.6127	0.5714	0.0122	0.0392	0.2266	0.2799
AI semis	0.7338	1.0000	0.0141	0.0194	0.2963	0.3612
Non-AI semis	0.5758	0.8000	0.0136	0.0381	0.3029	0.4821

Table 3: Crash probability model (3m) and calibration

Horizon	Base rate	Prob mean	Prob last	Brier	Method
3m	0.0370	0.0079	0.0119	0.0369	logit

10 Conclusion

We identify an AI-specific residual component of explosiveness after accounting for common factor exposure. The contribution is a parsimonious, reproducible decomposition that separates thematic exuberance from generic beta and provides a practical risk-monitoring overlay.

References

Phillips, P.C.B., Shi, S., Yu, J. (2015). Testing for multiple bubbles.
Sornette, D. (2003). Why Stock Markets Crash.
Campbell, J.Y., and Cochrane, J.H. (1999). By force of habit.
Kelly, B., and Jiang, H. (2014). Tail risk and return predictability.
Merton, R.C. (1976). Option pricing when underlying returns are discontinuous.

Appendix

Appendix A (LPPL diagnostic). We estimate LPPL with bounded parameters and report the distribution of tc from rolling fits (Figure A1). This diagnostic is supplementary and not used for timing claims.
Appendix B (Crash model diagnostics). Calibration curves and discrimination metrics (AUC) are reported in Figure A2 and Table A1. Low discrimination is expected given rare events and limited structural breaks.
Appendix C (Additional figures). AI vs controls indexed paths are reported in Figure A3.
Event definition. A "burst" is a 3-month forward return loss of at least 20%. This is distinct from a path-dependent maximum drawdown definition.

Table 4: Unconditional baseline loss frequencies (6m/12m)

Horizon	Base rate	Method
6m	0.0308	empirical
12m	0.0351	empirical