Algorithmic Exuberance: Do Algorithmic Markets Amplify Shocks?

Marc Schmitt
University of Oxford
August 2026

Abstract. Financial markets run on interacting trading and information algorithms, raising concerns that common models and concentrated AI infrastructure may synchronize responses and amplify shocks. I develop a theory of Algorithmic Exuberance in which amplification rises with common exposure, speed, information curation, and homogeneity. Using U.S. equities, 1980–2024, I test whether shocks propagate more strongly in high-feedback states. The volatility-implied feedback state is positive and rising, but cross-sectional and Russell reconstitution designs powered to detect 1% effects bound low-frequency amplification below 5% of realised volatility at every horizon. Algorithmic Exuberance is therefore present but historically subcritical, establishing a pre-frontier-AI benchmark.

JEL classification: G12, G14, C58, D83
Keywords: algorithmic markets, market reflexivity, volatility amplification, tail risk, financial stability

Marc Schmitt, University of Oxford, marc.schmitt@cs.ox.ac.uk. I have no relevant or material financial relationships or conflicts of interest that relate to the research described in this paper. The research received no external funding.

I. Introduction

Financial markets are now organized around interacting algorithms.[1] Algorithmic systems generate a substantial share of trading volume and process market-relevant information at machine speed (e.g., Hendershott et al., 2011; Chaboud et al., 2014; Budish et al., 2015). When algorithms trade with each other, price changes become inputs into future decisions, giving markets a recursive structure. Moreover, feedback operates through more than trading. Algorithms shape the information environment by selecting, ranking, and disseminating news, narratives, and sentiment on which other algorithms base their decisions (e.g., Tetlock, 2007; Bybee et al., 2024; Cookson et al., 2024). Trading and information algorithms therefore form a coupled feedback system. This paper studies the dynamics of this feedback.

Whether this transformation stabilizes or destabilizes markets is contested. For example, Hendershott et al. (2011) and Chaboud et al. (2014) document that automation improves liquidity and accelerates price discovery. In contrast, Kirilenko et al. (2017) show interacting algorithms withdrawing liquidity within minutes in the 2010 Flash Crash, and Ben-David et al. (2018) find that index-linked arbitrage transmits shocks into the underlying stocks. As artificial intelligence (AI) spreads through both trading and information systems, these dynamics have become a central financial-stability concern. The FSB, IOSCO, and the IMF warn that common models, shared data, and concentrated AI infrastructure may synchronize trading under stress.[2] Recent theory supports these concerns. Dou et al. (2025) show that AI speculators can learn to collude and degrade price efficiency (see also Cartea et al., 2026), and Yang (2024) shows that algorithmic coordination can seed self-fulfilling crises. Yet algorithmic shock amplification remains an unmeasured mechanism. Official assessments identify channels through which it may arise but provide neither an estimate of the amplification's magnitude nor a historical benchmark against which to assess it. This paper supplies the theory, the estimate, and the benchmark.

I term the resulting regime Algorithmic Exuberance, a market environment in which interactions between trading and information algorithms endogenously amplify shocks. When many systems

[1] Throughout, algorithmic markets are financial markets in which trading, information processing, and decision-making are substantially mediated by algorithmic systems. These systems span the full spectrum of machine-enabled decision technologies, from rule-based execution through machine-learning forecasting (Jordan and Mitchell, 2015; LeCun et al., 2015) to autonomous agentic systems built on modern foundation models (Vaswani et al., 2017). I develop algorithmic markets as a general framework for financial economics in Schmitt (2026).

[2] See, e.g., the Financial Stability Board's 2024 report on the financial-stability implications of artificial intelligence its subsequent work on responsible AI adoption, and the October 2024 *Global Financial Stability Report* chapter on artificial intelligence, which emphasize third-party concentration, market correlations, and the need to enhance supervisory monitoring. Former SEC Chair Gary Gensler has warned that AI-driven trading could trigger the next financial crisis (*The New York Times*, August 7, 2023).

adapt to common data, market outcomes, and one another's responses, individually rational decisions can become synchronized and self-reinforcing (e.g., Calvano et al., 2020; Dou et al., 2025). In the model, amplification arises through two channels of algorithmic reflexivity. Market-algorithmic reflexivity captures trading and forecasting systems adapting to one another's actions and forecasts. Information-algorithmic reflexivity captures algorithmic curation that transforms news, narratives, and sentiment into common trading inputs. Their joint strength is summarized by the system's *feedback gain*. Gain rises with common exposure and reaction speed on the trading side, algorithmic curation on the information side, the coupling between the two layers, and the model homogeneity that raises both at once.

I further provide a microfoundation for model homogeneity and show that the system crosses a stability boundary when feedback gain reaches a critical level. As algorithms converge on the model classes, data, and signals that perform best (e.g., Gu et al., 2020; Kelly et al., 2024; Bryzgalova et al., 2025), a larger share of their responses is common and survives aggregation (Kleinberg and Raghavan, 2021). Historical data measure the mechanism's trading-side footprint, common exposure, while the volatility-implied diagnostic tracks the system's combined feedback state. Algorithmic Exuberance concerns endogenous dynamics in markets shaped by interacting algorithms, not emotion or irrationality.[3] Its central implication is that market fragility can arise even without new fundamental information when trading and information algorithms mutually amplify their signals.

Reflexive amplification is difficult to quantify over long historical samples. I therefore construct a Reflexivity Index (RI), a volatility-implied monitoring diagnostic with two components. The Reflexivity Share of Variance (RSV) maps observed persistence into the share of volatility implied by feedback beyond a low-automation historical benchmark. Implied Feedback (IF) maps the same persistence into the model's combined feedback parameter, $\alpha + \delta\beta$. The index does not identify algorithmic feedback.[4] It translates observed volatility dynamics into the units of the theory. In other words, the index locates the feedback state, and the shock×state battery measures its amplification.

The model's central prediction is that the same shock propagates more strongly when the

[3]The mechanism differs from accounts that emphasize investor psychology, sentiment, and narrative contagion (see, e.g., Shiller, 2017). Human beliefs continue to matter. The point is that amplification can arise from algorithmic interaction alone. Because algorithms learn from human-generated data and signals, behavioral biases and narratives can moreover migrate into algorithmic systems and be propagated by the same feedback, with the 2021 GameStop episode as a case in point.

[4]Unlike event-time measures such as the Hawkes-process branching ratio (Filimonov and Sornette, 2012; Hardiman et al., 2013), which estimate endogenous activity over short windows, the Reflexivity Index summarizes the feedback state from low-frequency volatility persistence over decades. Both rely on the same stability logic, in which endogenous feedback intensifies as the relevant persistence parameter approaches unity.

pre-shock feedback state is elevated. I test this prediction in U.S. equities from 1980 to 2024 using complementary shock×state designs. At the aggregate level, local projections interact market shocks with the lagged feedback state. In the cross-section, the same shocks are interacted with stock-level return commonality, institutional ownership, and index exposure, with time fixed effects absorbing the aggregate regime. A Russell 1000/2000 reconstitution provides quasi-random variation in index-linked ownership around the cutoff (e.g., Chang et al., 2015; Appel et al., 2016; Ben-David et al., 2018). Estimates near zero are informative only when the design could have detected an economically meaningful effect, so I recast the battery as equivalence tests that bound the largest amplification consistent with the data.

These tests place any low-frequency amplification below five percent of unconditional realised volatility at every horizon, and the two best-powered designs detect effects as small as one percent. Elevated volatility in the algorithmic era is concentrated in major crisis episodes (the dot-com unwind, the global financial crisis, and COVID-19) and does not reflect a persistent increase in feedback gain. The 2010 Flash Crash and the 2021 GameStop surge (Kirilenko et al., 2017; Pedersen, 2022) were transient, contained in part by circuit breakers and related safeguards, and left no persistent imprint on low-frequency volatility. Algorithmic Exuberance is therefore not merely a prospective risk but a standing feature of algorithmic markets. The diagnostics indicate a persistently positive feedback state, individual episodes show it flaring locally, and the direct tests establish that its measurable low-frequency channels remained economically small through 2024. The result is a pre-frontier-AI baseline, not a pre-algorithmic one. By 2024, U.S. equities were already shaped by electronic execution, passive ownership, and machine-learning forecasting. What had not yet diffused widely were shared foundation models and autonomous decision systems spanning the trading and information layers. The historical bound therefore measures amplification in the mature algorithmic market and provides the benchmark against which new common decision architectures can be assessed.

The paper makes three contributions. First, it develops a theory of Algorithmic Exuberance and provides a microfoundation for the feedback gain that governs its intensity. Idiosyncratic, model-specific responses diversify away in the cross-section, while common responses survive aggregation and feed back into the market state. Stabilizing, contrarian strategies enter this aggregate with the opposite sign, so the feedback gain, and the amplification it produces, are net quantities. The theory delivers a sharp, testable prediction. The same shock propagates more strongly as the system approaches the stability boundary.

Second, the paper builds a transparent measurement battery for this prediction. Shock×state designs are implemented in the aggregate time series, in the within-date cross-section, and around

the Russell 1000/2000 reconstitution, with the Reflexivity Index serving as the monitoring diagnostic. It measures common exposure, the low-frequency trading route to high gain, and leaves intraday speed, the information layer, and strong model homogeneity to higher-frequency and post-2024 data.

Third, the paper provides the first comprehensive measurement of algorithmic shock amplification. The cross-sectional and Russell designs are powered to detect effects near one percent of realised volatility and place amplification below five percent at every horizon. Russell 2000 assignment raises institutional ownership by 3.1 percentage points without increasing shock propagation, and feedback gain is stable once major volatility episodes are treated symmetrically. These results yield the pre-frontier-AI baseline and a dated, falsifiable forward prediction. If shared foundation models, agentic systems, and common infrastructure raise gain through homogeneity, information curation, speed, and autonomy together, the amplification signature should grow in post-2024 data, and the framework provides concrete monitoring targets.

The paper proceeds as follows. Section II reviews the related literature. Section III develops the model, derives the stability boundary, and defines the monitoring diagnostics. Section IV reports the diagnostics and the amplification tests for U.S. equities from 1980 to 2024. Section V interprets the bound as a pre-frontier-AI baseline and states the forward prediction, and Section VI concludes.

II. Related Literature

Algorithmic Exuberance draws on several strands of research, which this section places in relation to the paper's theory and measurement.

The possibility that prices move in ways not fully explained by fundamental news is long established. Classic work by Black (1986) emphasizes the role of noise in trading and in prices, and models of positive-feedback investment strategies show that some traders buy after price increases and sell after declines, reinforcing directional movements, with rational speculators entering early in anticipation (De Long et al., 1990). Because arbitrage against such strategies is itself risky, prices can deviate from fundamentals for extended periods (Shleifer and Vishny, 1997). Feedback also runs through the information channel. Ozdenoren and Yuan (2008) show how the anticipation of information-driven trading feeds back into prices, and Carlin et al. (2007) document how cooperative liquidity provision can give way to predatory trading and episodic liquidity crises. Related work by Hong and Stein (1999) models newswatchers and momentum traders linked by gradual information diffusion, so that amplification arises from the interaction of distinct trader groups. Algorithmic Exuberance shares this amplification logic and embeds it in a different technological environment. Instead of distinct trader types linked by gradual diffusion, amplification arises from tightly coupled

learning systems trained on overlapping data and updating on common market outcomes, so that individually rational updating aggregates into computational synchronization and collective feedback.

What has changed is the technology of interaction. As automation diffused through execution and market design, algorithmic trading improved several dimensions of market quality, including liquidity and the speed with which public information is incorporated into prices (e.g., Hendershott et al., 2011; Brogaard et al., 2014; Chaboud et al., 2014). At the same time, machine-speed markets exhibit failure modes in which liquidity evaporates and prices move sharply with little apparent fundamental news, as in the 2010 Flash Crash (Kirilenko et al., 2017), and the safeguards introduced in response, such as circuit breakers, have effects on stability that are themselves studied (Chen et al., 2024b). Automation enhances micro-level efficiency, and it can also increase synchronization and strengthen feedback at the system level when many algorithms respond to common signals over short horizons (Budish et al., 2015).

In parallel, machine learning has become a general-purpose forecasting technology in asset pricing, with flexible models used to predict returns and risk (e.g., Gu et al., 2020; Bryzgalova et al., 2025) and the growing abundance of financial data reshaping how prices incorporate information (Dugast and Foucault, 2018; Farboodi and Veldkamp, 2020; Goldstein et al., 2021). Flexible methods discover predictive patterns directly from data (Jiang et al., 2023), and the frontier now extends to reinforcement-learning portfolio construction (Cong et al., 2021) and to the combination of human and machine analysts (Cao et al., 2024), while machine learning applied to price charts revives technical analysis (Murray et al., 2024). A key implication is that, as institutions adopt similar model classes trained on overlapping predictors and updated on shared market outcomes (Kelly et al., 2024), informational diversity may decline (Goldstein and Yang, 2015), raising the scope for synchronized responses and endogenous amplification. This homogenization concern is sharpened by work outside finance showing that algorithmic monoculture can lower welfare even when each shared model is individually better (Kleinberg and Raghavan, 2021), and by evidence that flexible, richly parameterized models trade internal refutability for out-of-sample fragility (Chen et al., 2024a).

Beyond trading, algorithmic systems shape the information environment itself, selecting, ranking, and disseminating narratives that enter models as tradable signals (e.g., Tetlock, 2007; Barber and Odean, 2008; Cookson et al., 2024). A growing body of work shows that news volume and tone, and the diffusion of attention across social networks, move risk premia and volatility (e.g., Bybee et al., 2024; Aleti and Bollerslev, 2025; Hirshleifer et al., 2025). This information layer introduces an additional feedback channel that classic models of feedback trading do not endogenize, and it is the

second layer of this paper's model.

A growing theoretical and simulation literature now examines how interacting learning algorithms reshape markets directly. Dou et al. (2025) show that AI-powered speculators can learn collusive trading strategies that lower price efficiency, whether through price-trigger punishment or through homogenized learning biases, and Colliard et al. (2026) find that reinforcement-learning market makers learn to quote supra-competitive spreads even without coordination (see also Cont and Xiong, 2024). Related work documents reinforcement-learning sellers that sustain supra-competitive prices without communication (Calvano et al., 2020; Hansen et al., 2021), co-learning externalities that bias equilibrium portfolios (Gufler et al., 2025), AI planning agents that manufacture and exploit bubbles when feedback trading is strong (Dou et al., 2026), and AI coordination that can seed self-fulfilling crises (Yang, 2024). These studies establish, in stylized environments, that algorithmic interaction generates emergent market outcomes that individual optimization does not predict. This literature centers on the trading layer, in which algorithms learn from one another's orders. Algorithmic Exuberance adds the information layer and its coupling to trading, so that the homogeneity of the algorithms on each side and the strength of the coupling between them govern whether the system amplifies shocks. The present paper supplies the empirical counterpart, a measurement over four decades of U.S. equity data of the shock-amplification signature such interaction leaves.

Finally, the paper connects to research on systemic fragility and financial stability, in which shocks are classically amplified through funding and liquidity spirals (Brunnermeier and Pedersen, 2009) and propagated through network interconnections that absorb small shocks yet amplify large ones (Acemoglu et al., 2015). Recent syntheses catalog the mechanisms through which modern markets become fragile (Goldstein et al., 2025), and a resilience perspective evaluates markets by their capacity to absorb shocks (Brunnermeier, 2024). Work linking model simplicity and homogeneity to systemic risk (Puri, 2025) is close in spirit to this paper's homogeneity route. The measurement in this paper contributes a magnitude to this literature for one specific channel, the low-frequency amplification produced by algorithmic common exposure.

III. A Model of Algorithmic Exuberance

Figure 1 illustrates the architecture underlying Algorithmic Exuberance. Information algorithms select, rank, and amplify news, sentiment, and narratives, shaping the information environment from which market algorithms construct predictive signals. Market algorithms process these signals to form forecasts and trading decisions. The resulting market volatility state feeds back into both layers through continuous learning and amplification. When information algorithms

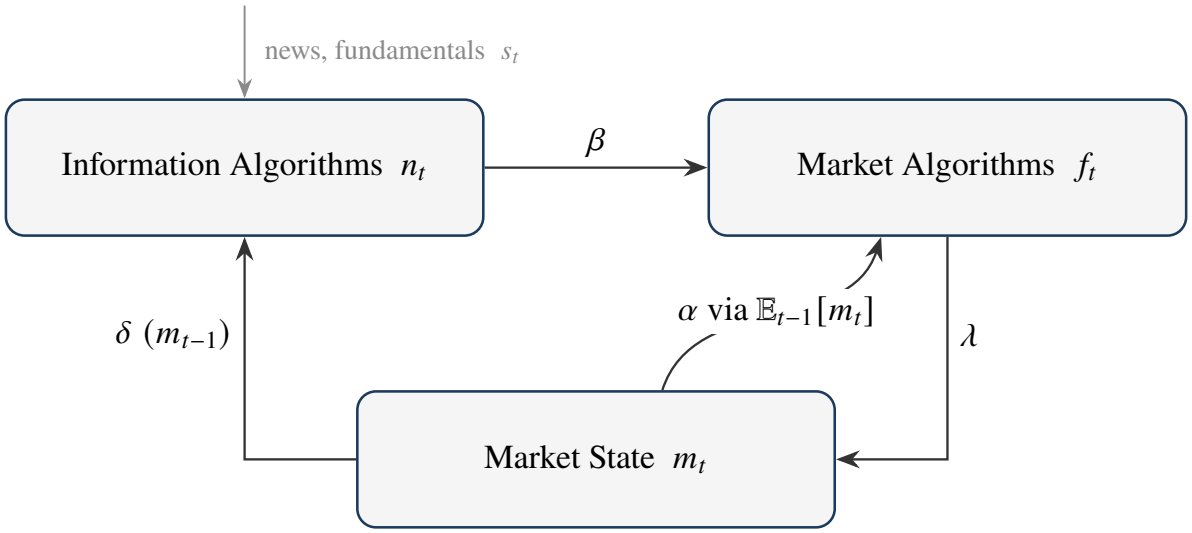


Figure 1: Dual-Channel Algorithmic Reflexivity. The system consists of two interacting layers, Information Algorithms (n_t) and Market Algorithms (f_t), which jointly shape the market state (m_t). Information algorithms process market signals and external inputs such as news, sentiment, and narratives, forming the information-reflexivity channel (δ). Market algorithms translate informational inputs into trading and forecasting behavior, forming the market-reflexivity channel (α). The coupling strength between the two layers is denoted by β, while λ represents the effective speed or weight with which the market volatility state adjusts toward algorithmic forecasts. When both layers reinforce one another, reflexive amplification can produce Algorithmic Exuberance, a market environment in which interactions between trading and information algorithms endogenously amplify shocks.

echo recent market conditions and trading or forecasting algorithms respond to those signals, the interaction becomes self-reinforcing. The strength of this coupling determines whether shocks are absorbed or amplified. This dynamic is captured in the formal model by a single stability boundary.

The signals entering these systems include price-based patterns such as momentum, measures of order-flow imbalance or liquidity, sentiment scores extracted from news or social media, and statistical forecasts of volatility or risk. In the context of the model, the informational variable n_t summarizes the intensity of algorithmically amplified information entering these forecasting systems. Market algorithms transform this informational input into forecasts and trading responses, which in turn shape the market volatility state m_t.

The informational variable n_t represents the information layer, summarizing the algorithmic selection, ranking, and amplification of news, sentiment, and narratives. Market algorithms f_t

form the trading and forecasting layer that uses this informational input to produce forecasts and trading decisions. The two layers are linked through the coupling parameter β, while the market volatility state m_t closes the loop by feeding back into both layers. The parameter δ measures how strongly the information layer reacts to lagged market conditions (m_{t-1}), capturing backward-looking narrative amplification, whereas α quantifies how strongly trading and forecasting algorithms act on expectations of the current volatility state via $\mathbb{E}_{t-1}[m_t]$, capturing forward-looking market reflexivity. The reflexivity index

$$\phi = (1 - \lambda) + \lambda(\alpha + \delta\beta) \tag{1}$$

aggregates these effects into a single measure of feedback intensity. Stability requires $\alpha + \delta\beta < 1$, ensuring that shocks are absorbed rather than self-amplified.

A. Model Intuition

Evidence shows that algorithmic trading can sustain and amplify mispricing without behavioral biases. Zhang and Zhang (2024) document that algorithmic participation, while improving execution efficiency, can increase short-horizon mispricings through correlated strategies and rapid feedback between order flow and price adjustments, so that instability emerges endogenously from rational optimization itself. This is the logic of Algorithmic Exuberance. When many systems rely on overlapping data and architectures, individually rational actions synchronize. Information continues to be incorporated rapidly (Hendershott et al., 2011; O'Hara, 2015), yet volatility becomes more persistent and small shocks propagate more strongly even in the absence of new fundamentals (Kirilenko et al., 2017). This coexistence of rapid information incorporation and endogenous instability motivates the formal model.

B. Setup and Timing

The model formalizes two interacting layers, an information layer in which algorithms process and amplify information relevant for trading, and a trading/forecasting layer in which that information is mapped into forecasts and trading decisions. Timing is as follows. At date t, the information signal n_t is formed using the exogenous shock s_t and the lagged volatility state m_{t-1}. Trading algorithms then form the forecast f_t, and the volatility state m_t is then updated as a weighted average of inertia and the algorithmic forecast.

Time is discrete, $t = 0, 1, 2, \ldots$. Let m_t denote the (log) market volatility state, s_t an exogenous fundamental volatility shock, n_t the algorithmically amplified information or sentiment signal, and f_t the one-step-ahead forecast of the volatility state formed by trading algorithms.

Information layer.

$$n_t = \eta s_t + \delta m_{t-1} + \xi_t, \tag{2}$$

where $\eta \geq 0$ measures sensitivity to fundamentals, $\delta \geq 0$ captures information reflexivity, and ξ_t is an idiosyncratic information shock. The signal n_t summarizes the equilibrium output of interacting curation and diffusion algorithms.

Trading layer.

$$f_t = \alpha \mathbb{E}_{t-1}[m_t] + \beta n_t + \varepsilon_t, \tag{3}$$

where $\alpha \geq 0$ captures market reflexivity, $\beta \geq 0$ is the sensitivity to informational inputs n_t, and ε_t is a forecast innovation.

Volatility-state formation.

$$m_t = (1 - \lambda) m_{t-1} + \lambda f_t, \qquad \lambda \in (0, 1], \tag{4}$$

where λ denotes the state-adjustment speed, reflecting how strongly the volatility state responds to algorithmic forecasts relative to inertia. When λ is small, the volatility state evolves slowly and is largely driven by past conditions, whereas larger values imply that algorithmic forecasts exert stronger influence on market dynamics.

This timing structure generates a closed feedback loop. Information algorithms respond to lagged market conditions through m_{t-1}, trading algorithms form forecasts about the current volatility state through $\mathbb{E}_{t-1}[m_t]$, and the realized state m_t feeds into the next period's information environment.

C. Reduced Form and Reflexivity

Under rational internal consistency,[5] $\mathbb{E}_{t-1}[m_t] = m_{t-1}$, substitute (2) and (3) into (4):

$$\begin{aligned} m_t &= (1 - \lambda) m_{t-1} + \lambda \Big[\alpha m_{t-1} + \beta(\eta s_t + \delta m_{t-1} + \xi_t) + \varepsilon_t \Big] \\ &= \underbrace{\Big[(1 - \lambda) + \lambda(\alpha + \delta\beta) \Big]}_{\phi} m_{t-1} + \lambda\beta\eta\, s_t + \lambda\beta\, \xi_t + \lambda\, \varepsilon_t. \end{aligned} \tag{5}$$

[5] For tractability, the model assumes $\mathbb{E}_{t-1}[m_t] = m_{t-1}$, corresponding to fixed-point internalization under one-step linear learning with public m_{t-1}. The qualitative results extend to richer linear prediction rules.

The coefficient on m_{t-1} in (5) defines the *reflexivity parameter*,

$$\phi \equiv (1-\lambda) + \lambda(\alpha + \delta\beta). \tag{6}$$

This parameter aggregates market reflexivity (α) and information reflexivity (δ, scaled by β) into a single persistence parameter. Larger values of α and δ increase ϕ, strengthen the persistence of the volatility-state process, and move the system closer to the stability boundary. Economically, ϕ captures the overall intensity of feedback between trading and information algorithms. In the empirical analysis, I map this persistence parameter to an observable proxy, *implied feedback* (IF), constructed from the persistence of $\log(RV_t)$.

Under these conditions, the shocks driving the system, $(s_t, \xi_t, \varepsilon_t)$, are assumed to be zero-mean, serially uncorrelated, and mutually independent with variances σ_s^2, σ_ξ^2, and σ_ε^2, respectively.

D. Stability and Volatility Amplification

With volatility-state dynamics $m_t = \phi m_{t-1} + u_t$ induced by the dual-channel model with information n_t, forecast f_t, and state formation m_t, the process is covariance stationary if and only if

$$|\phi| < 1 \quad \Longleftrightarrow \quad \alpha + \delta\beta < 1.$$

Under stationarity,

$$\mathrm{Var}(m_t) = \frac{\lambda^2\left(\beta^2\eta^2\sigma_s^2 + \beta^2\sigma_\xi^2 + \sigma_\varepsilon^2\right)}{1-\phi^2},$$

which is strictly increasing in ϕ and thus in both α and δ.[6]

These expressions characterize the stability properties of the volatility state. The system is stationary only if the combined feedback strength $\alpha + \delta\beta$ remains below one. This condition places a bound on the joint amplification generated by market algorithms (α) and information algorithms (δ, scaled by β). When feedback between the two layers becomes too strong, the persistence parameter ϕ approaches unity and the system moves toward instability.

Even within the stationary region, higher values of ϕ amplify volatility. The unconditional variance of the volatility state increases monotonically in ϕ, implying that stronger algorithmic feedback leads to larger fluctuations in market volatility. In economic terms, reflexivity does not merely increase persistence. It also magnifies the scale of volatility fluctuations generated by the

[6]Equation (5) is an AR(1) process with coefficient ϕ and composite innovation $u_t = \lambda(\beta\eta s_t + \beta\xi_t + \varepsilon_t)$. Stationarity requires $|\phi| < 1$, yielding the standard AR(1) variance formula, and monotonicity follows from $\partial\mathrm{Var}(m_t)/\partial\phi = 2\phi\,\mathrm{Var}(u_t)/(1-\phi^2)^2 > 0$ for $0 < \phi < 1$.

system.

To quantify the variance amplification implied by the model, let ϕ_0 denote a benchmark (low-reflexivity) value of ϕ. In the no-reflexivity limit ($\alpha = \delta = 0$), $\phi_0 = 1 - \lambda$. Because $\mathrm{Var}(m_t) \propto (1 - \phi^2)^{-1}$ for an AR(1) process, the Reflexivity Share of Variance is defined as

$$\mathrm{RSV}(\phi; \phi_0) \equiv 1 - \frac{1 - \phi^2}{1 - \phi_0^2} = \frac{\phi^2 - \phi_0^2}{1 - \phi_0^2}, \qquad \text{for } |\phi|, |\phi_0| < 1. \tag{7}$$

RSV measures the fraction of total variance in the volatility state attributable to reflexive amplification beyond baseline anchoring. The measure is monotone in ϕ and equals zero when $\phi = \phi_0$.

As either reflexivity channel intensifies ($\alpha \uparrow$ or $\delta \uparrow$), ϕ rises, increasing persistence and amplifying volatility nonlinearly through the $(1 - \phi^2)^{-1}$ term. Algorithmic Exuberance becomes progressively more consequential as feedback gain approaches the stability boundary. As $\alpha + \delta\beta \to 1^-$, small shocks generate large and persistent volatility bursts, the extreme expression of the regime. If $\alpha + \delta\beta \geq 1$, the system breaches the stability boundary and becomes nonstationary. Shocks are then self-amplifying rather than mean-reverting.

Algorithmic Exuberance arises from rational feedback amplification, and not from irrational behavior, consistent with evidence that algorithmic trading can amplify short-horizon mispricing through endogenous market dynamics (see Zhang and Zhang, 2024). When many systems respond to correlated signals, their collective actions reinforce price movements rather than offset them. Synchronized de-risking and faster-than-human exposure adjustment feed the trading channel (α), while algorithms reacting to trending news and sentiment feed the information channel (δ).

In this framework, volatility clustering arises from persistence in the volatility state. Define the composite innovation $u_t = \lambda(\beta\eta s_t + \beta\xi_t + \varepsilon_t)$ and rewrite the reduced form as

$$m_t - \phi\, m_{t-1} = u_t. \tag{8}$$

With homoskedastic innovations and constant ϕ, the one-step conditional variance is constant:

$$\mathrm{Var}(m_t \mid \mathcal{F}_{t-1}) = \mathrm{Var}(u_t). \tag{9}$$

Empirically, however, volatility clustering is state-dependent. Feedback strength varies with market conditions (i.e., ϕ becomes time-varying through α_t and δ_t), the innovation variance becomes state-dependent (e.g., ARCH or HAR-RV dynamics; Engle, 1982; Corsi, 2009), or both. In those cases, persistence in ϕ_t transmits directly to persistence in $\log(RV_t)$, generating the bursty dynamics

documented empirically. Economically, a higher ϕ steepens the unconditional-variance multiplier $(1-\phi^2)^{-1}$, so episodes with elevated ϕ_t display both higher average volatility and greater persistence.

Economic intuition. Periods of elevated feedback should exhibit higher and more persistent realized volatility and fatter return tails. Under a standard volatility-scaling observation equation for daily returns, $r_{d,t} = \sigma_t \varepsilon_{d,t}$ with $\sigma_t \propto \exp(m_t/2)$, the greater dispersion of the volatility state near the boundary translates into thicker unconditional return tails. The empirical analysis measures the first implication directly, and the second appears as a descriptive pattern in the data.

E. Routes to High Feedback Gain

The stability boundary is a statement about a single object, the system's feedback gain $\phi_t = (1-\lambda) + \lambda(\alpha_t + \delta_t \beta)$. Amplification intensifies as gain approaches the boundary, and any force that raises the trading-layer elasticity α_t or the information-layer elasticity δ_t raises gain. The mechanism is the composite gain, not any single force. The routes to high gain map onto the two layers of the model and their coupling, and they differ along two dimensions that govern what the historical record can say. The frequency at which each operates determines whether it is measurable in this sample, and the era in which each rose determines whether it belongs to the pre-frontier-AI baseline or to the forward prediction.

Trading-layer routes (raising α). Two forces synchronize trades. *Common exposure* arises when algorithms with overlapping holdings respond to the same signals or are forced into the same mechanical rebalancing, so a common component of their trading survives aggregation. It is observable at low frequency through return commonality, common institutional ownership, and index membership, and it rose over the sample. *Reaction speed* raises gain when algorithms react quickly relative to the rate at which liquidity and arbitrage capital replenish, so a shock loops through the system several times before stabilizing forces arrive. Its secular rise across the sample is reflected in the monthly data and is addressed directly by the feedback-gain test of Section IV.E. The intraday loop dynamics themselves, the 2010 Flash Crash being the canonical instance, operate below the monthly horizon and are left to higher-frequency data.

Information-layer route (raising δ). Algorithmic curation, ranking, and recommendation synchronize and amplify the signals that trading systems condition on, so narratives and sentiment enter many trading models as a common input before any trade is placed. This is the information layer of the model, δ, and it operates as signal synchronization among machines, distinct from human sentiment or emotion. A clean per-period measure of algorithmic narrative amplification over 1980 to 2024 is not available, and large-scale algorithmic curation is itself largely a post-2010

development, so this route enters the historical analysis only in combination, through its contribution to the implied-feedback diagnostic (Section III.F), and stands with model homogeneity among the forward channels.

Cross-layer coupling. Gain rises with the strength of the link between the layers, the magnitude of β that carries information feedback into trading and trading outcomes back into the information environment, even when the agents on each side are diverse. Coupling is structural and cross-cutting, and the aggregate feedback state reflects it.

Model homogeneity (the AI-era channel). Homogeneity is distinct from the routes above because it raises both layers at once. When trading and information algorithms share model classes, training data, and signals, the common component grows on the trading side (α, through ρ_M) and on the information side (δ, through ρ_I) together (Section III.F). Shared foundation models and converging architectures are exactly such a force, which is why homogeneity is the dominant forward concern for the AI era. Over the historical sample its trading-side footprint is visible through common exposure. The homogeneity primitives ρ_M and ρ_I themselves are left latent.

These routes separate along two axes the empirical analysis keeps distinct. By *measurability*, the low-frequency trading route, common exposure, is identified at the monthly horizon, while the intraday speed and coupling dynamics and the information-layer route lie outside the standard market data this sample uses. By *era*, common exposure and the secular automation of trading rose across 1980 to 2024 and form the measured record, while strong model homogeneity and large-scale algorithmic information curation are primarily forward forces that had not meaningfully risen by the end of the sample. The empirical analysis is organized accordingly. It measures the common-exposure route directly and bounds its amplification, reads the combined feedback state from the diagnostics, and names the speed, information-layer, and strong-homogeneity routes as the frontier the AI era will test.

F. The Homogeneity Microfoundation

The homogeneity primitive of Section III.E is the structural force behind the common component on both the trading and information layers, and this subsection develops it in detail. Having derived the stability boundary in terms of the aggregate feedback parameters α and δ, I now provide a microfoundation for those parameters that explains why the elasticities rise as trading and information systems become more homogeneous. The reduced form treats α and δ as elasticities of the aggregate market–information system. This subsection provides a microfoundation for those elasticities and identifies their structural determinant: *model homogeneity*, the extent to which trading and information algorithms use common model classes, overlapping data, and similar state

variables. When homogeneity is low, algorithm-specific responses are diversified in the cross-section and largely wash out in aggregation. When homogeneity is high, responses contain a common component that survives aggregation. The result is an aggregate feedback parameter that rises with model homogeneity even though each algorithm is individually optimizing and no behavioral bias is required. Aggregate reflexivity is therefore endogenous to the ecology of models: it rises with the commonality of the algorithmic systems that process market and information signals, climbing as that ecology becomes more uniform.

Definition 1 (Model homogeneity) *Model homogeneity is the share of an algorithmic signal or forecasting rule that is common across systems. Let* $\rho_M \in [0,1]$ *denote homogeneity among market algorithms and* $\rho_I \in [0,1]$ *denote homogeneity among information algorithms. A higher value means that a larger fraction of the relevant algorithmic input is common rather than idiosyncratic.*

Trading algorithms. There is a unit mass of trading and forecasting algorithms indexed by $i \in [0,1]$. At date t, algorithm i forms a forecast or trading signal $f_{i,t}$ using the amplified information signal n_t and a market-state predictor $x^M_{i,t}$ that combines a common model component and an idiosyncratic component:

$$x^M_{i,t} = \rho_M \, E_{t-1}[m_t] + (1-\rho_M)\, \tilde{x}^M_{i,t}, \qquad \int_0^1 \tilde{x}^M_{i,t}\, di = 0. \tag{10}$$

The common component reflects the use of similar state variables, risk models, momentum or reversal features, volatility forecasts, order-flow indicators, and machine-learning architectures across systems. The idiosyncratic component reflects model-specific signals and implementation choices.[7] Each algorithm is rational but restricted to a linear forecasting technology, choosing coefficients that minimize expected squared forecast error within its model class:

$$f_{i,t} = \alpha_0 \, E_{t-1}[m_t] + \alpha_1 x^M_{i,t} + \beta n_t + \varepsilon_{i,t}, \qquad \alpha_0, \alpha_1, \beta \geq 0, \tag{11}$$

where α_0 is the baseline response to the public market state, α_1 the additional response through the algorithm-specific predictor, β the sensitivity to the information signal, and $\varepsilon_{i,t} = \varepsilon_t + \tilde{\varepsilon}_{i,t}$ with

[7]The zero cross-sectional mean of $\tilde{x}^M_{i,t}$ is a normalization. Any common mean is absorbed into the common component. Equivalently, if idiosyncratic components are independent across algorithms, the integral vanishes by the exact law of large numbers for a continuum (Sun, 2006).

$\int_0^1 \tilde{\varepsilon}_{i,t}\, di = 0$. The aggregate trading forecast $f_t = \int_0^1 f_{i,t}\, di$ is then

$$f_t = \underbrace{\left(\alpha_0 + \alpha_1 \rho_M\right)}_{\alpha(\rho_M)} E_{t-1}[m_t] + \beta n_t + \varepsilon_t, \tag{12}$$

so the aggregate market-reflexivity coefficient is

$$\alpha(\rho_M) = \alpha_0 + \alpha_1 \rho_M, \qquad \alpha'(\rho_M) = \alpha_1 \geq 0. \tag{13}$$

Idiosyncratic model components diversify away. The common component remains in the aggregate forecast and feeds back into the market state. The baseline term α_0 permits feedback even under full diversity ($\rho_M = 0$), all algorithms still observe and respond to the public price, while homogeneity adds the amplification term $\alpha_1 \rho_M$.

Information algorithms. A parallel logic applies to the information layer. A unit mass of information algorithms indexed by $j \in [0, 1]$ select, rank, summarize, and disseminate news, sentiment, and narratives. Algorithm j produces

$$n_{j,t} = \eta s_t + \delta_0 m_{t-1} + \delta_1 x^I_{j,t} + \xi_{j,t}, \qquad \eta, \delta_0, \delta_1 \geq 0, \tag{14}$$

where s_t is an exogenous fundamental shock and the information-side predictor is

$$x^I_{j,t} = \rho_I\, m_{t-1} + (1 - \rho_I)\, \tilde{x}^I_{j,t}, \qquad \int_0^1 \tilde{x}^I_{j,t}\, dj = 0. \tag{15}$$

The common component reflects shared ranking systems, common training corpora, and similar sentiment models. The idiosyncratic component reflects source-specific coverage and interpretation. With $\xi_{j,t} = \xi_t + \tilde{\xi}_{j,t}$ and $\int_0^1 \tilde{\xi}_{j,t}\, dj = 0$, the aggregate information signal $n_t = \int_0^1 n_{j,t}\, dj$ is

$$n_t = \eta s_t + \underbrace{\left(\delta_0 + \delta_1 \rho_I\right)}_{\delta(\rho_I)} m_{t-1} + \xi_t, \tag{16}$$

so the aggregate information-reflexivity coefficient is

$$\delta(\rho_I) = \delta_0 + \delta_1 \rho_I, \qquad \delta'(\rho_I) = \delta_1 \geq 0. \tag{17}$$

When information systems rank and retransmit market conditions through common narratives or sentiment models, the market state is echoed back into the information environment and re-enters trading models through n_t.

Proposition 1 (Aggregation generates the dual-channel reduced form) *Suppose the idiosyncratic components in* (10) *and* (15) *have zero cross-sectional mean and algorithms follow the linear rules* (11) *and* (14)*. Then the aggregate layers are*

$$f_t = \alpha(\rho_M)\, E_{t-1}[m_t] + \beta n_t + \varepsilon_t, \qquad n_t = \eta s_t + \delta(\rho_I)\, m_{t-1} + \xi_t,$$

with $\alpha(\rho_M) = \alpha_0 + \alpha_1 \rho_M$ *and* $\delta(\rho_I) = \delta_0 + \delta_1 \rho_I$*, both weakly increasing in the corresponding homogeneity parameter.*

Proof. Integrate (11) over i using $\int_0^1 \tilde{x}^M_{i,t}\, di = 0$ and $\int_0^1 \tilde{\varepsilon}_{i,t}\, di = 0$; integrate (14) over j using $\int_0^1 \tilde{x}^I_{j,t}\, dj = 0$ and $\int_0^1 \tilde{\xi}_{j,t}\, dj = 0$. The comparative statics follow from (13) and (17). □

The reduced-form coefficients α and δ are thus equilibrium aggregation objects rather than behavioral parameters. They arise because individually optimal algorithms use overlapping model components that do not diversify away in the aggregate.[8] The market-level externality is that each algorithm treats the market state and information environment as inputs to its own forecast, while the aggregate of such forecasts moves the market state and information environment themselves.

Combining Proposition 1 with the volatility-state equation $m_t = (1 - \lambda)m_{t-1} + \lambda f_t$ and the internal-consistency assumption $E_{t-1}[m_t] = m_{t-1}$ reproduces the baseline reduced form, with reflexivity parameter

$$\phi(\rho_M, \rho_I) = (1 - \lambda) + \lambda\big(\alpha(\rho_M) + \beta\, \delta(\rho_I)\big). \tag{18}$$

More generally, if $E_{t-1}[m_t] = \psi\, m_{t-1}$ for some $\psi \in [0, 1]$, all results below hold with $\alpha(\rho_M)$ replaced by $\psi\, \alpha(\rho_M)$. I set $\psi = 1$ to keep the empirical mapping transparent.

Proposition 2 (Homogeneity moves the system toward the stability boundary) *With* $\alpha_1, \delta_1, \beta \geq 0$ *and* $\lambda \in (0, 1]$*,*

$$\frac{\partial \phi}{\partial \rho_M} = \lambda \alpha_1 \geq 0, \qquad \frac{\partial \phi}{\partial \rho_I} = \lambda \beta \delta_1 \geq 0,$$

[8] This is a *reduced* microfoundation. Optimizing algorithms and aggregation deliver the reduced-form reflexivity parameters and their dependence on homogeneity, while the volatility state is modeled directly rather than derived from market clearing in a fully specified asset-pricing equilibrium. A price-based equilibrium in which strategy correlation raises price impact, with volatility diverging as $\rho_M \to \rho^*_M$, is a natural extension.

and the volatility-state process is covariance stationary if and only if

$$\alpha(\rho_M) + \beta\,\delta(\rho_I) < 1. \tag{19}$$

Higher homogeneity among trading or information algorithms moves the system closer to the stability boundary.

Proof. Differentiate (18). Since $\lambda \in (0, 1]$, $\phi < 1 \iff$ (19); the non-negativity restrictions give $\phi > -1$, so covariance stationarity is equivalent to $\phi < 1$. □

Setting $\rho_M = \rho_I = \rho$ collapses commonality on both layers into a single homogeneity index:

$$\phi(\rho) = (1 - \lambda) + \lambda\big[\alpha_0 + \beta\delta_0 + (\alpha_1 + \beta\delta_1)\rho\big]. \tag{20}$$

If $\alpha_1 + \beta\delta_1 > 0$, the critical level of model homogeneity is

$$\rho^* = \frac{1 - \alpha_0 - \beta\delta_0}{\alpha_1 + \beta\delta_1}. \tag{21}$$

When $\rho^* \in (0, 1)$, the market is stable for $\rho < \rho^*$ and breaches the stability boundary for $\rho \geq \rho^*$. This critical-homogeneity representation connects the stability boundary to observable market-structure primitives, common data sources, common machine-learning architectures, common vendor models and risk controls, and concentration in information or execution technologies.

Corollary 1 (Homogeneity amplifies volatility) *Let $u_t = \lambda(\beta\eta s_t + \beta\xi_t + \varepsilon_t)$ be serially uncorrelated with variance σ_u^2. If $0 < \phi(\rho_M, \rho_I) < 1$, then $\mathrm{Var}(m_t) = \sigma_u^2/(1 - \phi^2)$, and*

$$\frac{\partial\,\mathrm{Var}(m_t)}{\partial\rho_M} = \frac{2\phi\sigma_u^2}{(1 - \phi^2)^2}\,\lambda\alpha_1 \geq 0, \qquad \frac{\partial\,\mathrm{Var}(m_t)}{\partial\rho_I} = \frac{2\phi\sigma_u^2}{(1 - \phi^2)^2}\,\lambda\beta\delta_1 \geq 0.$$

Model homogeneity raises volatility amplification even when the variance of primitive shocks is unchanged. Common algorithmic responses make a given shock propagate more strongly through the market–information system. This is the force the Reflexivity Index measures.

What the index identifies. The empirical measure maps observed volatility persistence into the composite feedback term

$$\alpha(\rho_M) + \beta\,\delta(\rho_I) = \alpha_0 + \alpha_1\rho_M + \beta(\delta_0 + \delta_1\rho_I). \tag{22}$$

The Reflexivity Index is therefore a composite diagnostic statistic for aggregate feedback intensity, *not* a decomposition of the trading- and information-side channels. Separate identification of ρ_M and ρ_I would require direct data on trading-system commonality, message traffic, order-flow synchronization, information-vendor concentration, or sentiment-model commonality. The present implementation instead measures the combined feedback state over a long historical sample using market data alone. The information-layer route therefore enters the historical analysis only through this composite, contributing to measured implied feedback without being separately identified. Isolating it is a task for higher-frequency and information-side data.

Stabilizing algorithms and net gain. Not all algorithmic strategies amplify. Contrarian and mean-reversion systems trade against recent price movements, and their responses enter the aggregate forecast with the opposite sign, offsetting part of the common component that survives aggregation and lowering the aggregate elasticity α. The feedback gain the model summarizes, and the amplification the battery measures, are therefore net quantities, the balance of amplifying and stabilizing algorithmic responses. The net is the object that matters for stability, and it sharpens the reading of the historical bound. Low measured amplification through 2024 is consistent with substantial gross algorithmic activity on both sides of the balance. The forward concern is a shift in that balance, if homogeneous models and shared signals raise the correlation of responses faster than contrarian capital expands.

Homogeneity and the secular drift. This microfoundation reframes the empirical patterns documented below. The secular rise in implied feedback can arise without any change in the speed or fundamentals of trading. It is consistent with a rise in model homogeneity as institutions adopted common machine-learning model classes trained on overlapping data, a mechanism related to work on model simplicity and fragility (Puri, 2025) and to simulation evidence that shared model architectures generate correlated trading (Lopez-Lira, 2025). The microfoundation also rationalizes the quasi-experimental evidence that removing heterogeneous human intermediaries raises synchronization and volatility (Brogaard et al., 2025). Suspending floor trading lowers cross-sectional diversity and raises effective ρ_M. Conversely, governance that preserves diversity or interrupts synchronized feedback lowers effective ρ_M and ρ_I, consistent with the interpretation of the post-2010 moderation in Section V.

Implication for moderate conditional exceedances. The same logic clarifies why the model speaks to the clustering of *moderate* shocks rather than to the forecasting of rare disasters. Suppose

daily returns within a month satisfy $r_{d,t} = \sigma_t z_{d,t}$ with $\log \sigma_t^2 = c + m_t$. A higher $\phi(\rho_M, \rho_I)$ makes shocks to m_t more persistent, so, conditional on an ex-ante volatility forecast, a high-reflexivity state should generate more repeated volatility-scaled exceedances of moderate thresholds. Qualitatively, the model therefore implies

$$\frac{\partial}{\partial \phi_t} E_t\big[\log(1 + \mathrm{Exc}_{k,t+1})\big] > 0 \quad \text{for moderate } k, \qquad \mathrm{Exc}_{k,t+1} = \sum_{d \in t+1} \mathbf{1}\{|r_d| > k\,\widehat{\sigma}_{t+1|t}\}, \quad (23)$$

with a weaker effect at very high thresholds, where rare far-tail events may depend on crisis-scale fundamentals, jumps, or liquidity shocks outside the linear feedback mechanism. This is a qualitative, directional implication of higher persistence, and it motivates the conditional-exceedance test reported in the Internet Appendix. As a matter of theory, Algorithmic Exuberance speaks to the clustering and propagation of moderate shocks. It does not require that the index forecast every far-tail realization.

G. Testable Predictions

The model links algorithmic feedback to volatility dynamics through the composite feedback parameter, or gain,

$$\phi_t = (1 - \lambda) + \lambda(\alpha_t + \beta\delta_t).$$

Gain rises through several routes (Section III.E). The microfoundation develops the homogeneity primitive, showing that the aggregate elasticities α_t and δ_t increase with model homogeneity in the trading and information layers. Its trading-side footprint, common exposure, is what the historical data measure. Because ρ_M and ρ_I are not directly observed over the full historical sample, the empirical analysis focuses on the volatility-implied feedback state summarized by the Reflexivity Index. This yields two implications, a descriptive one that motivates the diagnostics and a sharp one that the amplification battery measures directly.

1. **Rising feedback diagnostics.** Higher model homogeneity raises the aggregate elasticities $\alpha(\rho_M)$ and $\delta(\rho_I)$, and therefore the composite gain ϕ. Because the diagnostics are constructed from volatility persistence, this implies a measurable upward drift over the algorithmic era, more persistent realized volatility, a higher Reflexivity Share of Variance, and higher Implied Feedback. These translate persistence into the units of the theory and locate where amplification, if present, should appear. They motivate the direct test without establishing it.

2. **State-dependent shock amplification.** The central prediction is comparative. The *same*

shock propagates more strongly when the pre-shock feedback state is elevated, and more strongly still as gain approaches the stability boundary. It concerns a magnitude that a volatility forecast does not reproduce mechanically, so it can be taken to the data directly and either detected or bounded. The microfoundation makes it route-specific. Amplification should rise with a stock's common exposure, its loading on the common, non-diversifying component of algorithmic trading, which is observable at low frequency through return commonality, institutional ownership, and index membership. The empirical core estimates this shock×state interaction at the aggregate level, within the cross-section, and in a Russell 1000/2000 reconstitution, and recasts the result as an equivalence bound. Separating the trading- and information-side channels, or reaching the high-frequency speed route, would require data on order-flow synchronization, vendor and model concentration, and intraday activity that this sample leaves open.

IV. Empirical Analysis

The empirical analysis measures the strength of the amplification the model predicts, the degree to which the same shock propagates more strongly when the system is closer to the stability boundary. It proceeds in two parts. The first part is descriptive motivation. It constructs the Reflexivity Index (RI) and its components, the Reflexivity Share of Variance (RSV) and Implied Feedback (IF), documents their trajectory, and adds two corroborating patterns, the co-movement of implied feedback with an independent trading-layer commonality measure and the thickening of return tails. These express volatility persistence in the model's own quantities and document that markets behave as if feedback intensity rose over the algorithmic era. They are suggestive and call for a direct test of magnitude. The second part is the empirical core that supplies it, a battery of shock×state amplification tests designed so that the feedback state must do work beyond what ordinary volatility dynamics already explain.

The battery is organized by the routes to high feedback gain it measures (Section III.E). The within-date tests based on commonality, ownership, and index exposure, including the Russell reconstitution, all measure the common-exposure route, the low-frequency trading-layer force that rose over the sample. The aggregate local projection measures the coupling route, and the feedback-gain test speaks to the secular speed and automation of trading directly. Each operates at the monthly horizon at which these slow-moving forces are identifiable. The intraday speed dynamics, the information-layer route (algorithmic curation and narrative amplification), and strong model homogeneity enter only through the combined feedback diagnostics, and together they are the

frontier this sample leaves to higher-frequency data and to the post-2024 record.

The battery produces four results. The aggregate test traces a positive, horizon-rising response that is suggestive though not statistically decisive. The within-date cross-sectional tests, using return commonality, institutional ownership, and index exposure as the feedback state, place the amplification within a narrow band around zero, while the same designs detect amplification on size and market beta. A Russell 1000/2000 reconstitution quasi-experiment, in which index assignment raises institutional ownership by 3.1 percentage points with a strong first stage, holds the reduced-form amplification within the same band. A direct feedback-gain test finds the gain stable across the sample once volatility episodes are treated symmetrically. The evidence establishes a pre-frontier-AI baseline. Through 2024, the standing-commonality and common-ownership routes hold low-frequency amplification to an economically small magnitude. The descriptive diagnostics and two illustrative episodes, the 2010 Flash Crash and the 2021 GameStop surge, provide monitoring context.

A. Data and Measurement

All empirical measures are constructed from daily returns on the CRSP value-weighted market portfolio over the 1980 to 2024 sample period. Daily log returns are used to compute realized volatility within each month as the standard deviation of daily returns, annualized by multiplying by $\sqrt{252}$. This procedure yields a monthly time series of market-level realized volatility. Measuring volatility directly from realized returns in this way follows the realized-volatility tradition (Andersen et al., 2003; Barndorff-Nielsen and Shephard, 2002), applied at a monthly frequency suited to the long historical sample.

Volatility persistence is measured as the first-order autocorrelation of $\log(RV_t)$ computed with a 24-month rolling window. The baseline level of persistence, denoted ϕ_0, is defined as the median autocorrelation during the pre-automation benchmark period from 1980 to 2002. This empirical ϕ_0 is a low-automation historical benchmark, distinct from the model's literal no-reflexivity limit ($\alpha = \delta = 0$, $\phi_0 = 1 - \lambda$). It provides a reference point for evaluating changes in reflexivity following the diffusion of automated trading and does not assert that pre-2003 markets had zero feedback. The Reflexivity Share of Variance (RSV) is calculated as

$$\mathrm{RSV}_t = 1 - \frac{1 - \phi_t^2}{1 - \phi_0^2},$$

which measures the share of volatility behaving as if attributable to reflexive feedback relative to the

historical baseline. Implied Feedback (IF) is derived from the reduced-form mapping

$$\phi = (1 - \lambda) + \lambda(\alpha + \delta\beta).$$

In the baseline implementation I set $\lambda = 1$, so $IF_t = \phi_t$. Naturally, because λ only rescales the mapping between persistence and feedback intensity, the qualitative time-series pattern and the post-2003 increase remain similar for $\lambda \in \{0.6, 0.8, 1.0\}$. This mapping provides an empirical proxy for the combined feedback intensity between market and information algorithms. IF is a reduced-form monitoring statistic for the combined feedback intensity. It does not provide a structural decomposition of the trading and information channels. Standard errors for regime means and confidence intervals are computed using HAC(12) estimators (Newey and West, 1987) on monthly data.

Consistent with gradual technological diffusion, IF_t is interpreted as a continuous state variable rather than as a discrete regime indicator. For descriptive comparisons, the analysis uses 2003 as a reference point within an extended electronification transition, corresponding to the diffusion of electronic order books, the move to decimal pricing, and the shift from manual to automated execution, consistent with Hendershott et al. (2011). Data-driven (Bai–Perron) breaks (Bai and Perron, 1998) place the acceleration in implied feedback earlier, at 1997:Q2 (coinciding with the SEC Order Handling Rules and the ECN era), and the shift in the variance-share measure over a broad 2003–2007 window (Table 3), consistent with gradual diffusion rather than a single discrete switch.

Table 1 reports summary statistics for the core monthly series used throughout the analysis.

B. Implied Feedback and Volatility Persistence

This subsection examines the evolution of volatility persistence and the diagnostics built from it. These diagnostics map persistence into the model's feedback parameter and trace whether markets behave as if feedback intensity is rising. They are monitoring statistics that motivate the direct amplification tests of Section IV.E, which measure the magnitude of any amplification separately. Throughout, the diagnostics describe the feedback state, and Section IV.E establishes how strongly shocks actually propagate within it.

Figure 2 aligns the empirical trajectory of the Reflexivity Share of Variance (RSV) with the historical diffusion of algorithmic trading. Prior to 2003, volatility shocks were typically short-lived and RSV remained close to zero. Beginning in the early 2000s, volatility persistence increases, a pattern that coincides with the expansion of electronic communication networks and the migration

Table 1: Descriptive statistics

This table reports summary statistics for the U.S. equity market over 1980–2024. *Panel A* (aggregate monthly series): RV_t is realized volatility; Z_t the market shock (largest standardized daily return in the month); RSV_t and Implied Feedback IF_t the diagnostics of Section IV.A; AR_t the top-500 absorption ratio. AC(1) is the lag-one autocorrelation. *Panel B* (stock-months): the four cross-sectional feedback states and three firm controls; $p1$, $p99$ are the 1st and 99th percentiles. *Panel C* (Russell boundary firm-years, ±250 rank band): R2000 assignment, the float-rank running variable, and the index weight each assignment confers (basis points), whose ∼9× jump the first stage exploits. N is the number of observations.

Panel A. Aggregate monthly series

Variable	Mean	Median	SD	Min	Max	AC(1)	N
Realized volatility, RV_t	0.143	0.121	0.089	0.044	0.885	0.622	540
Market shock, Z_t	2.51	2.28	1.02	1.00	10.84	−0.12	539
Reflexivity share of variance, RSV_t	0.111	0.040	0.191	−0.088	0.706	0.965	518
Implied feedback, IF_t	0.360	0.340	0.224	0.000	0.854	0.964	518
Absorption ratio, AR_t	0.397	0.376	0.116	0.211	0.622	0.986	518

Panel B. Cross-sectional stock-months

Variable	Mean	Median	SD	$p1$	$p99$	N
Feedback states						
Return commonality, $\rho_{i,t}^{PC}$	0.380	0.358	0.183	0.068	0.793	251,350
Institutional ownership, $IO_{i,t}$	0.595	0.630	0.265	0.015	1.000	583,156
Log ownership breadth (# managers)	5.08	5.22	1.20	1.61	7.50	583,156
Index weight (bps)	11.0	5.0	23.5	0.4	109.5	502,381
Firm controls						
Market beta, $\beta_{i,t}$	1.04	0.99	0.46	0.14	2.46	251,350
Log size, $\log ME_{i,t}$	16.02	15.98	1.330	13.45	19.40	251,350
Log idiosyncratic volatility, $\log RV_{i,t}^{idio}$	−1.581	−1.603	0.515	−2.689	−0.243	251,350

Panel C. Russell boundary sample

Variable	Mean	Median	SD	Min	Max	N
Russell 2000 indicator	0.40	0	0.49	0	1	7,767
Running variable (rank − 1000)	−37.5	−55.0	142.3	−250	250	7,767
R2000 index weight (bps)	13.47	12.80	3.19	7.84	26.99	3,104
R1000 index weight (bps)	1.49	1.48	0.39	0.49	2.57	4,663

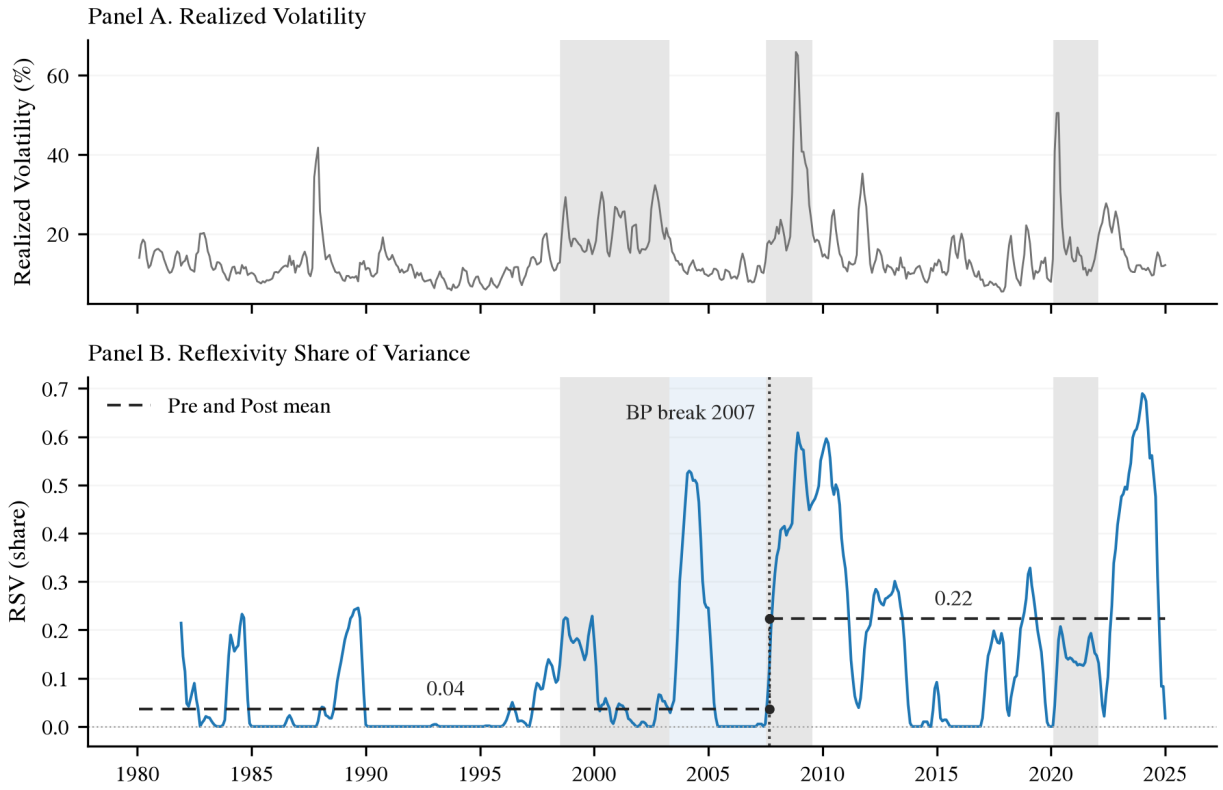


Figure 2: Realized Volatility and the Rising Reflexivity Share of Variance. Panel A plots monthly realized volatility of the CRSP value-weighted market over 1980–2024. Panel B plots the Reflexivity Share of Variance (RSV) derived from the persistence of log realized volatility relative to the 1980–2002 baseline. Shaded bands mark major market stress episodes (Dot-com bubble, Global Financial Crisis, and COVID-19). The vertical dotted line and light shaded band mark a data-driven (Bai–Perron) structural break in RSV and its near-optimal band, which falls within a broad 2003–2007 window. Dashed horizontal lines indicate mean RSV before and after the estimated break.

of liquidity provision toward automated systems. The subsequent moderation around 2005 to 2006 broadly aligns with the introduction of hybrid market models and early regulatory safeguards affecting algorithmic trading. Despite these changes, RSV remains positive and rises again during 2008 to 2010, a period that includes the 2010 Flash Crash (see Kirilenko et al., 2017), widely discussed as a prominent example of algorithmically amplified market dynamics.

Mean RSV is about 0.22 after the data-driven break, compared with about 0.04 before it (Table 3), and elevated levels persist thereafter. The persistence diagnostic behaves as the model's monitoring logic anticipates. A substantially larger fraction of volatility behaves as if generated through reflexive feedback in the more automated sample. Whether this elevated feedback state translates into stronger

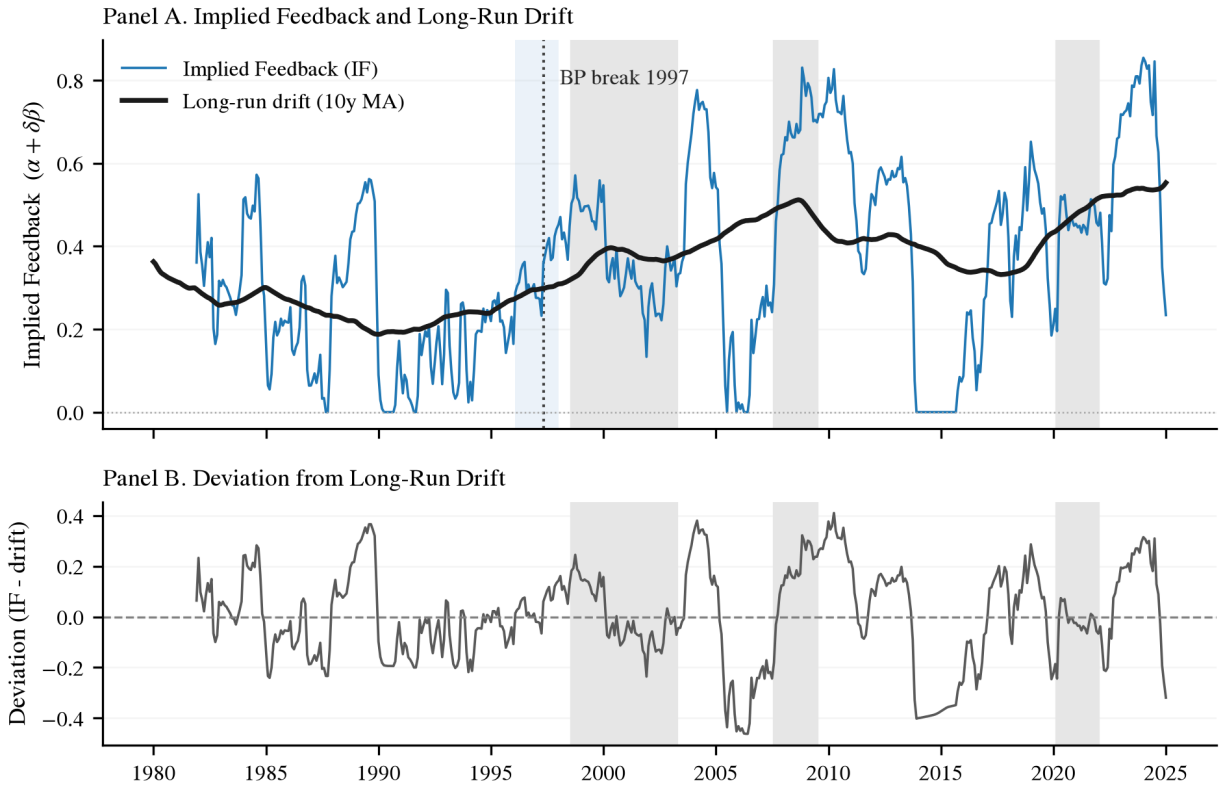


Figure 3: Implied Feedback Drifts Upward over the Algorithmic Era. Panel A plots the estimated implied feedback parameter derived from the persistence of log realized volatility of the CRSP value-weighted market over 1980–2024. The black line shows a centered 10-year moving average of implied feedback. Panel B plots deviations of implied feedback from its long-run trend. Shaded bands mark major market stress episodes (Dot-com bubble, Global Financial Crisis, and COVID-19). The vertical dotted line and light shaded band mark a data-driven (Bai–Perron) acceleration in implied feedback at 1997:Q2 and its near-optimal band, coinciding with the SEC Order Handling Rules and the rise of electronic communication networks.

shock propagation is a separate, quantitative question, measured directly in Section IV.E, where the amplification carried by the state is bounded to an economically small magnitude.

Figure 3 traces the evolution of the implied-feedback parameter derived from the persistence of log realized volatility. Panel A shows estimated implied feedback together with its long-run trend, while Panel B plots deviations from that trend. Two patterns emerge. First, implied feedback exhibits a clear upward drift over time, consistent with the diffusion of algorithmic trading and information processing. Second, the trajectory is not monotonic. Periods of acceleration and moderation occur around major market and technological transitions.

Prior to the early 2000s, implied feedback fluctuates around relatively low levels, consistent with

markets in which human trading and manual liquidity provision dominated. Beginning in the early 2000s, the series rises markedly as electronic communication networks expanded and automated execution systems began to reshape liquidity provision and price discovery. This shift coincides with the onset of large-scale algorithmic trading and the first sustained increase in implied endogenous feedback within market dynamics.

Table 2: Implied Feedback Across Market Phases

This table reports the implied feedback parameter $\alpha + \delta\beta$ across successive phases of market development. Implied feedback is derived from the persistence of log realized volatility using $\phi = (1 - \lambda) + \lambda(\alpha + \delta\beta)$ with $\lambda = 1$. Standard errors are Newey–West (HAC) with 12 monthly lags. The phases reflect the diffusion of algorithmic market infrastructures: early electronic adoption (2003–2010), mature algorithmic trading (2011–2019), and the AI-intensive phase after 2020. The 2003 boundary is an a-priori reference for the electronification transition; the data-driven (Bai–Perron) break dates for the underlying series are reported in Table 3.

Market Phase	Mean	SE (HAC)	CI 2.5%	CI 97.5%	N
1980–2002	0.27	0.03	0.21	0.33	276
2003–2010	0.49	0.09	0.32	0.67	96
2011–2019	0.28	0.10	0.09	0.46	108
2020–2024	0.56	0.06	0.43	0.68	60
Post–2003 (All)	**0.44**	0.04	0.35	0.49	264

Subsequent years show alternating phases of rising and moderating implied feedback. The diagnostic remains elevated through 2003 to 2010, a period associated with the diffusion of early electronic and algorithmic market structures, before moderating during 2011 to 2019, a moderation whose interpretation I return to in Section V. It rises again after 2020, coinciding with the broader expansion of machine-learning-based trading and forecasting strategies (e.g., Gu et al., 2020; Bryzgalova et al., 2025).

Table 2 summarizes these dynamics across successive market regimes. Mean values increase from 0.27 in the pre-2003 baseline to 0.49 during the early automation phase, stabilize around 0.28 during the mature algorithmic period, and rise again to 0.56 in the AI-driven phase after 2020. The post-2003 aggregate mean of 0.44 is significantly higher than the historical benchmark, consistent with the view that a larger share of volatility persistence behaves as if it reflects endogenous feedback generated by algorithmic interaction. Because the model implies non-negative feedback intensity, I also estimated a shape-restricted version. The resulting regime means are 0.27, 0.50, 0.33, and 0.56, reinforcing the upward trend in implied feedback.

The figure and regime estimates indicate a gradual increase in the diagnostic measure of feedback

intensity over the algorithmic era, the trajectory the monitoring logic is designed to surface. Its interpretation is disciplined by the direct tests of Section IV.E. There, the same upward drift is traced to volatility persistence together with the clustering of stress episodes, while the strength of shock amplification in the elevated-feedback state stays economically small. The diagnostics therefore locate where to look. Two further descriptive signatures track the same drift. First, return tails thicken over the sample. The rolling 99th percentile of the absolute daily market return rises from a mean of 2.78% before 2003 to 3.69% after, with the largest extremes remaining concentrated in the Global Financial Crisis and the COVID-19 shock, where fundamentals, jumps, and liquidity dominate. Second, implied feedback co-moves with an independent trading-layer commonality measure (Section IV.C). The amplification tests of Section IV.E then measure what is found.

C. Evidence on Trading-Layer Commonality

The microfoundation traces aggregate reflexivity to model homogeneity, with ρ_M governing the trading layer (Definition 1). The Reflexivity Index is a composite diagnostic of this feedback state, not a decomposition of the two channels. As a model-free check on the trading-layer channel, I compare implied feedback to the absorption ratio of Kritzman et al. (2011), the share of cross-sectional return variance captured by the top principal components of the top-500 large-cap universe, on the same trailing 24-month window as IF_t (Table 8, reported in the Internet Appendix). The two series co-move strongly ($\rho = 0.61$ in levels), and the co-movement survives removal of a common linear trend ($\rho = 0.51$) and remains positive and significant in first differences ($b = 0.67$, HAC $p = 0.001$), so it is not merely a shared secular rise. Because the absorption ratio is an *outcome* of synchronized trading, I read this as suggestive fingerprint evidence that IF_t tracks the rising common component of trading-layer activity, the observable footprint of ρ_M, and not as identification. Its magnitude is measured directly in the within-date cross-section of Section IV.E, which uses the per-stock decomposition $\rho^{PC}_{i,t-1}$ of this very absorption ratio as the feedback state. The information-layer component (ρ_I) is not measurable over this sample and is left for future work.[9]

[9]Two external results point the same way. Brogaard et al. (2025) exploit the pandemic suspension of NYSE floor trading and find that when human intermediaries are absent, markets exhibit higher short-horizon volatility, stronger cross-security correlation, and wider effective spreads. Removing heterogeneous discretionary traders raises effective trading-layer commonality (ρ_M) and synchronizes the algorithms that remain. Bogousslavsky et al. (2024) show that periods of elevated information-based order flow coincide with greater short-horizon volatility and wider spreads, the pattern predicted when α rises. Both are corroboration from external designs, and neither is evidence generated by this paper.

D. Limits of the Aggregate Diagnostics

The descriptive diagnostics are constructed from one aggregate series, the persistence of realized volatility, so their upward drift could in principle reflect forces correlated with automation yet distinct from it. Three are leading. The first is the changing macroeconomic volatility environment, including the Great Moderation and its reversal, shifting monetary-policy regimes, and the leverage cycle (Geanakoplos, 2010), which move realized volatility and its persistence. This is why the index is framed as a diagnostic of the combined feedback state, with the separate algorithmic channel left to be identified with direct data. The second is crisis clustering. The high-reflexivity regimes coincide with the dot-com unwind, the Global Financial Crisis, and COVID-19, so the post-2003 drift could be carried by a few stress episodes. The third is structural change in market composition, including the growth of the volatility-derivatives complex, exchange-traded and index investing (Ben-David et al., 2018), dealer hedging, and market fragmentation. A single aggregate series cannot separate any of these from algorithmic feedback.

This limitation is intrinsic to one aggregate series, and it is what the direct tests are built to overcome. The within-date cross-section absorbs the entire aggregate regime, the macroeconomic environment, the crisis, and the common shock alike, through origin-month fixed effects, so identification rests only on cross-sectional variation in the feedback state. The Russell reconstitution adds quasi-random variation in common ownership, and the feedback-gain test treats volatility episodes symmetrically across eras, meeting the crisis-clustering concern directly. The timing of the drift is itself gradual. A data-driven (Bai–Perron) break places the acceleration in implied feedback at 1997:Q2, coinciding with the SEC Order Handling Rules and the rise of electronic communication networks, with the variance-share measure shifting over a broader 2003–2007 window (Table 3), consistent with algorithmification arriving in successive technological waves.

A related question is whether implied feedback merely proxies for option-implied volatility, so that the VIX would serve equally well. The CBOE VIX is available only from 1990. In the point-forecast horse race over the overlapping sample (Internet Appendix), option-implied volatility, like the HAR controls, absorbs most of IF_t's forecast content, reinforcing the diagnostic reading. The distinct value of implied feedback is that it is constructible from standard equity-return data back to 1980, well before option-implied measures exist.

E. Direct Tests of Shock Amplification

The aggregate mechanism is algorithmic feedback. Trading and information systems react to shared signals, and shocks amplify as feedback gain rises. Taking this to the cross-section requires an

Table 3: Data-driven structural breaks in implied feedback

This table reports, for each series, the single least-squares break date of Bai and Perron (1998), the mean on each side of the break, and the near-optimal band (all candidate dates with SSR within 1% of the minimum). The narrow IF band (≈1997, coinciding with the SEC Order Handling Rules and the ECN era) and the broad RSV band (2003–2007, spanning Reg NMS and the rise of algorithmic trading) are consistent with algorithmification arriving in successive technological waves.

Series	Break	Mean (pre)	Mean (post)	Near-optimal band	Width (mo.)
Implied feedback (IF)	1997-04	0.234	0.430	1996-01 to 1997-12	24
Reflexivity share (RSV)	2007-08	0.036	0.223	2003-03 to 2007-11	14

observable measure of how exposed a given stock is to that feedback, and no direct per-stock measure of algorithmic feedback exists. I therefore use common exposure as the proxy. Stocks with high return commonality, high common institutional ownership, and high index membership are the stocks that algorithmic and rules-based strategies trade in correlated ways. Active algorithmic strategies react to shared price and information signals, and index-linked and passive strategies follow index flows into synchronized, mechanical trading across constituents (e.g., Appel et al., 2016; Ben-David et al., 2018). Both produce correlated, rule-driven trading, the trading-layer synchronization the model formalizes. A stock's common-exposure loading therefore proxies the algorithmic feedback surrounding it. The interaction of an aggregate shock with a stock's pre-shock common exposure then tests whether shocks propagate more strongly precisely where algorithmic co-trading is most concentrated, which is where the feedback mechanism should be strongest if it operates at all.

Common exposure also arises from ordinary factor structure, so the Russell 1000/2000 reconstitution provides the decisive test. Assignment moves index-linked, mechanical ownership quasi-randomly at the cutoff while holding fundamentals approximately fixed, isolating the rules-based, algorithmic component of common exposure from generic factor commonality.

The direct tests measure the magnitude the diagnostics leave open. The object of interest is the comparative dynamic at the heart of the model, whether the *same shock propagates more strongly* when the pre-shock feedback state is elevated, a quantity that a volatility forecast does not mechanically reproduce.

Design. I summarize each month by a market shock Z_t, the largest daily absolute return on the CRSP value-weighted index standardized by an ex-ante volatility estimate (RiskMetrics EWMA, using information through $d-1$). The standardization makes Z_t a surprise rather than a level. The feedback state enters lagged and predetermined. Outcomes are future log realized volatility (total

and idiosyncratic) or the cumulative HAR forecast error of log RV. All forecasts and states use information through $t-1$. The amplification coefficient is the interaction of the shock with the lagged state. Standard errors are HAC at the aggregate level and two-way clustered (firm and month) in the panel. There is no imputation and no look-ahead.

Aggregate test. I first estimate state-dependent local projections (Jordà, 2005) of the cumulative HAR forecast error of log RV on Z_t, on Z_t interacted with the lagged absorption ratio AR_{t-1}, and on controls (lagged log RV, HAR components, and $Z_t \times \log RV_{t-1}$):

$$Y_{t,h} = a_h + b_h Z_t + c_h \left(Z_t \times AR_{t-1}\right) + d_h AR_{t-1} + e_h \left(Z_t \times \log RV_{t-1}\right) + \Gamma_h' X_{t-1} + \varepsilon_{t+h}. \tag{24}$$

The amplification coefficient c_h rises monotonically with the horizon and is individually significant at long horizons (Table 4, Panel A), the qualitative shape the model predicts. The pre-registered pooled test of $\sum_{h=1}^{12} c_h = 0$ returns $p = 0.25$ under a moving-block bootstrap, and a shuffled-state placebo supports reading the aggregate result as suggestive rather than decisive. With one regime variable and a single macro time series, the aggregate design carries limited power, which motivates moving to the cross-section, where power is abundant.

Cross-sectional tests. The model's amplification prediction concerns the aggregate feedback gain, which the microfoundation traces to the common, non-diversifying component of algorithmic responses ($\alpha_1 \rho_M$ and $\beta \delta_1 \rho_I$). That aggregate prediction has a cross-sectional counterpart. Stocks differ in how strongly they load on this common component. Those more exposed to it, through higher return commonality, common ownership, or index membership, should respond more strongly to the *same* common shock if the route amplifies. The within-date cross-section tests for that footprint. It also separates amplification from a competing reading of the aggregate result, in which high-commonality periods merely coincide with persistent-volatility crises. With origin-month fixed effects τ_t absorbing the entire aggregate regime (and the common shock Z_t), identification comes only from whether stocks with higher pre-shock commonality respond more strongly to the same shock in the same month. I estimate

$$Y_{i,t+h} = \mu_i + \tau_t + b_h \left(Z_t \times S_{i,t-1}\right) + \gamma_h' \left(Z_t \times C_{i,t-1}\right) + \kappa_h' C_{i,t-1} + \varepsilon_{i,t+h}, \tag{25}$$

where $C_{i,t-1}$ collects market beta, log size, and log idiosyncratic volatility, and the regression is run separately for each of four feedback states $S_{i,t-1}$: stock-level principal-component commonality $\rho_{i,t-1}^{PC}$ (the per-stock decomposition of the absorption ratio), institutional ownership and ownership

breadth from 13F filings, and FTSE index weight. Across all four states and both outcomes, the amplification coefficient sits within a narrow band around zero at every horizon (Table 4, Panel B; roughly 250,000 stock-months, two-way clustered). The design carries ample power to detect amplification. In the *same* regressions, the shock interacted with market beta and with size is significant ($t = -2.6$ and $t = 4.0$), and the level controls load strongly ($t \approx 35$). Cross-sectional heterogeneity in shock response is readily detected, and the common-exposure route contributes an economically small amount of it. This footprint is a proxy for the route's contribution to aggregate gain. A near-zero cross-sectional response bounds that contribution, while the level of the aggregate parameter ϕ is read from the diagnostics and the aggregate design.

Russell reconstitution. The sharpest test exploits the annual Russell 1000/2000 reconstitution, a design widely used to identify the price and ownership effects of index membership (e.g., Chang et al., 2015; Appel et al., 2016; Ben-David et al., 2018). Near the cutoff, stocks pushed into the Russell 2000 receive mechanically higher passive and index-linked trading, raising the common exposure to which the theory assigns amplification. With actual FTSE membership, a rank band around the cutoff, and firm and month fixed effects, the first stage is strong. Russell 2000 assignment raises institutional ownership by 3.1 percentage points ($t = 4.43$). The reduced-form amplification interaction $Z_t \times \mathbf{1}\{\text{R2000}\}$ stays close to zero at every horizon for both total and idiosyncratic volatility (Table 4, Panel C). Where index-linked ownership is quasi-randomly higher, the strength of shock propagation is essentially unchanged. The design is well powered. It detects an amplification as small as roughly one percent of realised volatility (Table 5), so the estimate is a tight bound. I report the reduced form because its units, percent of realised volatility, are the units in which the equivalence margin is defined. With a strong first stage, a two-stage estimate would restate the same bound per point of institutional ownership. A deliberately sharpened specification, with a narrower band, an interior donut, and local-linear control in the running variable, reproduces the bound with a wider interval on the smaller sample. The first stage remains strong in this specification (4.1 percentage points, $t = 3.7$). The equivalence analysis at the end of this subsection formalizes the bound for both the Russell and the cross-sectional designs.

Feedback gain over time. Finally, I test directly whether the system's loop gain rose with automation, without conditioning on a homogeneity state. The cumulative impulse response of log RV to Z_t appears larger in the post-2003 era than before, about twice as large in raw form. A symmetric treatment of volatility episodes accounts for this appearance (Table 4, Panel D). Excluding major stress episodes in *both* eras (1987, 1997–98, the dot-com unwind, the global financial crisis,

and COVID-19), single-window persistence is essentially identical ($\phi = 0.49$ versus 0.52) and the cumulative response falls from 2.0× to a residual 1.3×, which a continuous time-trend interaction leaves statistically flat at every horizon. Rolling persistence, moreover, tracks the contemporaneous volatility level. Feedback gain is stable across the sample, and the apparent increase reflects volatility persistence together with the clustering of stress episodes.

Equivalence bounds. An estimate near zero carries information in proportion to the effect the design could have detected. A conventional significance test can detect amplification but cannot, on its own, bound it. I therefore report the amplification as an equivalence test (Schuirmann, 1987; Lakens, 2017), which states the magnitude the data rule out and thus measures the ceiling on the effect. The policy and academic literatures treat algorithmic amplification as a live systemic risk and supply a verb without a number. I pre-specify the equivalence margin δ in the units in which stability is assessed, as a fraction of unconditional realised volatility, the volatility a typical month already carries. An effect of δ% then states how much the feedback channel adds to the volatility response that a one-standard-deviation shock would otherwise produce. Two benchmarks fix the economically meaningful range. An amplification contributing on the order of 5 to 20% of unconditional realised volatility per shock, compounded across repeated shocks and horizons, can account for a meaningful share of the several-fold volatility increases that define stress episodes. A channel below this range therefore cannot by itself convert an ordinary shock into a stress event.[10] Independently, the same regressions detect the cross-sectional amplification carried by firm size and market beta, so a margin whose floor is 5% would register an algorithmic route even a fraction the size of these well-documented amplifiers. The size and beta controls thus pin the 5% floor, while the stress-episode benchmark pins the 10 to 20% ceiling. I therefore report results at $\delta \in \{5\%, 10\%, 20\%\}$. The amplification effect is expressed as the additional percentage by which a one-standard-deviation market shock raises future realised volatility when the feedback state moves across its interquartile range (cross-section) or under Russell 2000 assignment.

Table 5 reports, for each design, the minimum detectable effect (MDE) at 80% and 90% power and the two-one-sided-test (TOST) verdict at each margin. The cross-sectional and Russell designs are powered to detect amplification as small as roughly 1% of realised volatility, and both *reject* amplification of 5%, 10%, and 20% at every horizon. The data establish that any amplification is bounded below 5% of unconditional realised volatility, and below about 1% at the binding horizon.

[10]The compounding follows the model's own persistence. A per-shock increment to the volatility response accumulates over its decay to roughly $1/(1-\phi)$ times its impact, on the order of two at the measured ϕ, and summed across the several large shocks that cluster in a stress month a per-shock channel of 5 to 20% then contributes on the order of tens of percent to cumulative volatility.

Table 4: Direct tests of shock amplification

This table reports how much more strongly a market shock Z_t propagates when the pre-shock feedback state is elevated, across four designs of increasing power. Panels A–C report the amplification coefficient on Z_t times the pre-shock state (the lagged absorption ratio in A, the four feedback states in B, Russell-2000 assignment in C), cumulative over $h = 1$–12 in Panels B and C; Panel D asks whether loop gain rose over time. The aggregate projection (Panel A, $N = 456$) is suggestive (pooled $\sum_h c_h = 0$: bootstrap $p = 0.25$); the cross-section (Panel B, ≈250,000 stock-months) detects size and beta in the same regressions ($t = 4.0$ and -2.6); the Russell first stage (Panel C) raises institutional ownership by 3.1pp ($t = 4.43$). Across the powered designs the amplification stays within a narrow band around zero; standard errors are HAC in the time series and two-way clustered by firm and month in the panels.

Panel A. Aggregate local projection

Horizon h	1	3	6	8	10	12
Amplification c_h	−0.001	0.003	0.093	0.145	0.212	0.292
HAC s.e.	0.021	0.041	0.061	0.087	0.109	0.124
t-statistic	−0.07	0.06	1.52	1.66	1.94	2.35

Panel B. Within-date cross-section

Feedback state $S_{i,t-1}$	Idiosyncratic vol	Total vol
Return-PCA commonality ρ^{PC}	0.004	−0.003
Institutional ownership (13F)	0.000	0.001
Ownership breadth (# managers)	−0.022	−0.025
FTSE index weight	0.007	0.008

Panel C. Russell reconstitution

Horizon h	1	3	6	12
Amplification b_h, total vol	0.0019	0.0040	−0.0021	−0.0036
s.e.	0.0028	0.0027	0.0037	0.0035
idiosyncratic vol	0.0008	0.0030	−0.0011	−0.0027

Panel D. Feedback gain over time

	Pre-2003	Post-2003
Volatility-state persistence ϕ, full sample	0.643	0.646
stress episodes excluded	0.488	0.520
Cumulative IRF of log RV to Z_t (post ÷ pre):		
full sample	2.0×	
stress episodes excluded	1.3×	
Continuous time-trend $\times Z_t$	flat ($h = 1, 3, 6, 12$)	

The bound survives dropping warm-up observations and restricting the sample period, so it does not hinge on any single preprocessing choice. Because its outcome is a twelve-month cumulative forecast error on a single regime series, the aggregate design carries a minimum detectable effect above 50%. It supports a wide bound and is read as suggestive. Figure 4 plots the per-horizon point effects with 95% confidence intervals against the δ thresholds.

Table 5: Equivalence tests: bounding the amplification

This table reports the minimum detectable effect (MDE) at 80% and 90% power and the two-one-sided-test (TOST) verdict at each margin δ, by design. The amplification effect and δ are expressed as a percentage of unconditional realised volatility, the additional response of future realised volatility to a one-standard-deviation shock under the relevant state shift. MDE is the most conservative value across horizons $h \in \{1, 3, 6, 12\}$, and a verdict of *reject* establishes equivalence at every horizon.

	MDE (power)		TOST verdict at margin δ		
Design	80%	90%	5%	10%	20%
Aggregate LP (cumulative)	54.8%	63.4%	n/a	partial	partial
Cross-section (firm×date FE)	1.27%	1.47%	reject	reject	reject
Russell reconstitution	1.06%	1.22%	reject	reject	reject

Together, the four tests deliver a consistent measurement. Through 2024, the low-frequency routes hold shock amplification to an economically small magnitude, and the two best-powered designs place it below the level that would matter for stability. This establishes a clean baseline for the mechanism, an interpretation developed in Section V.

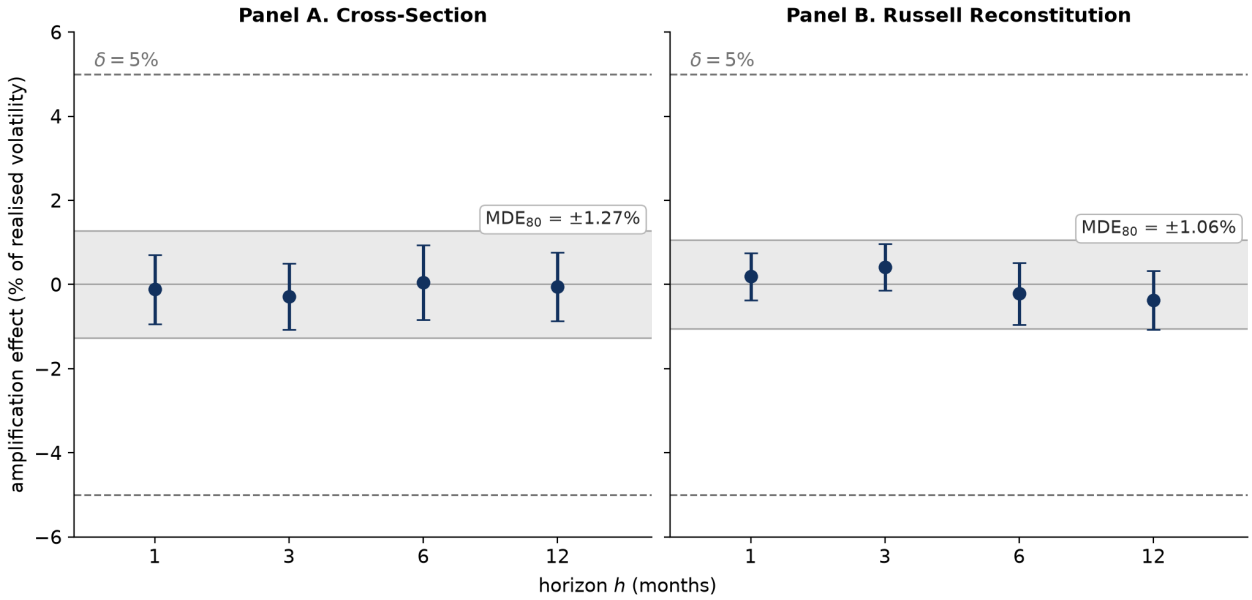


Figure 4: Bounding Amplification Against the Equivalence Margin. Per-horizon amplification effect (percent of unconditional realised volatility) with 95% confidence intervals at horizons of 1, 3, 6, and 12 months. **Panel A** reports the cross-sectional design and **Panel B** the Russell reconstitution; the two panels share a common vertical axis. The axis is zoomed to ±6% so the intervals are visible. The shaded grey band is $\pm\text{MDE}_{80}$, the smallest effect each design detects at 80% power, with edges at ±1.27% (cross-section) and ±1.06% (Russell); the estimates sit at the edge of what the designs could detect. Dashed lines mark the conservative equivalence margin $\delta = 5\%$; every interval lies well inside it, and the $\delta = 10\%$ and 20% margins lie beyond the plotted range. The two-one-sided tests reject $\delta = 5\%, 10\%, 20\%$ at every horizon in both designs.

F. Illustrative Mechanisms

The sample's two vivid algorithmic episodes illustrate the model's channels firing episodically even as the low-frequency signature stays bounded. The 2010 Flash Crash illustrates the market channel (α). Within minutes, interacting execution algorithms withdrew liquidity, amplified order-book imbalances, and generated a self-reinforcing collapse across equities and futures, a transient instability that unwound once liquidity returned and that subsequent safeguards are designed to interrupt (Kirilenko et al., 2017). The 2021 GameStop episode illustrates the information channel (δ). A narrative originating in retail discussion was propagated by ranking, recommendation, and sentiment algorithms into a self-reinforcing loop with market behavior, attention itself becoming a tradable signal as rising prices re-entered the models as fresh input (Pedersen, 2022). Both episodes are illustrative and do not identify the channels separately. They operate at frequencies and in scopes the monthly designs do not reach. The Internet Appendix collects the detail, including the

narrative-diffusion evidence behind the information channel (Shiller, 2017; Bybee et al., 2024; Aleti and Bollerslev, 2025; Hirshleifer et al., 2025) and the specific mechanisms linking narrative diffusion to volatility.

V. Discussion and Implications

The central result is a separation between feedback and amplification. Algorithmic markets exhibit a persistently positive feedback state, and the monitoring diagnostics drift upward over time, but this feedback has not translated into economically important low-frequency shock propagation.[11] The aggregate evidence is suggestive, while the powered cross-sectional and Russell designs place amplification below five percent of unconditional realised volatility at every horizon. Feedback gain is also stable once major volatility episodes are treated symmetrically. Together, these findings locate the historical market on the subcritical portion of the model's amplification curve. Feedback is present but sufficiently distant from the stability boundary that its measurable low-frequency effects remain small. The separation also reconciles the two views that frame the literature. Micro-level gains in liquidity and price discovery (e.g., Hendershott et al., 2011; Brogaard et al., 2014) coexist with a rising feedback state, because gain translates into material amplification only near the boundary.

Figure 5 provides the model counterpart to this separation. Amplification is convex in feedback gain. The variance multiplier changes slowly at low and moderate gain, a factor of roughly 1.2 to 1.7 at the measured levels, and rises rapidly as the system approaches the stability boundary, the same threshold structure through which financial networks shift from absorption to amplification (Acemoglu et al., 2015). The historical estimates place algorithmic markets on the relatively flat portion of this curve. Algorithmification has put the feedback environment in place, but through 2024 its measurable low-frequency channels remained too far from the boundary to generate economically meaningful market-wide amplification.

I interpret these findings as a pre-frontier-AI baseline. Algorithmic Exuberance was already a standing feature of algorithmic markets, but its measurable low-frequency routes remained subcritical through 2024. The bound is deliberately specific. It applies to common exposure, the trading-side route observed through return commonality, institutional ownership, and index membership, and it is strongest in the Russell design, where index-linked ownership varies quasi-randomly around the cutoff. The information layer and model homogeneity are not separately identified in the historical data, while shared foundation models and autonomous systems had not yet diffused widely across

[11] Consistent with this separation, the index adds no incremental forecast power for volatility-scaled exceedances once an ex-ante volatility forecast is used (Internet Appendix).

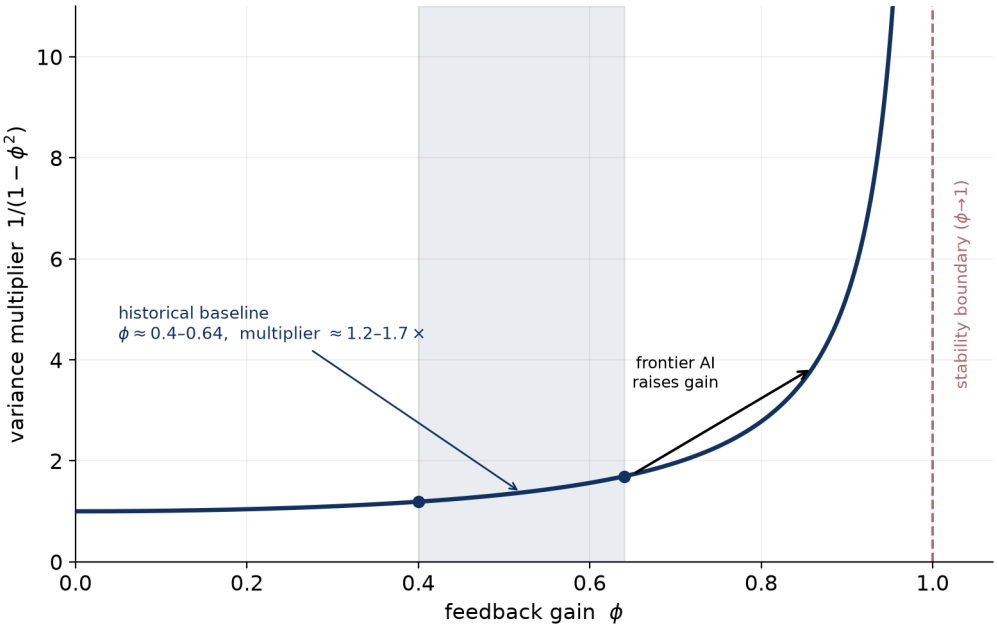


Figure 5: Amplification is convex in feedback gain. The model-implied variance multiplier of the volatility state, $1/(1-\phi^2)$, is nearly flat for low and moderate gain and steepens only as $\phi \to 1$. Here ϕ is the persistence of the volatility state, equal to the feedback gain $\alpha + \delta\beta$ under the baseline $\lambda = 1$. The shaded band marks the historical baseline, the central range of the sustained feedback gain from the post-2003 mean through the full-sample persistence ($\phi \approx 0.4$ to 0.64), where the multiplier is roughly 1.2 to 1.7, far below the divergence at the boundary. Transient monthly implied feedback peaks higher during stress episodes, consistent with the reading that the diagnostic drift reflects volatility persistence and crisis clustering. Equal increments in gain raise amplification progressively more as ϕ rises.

both layers. Their future contribution will be read against this baseline. The appropriate posture is precise vigilance rather than alarm or complacency. The historically measurable routes contribute little, and the framework turns future amplification from a qualitative concern into a detectable event.

The forward-looking implication follows from the same convexity. Equal increases in feedback gain become progressively more consequential as the system moves toward the boundary, and maturing intelligent algorithms and AI agents are the force most likely to move it. For example, in the model economies of Dou et al. (2026), AI speculators with planning modules learn, without communication or shared intent, to create and exploit bubbles when feedback trading is strong, the region where the amplification curve steepens. If frontier AI raises gain jointly through model homogeneity, information curation, speed, and autonomy, the amplification signature should grow in post-2024 data. The shock×state interactions should turn positive and persistent, cumulative

impulse responses should strengthen after symmetric treatment of crisis episodes, and the diagnostics should rise beyond what the volatility level alone explains. The battery can be re-estimated as these technologies diffuse, extended to high-frequency data for the speed channel, where event-time reflexivity diagnostics already exist (Filimonov and Sornette, 2012; Hardiman et al., 2013), and combined with direct measures of model concentration and data commonality. The Reflexivity Index monitors the feedback state, and the historical bound anchors every future test.

A more speculative interpretation concerns the nonmonotonic trajectory of implied feedback. Its moderation from 2011 to 2019 coincides with a concentrated wave of market-structure safeguards aimed at algorithmic trading, including circuit breakers, market-access controls, and system-resilience requirements.[12] If reflexive amplification depends partly on how trading and information systems are permitted to interact (Brogaard et al., 2025), rules that interrupt synchronized feedback should lower effective feedback gain. The timing is consistent with this mechanism and connects to the resilience perspective of Brunnermeier (2024), and the endogenous-risk tradition adds the caution that risk can build when measured volatility is low (Brunnermeier and Sannikov, 2014). If resilience can be engineered, market reflexivity is partly a policy variable.

VI. Conclusion

Algorithmic markets combine machine-driven trading with machine-mediated information. Whether the interaction between these systems amplifies shocks is now a first-order question for financial stability. Until now, that question had remained unmeasured. This paper has provided the first comprehensive measurement of algorithmic shock amplification. In the theory, trading and information algorithms form a coupled feedback system whose gain rises with common exposure, reaction speed, information curation, cross-layer coupling, and model homogeneity. As gain approaches its critical level, shocks dissipate more slowly and the market moves toward a stability boundary beyond which amplification becomes self-sustaining.

The evidence separates the feedback state from its consequences. In U.S. equities from 1980 to 2024, the volatility-implied diagnostic is persistently positive and drifts upward. Yet the cross-sectional and Russell designs, powered to detect effects near one percent of realised volatility, place

[12]The relevant measures include the 2010 market-access rule and single-stock circuit breakers, whose stability effects are themselves contested (Chen et al., 2024b), the Limit Up–Limit Down mechanism phased in over 2012–2013, Regulation SCI in 2014, and the MiFID II algorithmic-trading controls in 2018. This interpretation is not identified. The 2011–2019 period is also the calmest realized-volatility regime in the sample, and because implied feedback is constructed from volatility persistence, tranquility mechanically lowers the diagnostic. The safeguards followed the 2008–2010 stress episodes, so the moderation may also reflect ordinary mean reversion. Separating governance from the volatility regime would require exogenous policy variation, most naturally the staggered international adoption of algorithmic-trading rules.

low-frequency amplification below five percent of unconditional realised volatility at every horizon. The same regressions detect amplification on size and beta, showing that the near-zero estimates do not reflect limited power. Finally, feedback gain proves stable once major volatility episodes are treated symmetrically.

The result is a pre-frontier-AI baseline. Historical algorithmification, passive and index-linked ownership, and return-covariance commonality were not, by themselves, sufficient to generate measurable low-frequency shock amplification. The baseline disciplines the financial-stability concern without dismissing it, and it converts the question into a dated, falsifiable prediction. If shared foundation models, common data pipelines, agentic trading systems, and concentrated AI infrastructure raise feedback gain jointly through homogeneity, information curation, speed, and autonomy, the amplification signature should grow in post-2024 data. The measurement battery specifies what to monitor, and the baseline makes future amplification detectable.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2015, Systemic risk and stability in financial networks, *American Economic Review* 105, 564–608.

Aleti, Saketh, and Tim Bollerslev, 2025, News and asset pricing: A high-frequency anatomy of the SDF, *The Review of Financial Studies* 38, 712–759.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2003, Modeling and forecasting realized volatility, *Econometrica* 71, 579–625.

Appel, Ian R., Todd A. Gormley, and Donald B. Keim, 2016, Passive investors, not passive owners, *Journal of Financial Economics* 121, 111–141.

Bai, Jushan, and Pierre Perron, 1998, Estimating and testing linear models with multiple structural changes, *Econometrica* 66, 47–78.

Barber, Brad M., and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* 21, 785–818.

Barndorff-Nielsen, Ole E., and Neil Shephard, 2002, Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society: Series B* 64, 253–280.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do ETFs increase volatility?, *The Journal of Finance* 73, 2471–2535.

Black, Fischer, 1986, Noise, *The Journal of Finance* 41, 528–543.

Bogousslavsky, Vincent, Vyacheslav Fos, and Dmitriy Muravyev, 2024, Informed trading intensity, *The Journal of Finance* 79, 903–948.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-frequency trading and price discovery, *The Review of Financial Studies* 27, 2267–2306.

Brogaard, Jonathan, Matthew C. Ringgenberg, and Dominik Roesch, 2025, Does floor trading matter?, *The Journal of Finance* 80, 375–414.

Brunnermeier, Markus K., 2024, Presidential address: Macrofinance and resilience, *The Journal of Finance* 79, 3683–3728.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *The Review of Financial Studies* 22, 2201–2238.

Brunnermeier, Markus K., and Yuliy Sannikov, 2014, A macroeconomic model with a financial sector, *American Economic Review* 104, 379–421.

Bryzgalova, Svetlana, Markus Pelger, and Jason Zhu, 2025, Forest through the trees: Building cross-sections of stock returns, *The Journal of Finance* 80, 2447–2506.

Budish, Eric, Peter Cramton, and John Shim, 2015, The high-frequency trading arms race: Frequent batch auctions as a market design response, *The Quarterly Journal of Economics* 130, 1547–1621.

Bybee, Leland, Bryan Kelly, Asaf Manela, and Dacheng Xiu, 2024, Business news and business cycles, *The Journal of Finance* 79, 3105–3147.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello, 2020, Artificial intelligence, algorithmic pricing, and collusion, *American Economic Review* 110, 3267–3297.

Cao, Sean, Wei Jiang, Junbo Wang, and Baozhong Yang, 2024, From man vs. machine to man + machine: The art and AI of stock analyses, *Journal of Financial Economics* 160, 103910.

Carlin, Bruce Ian, Miguel Sousa Lobo, and S. Viswanathan, 2007, Episodic liquidity crises: Cooperative and predatory trading, *The Journal of Finance* 62, 2235–2274.

Cartea, Álvaro, Patrick Chang, José Penalva, and Harrison Waldon, 2026, Algorithmic collusion and a folk theorem from learning with bounded rationality, *Games and Economic Behavior* 157.

Chaboud, Alain P., Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega, 2014, Rise of the machines: Algorithmic trading in the foreign exchange market, *The Journal of Finance* 69, 2045–2084.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *The Review of Financial Studies* 28, 212–246.

Chen, Hui, Winston Wei Dou, and Leonid Kogan, 2024, Measuring "dark matter" in asset pricing models, *The Journal of Finance* 79, 843–902.

Chen, Hui, Anton Petukhov, Jiang Wang, and Hao Xing, 2024, The dark side of circuit breakers, *The Journal of Finance* 79, 1405–1455.

Colliard, Jean-Edouard, Thierry Foucault, and Stefano Lovo, 2026, Algorithmic pricing and liquidity in securities markets, *The Review of Financial Studies*, forthcoming.

Cong, Lin William, Ke Tang, Jingyuan Wang, and Yang Zhang, 2021, AlphaPortfolio: Direct construction through deep reinforcement learning and interpretable AI, Working paper.

Cont, Rama, and Wei Xiong, 2024, Dynamics of market making algorithms in dealer markets: Learning and tacit collusion, *Mathematical Finance* 34, 467–521.

Cookson, J. Anthony, Runjing Lu, William Mullins, and Marina Niessner, 2024, The social signal, *Journal of Financial Economics* 158, 103870.

Corsi, Fulvio, 2009, A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics* 7, 174–196.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *The Journal of Finance* 45, 379–395.

Dou, Winston Wei, Itay Goldstein, and Yan Ji, 2025, AI-Powered trading, algorithmic collusion, and price efficiency, Working Paper 34054, National Bureau of Economic Research.

Dou, Winston Wei, Itay Goldstein, and Yan Ji, 2026, Financial market fragility in the era of AI planning, Working paper.

Dugast, Jérôme, and Thierry Foucault, 2018, Data abundance and asset price informativeness, *Journal of Financial Economics* 130, 367–391.

Engle, Robert F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987–1007.

Farboodi, Maryam, and Laura Veldkamp, 2020, Long-run growth of financial data technology, *American Economic Review* 110, 2485–2523.

Filimonov, Vladimir, and Didier Sornette, 2012, Quantifying reflexivity in financial markets: Toward a prediction of flash crashes, *Physical Review E* 85, 056108.

Geanakoplos, John, 2010, The leverage cycle, *NBER Macroeconomics Annual* 24, 1–66.

Goldstein, Itay, Chong Huang, and Liyan Yang, 2025, Fragility of financial markets, *Annual Review of Financial Economics*, forthcoming.

Goldstein, Itay, Chester S. Spatt, and Mao Ye, 2021, Big data in finance, *The Review of Financial Studies* 34, 3213–3225.

Goldstein, Itay, and Liyan Yang, 2015, Information diversity and complementarities in trading and information acquisition, *The Journal of Finance* 70, 1723–1765.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *The Review of Financial Studies* 33, 2223–2273.

Gufler, Ivan, Francesco Sangiorgi, and Emanuele Tarantino, 2025, (Deep) learning to trade: An analysis of AI trading and market outcomes, Working paper.

Hansen, Karsten T., Kanishka Misra, and Mallesh M. Pai, 2021, Frontiers: Algorithmic collusion: Supra-competitive prices via independent algorithms, *Marketing Science* 40, 1–12.

Hardiman, Stephen J., Nicolas Bercot, and Jean-Philippe Bouchaud, 2013, Critical reflexivity in financial markets: A Hawkes process analysis, *The European Physical Journal B* 86, 442.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of Finance* 66, 1–33.

Hirshleifer, David, Lin Peng, and Qiguang Wang, 2025, News diffusion in social networks and stock market reactions, *The Review of Financial Studies* 38, 883–937.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of Finance* 54, 2143–2184.

Jiang, Jingwen, Bryan Kelly, and Dacheng Xiu, 2023, (Re-)Imag(in)ing price trends, *The Journal of Finance* 78, 3193–3249.

Jordà, Òscar, 2005, Estimation and inference of impulse responses by local projections, *American Economic Review* 95, 161–182.

Jordan, Michael I., and Tom M. Mitchell, 2015, Machine learning: Trends, perspectives, and prospects, *Science* 349, 255–260.

Kelly, Bryan, Semyon Malamud, and Kangying Zhou, 2024, The virtue of complexity in return prediction, *The Journal of Finance* 79, 459–503.

Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2017, The Flash Crash: High-frequency trading in an electronic market, *The Journal of Finance* 72, 967–998.

Kleinberg, Jon, and Manish Raghavan, 2021, Algorithmic monoculture and social welfare, *Proceedings of the National Academy of Sciences* 118, e2018340118.

Kritzman, Mark, Yuanzhen Li, Sébastien Page, and Roberto Rigobon, 2011, Principal components as a measure of systemic risk, *The Journal of Portfolio Management* 37, 112–126.

Lakens, Daniël, 2017, Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses, *Social Psychological and Personality Science* 8, 355–362.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, 2015, Deep learning, *Nature* 521, 436–444.

Lopez-Lira, Alejandro, 2025, Can large language models trade? Testing financial theories with LLM agents in market simulations, arXiv:2504.10789.

Murray, Scott, Yusen Xia, and Houping Xiao, 2024, Charting by machines, *Journal of Financial Economics* 153, 103791.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Ozdenoren, Emre, and Kathy Yuan, 2008, Feedback effects and asset prices, *The Journal of Finance* 63, 1939–1975.

O'Hara, Maureen, 2015, High frequency market microstructure, *Journal of Financial Economics* 116, 257–270.

Pedersen, Lasse Heje, 2022, Game on: Social networks and markets, *Journal of Financial Economics*

146, 1097–1119.

Puri, Indira, 2025, Simplicity and risk, *The Journal of Finance* 80, 1029–1080.

Schmitt, Marc, 2026, Algorithmic markets, Working paper, University of Oxford.

Schuirmann, Donald J., 1987, A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability, *Journal of Pharmacokinetics and Biopharmaceutics* 15, 657–680.

Shiller, Robert J., 2017, Narrative economics, *American Economic Review* 107, 967–1004.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

Sun, Yeneng, 2006, The exact law of large numbers via Fubini extension and characterization of insurable risks, *Journal of Economic Theory* 126, 31–69.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *The Journal of Finance* 62, 1139–1168.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin, 2017, Attention is all you need, in *Advances in Neural Information Processing Systems*, vol. 30, 5998–6008.

Yang, Hao, 2024, AI coordination and self-fulfilling financial crises, Working paper.

Zhang, Lihong, and Xiaoquan (Michael) Zhang, 2024, Mispricing and algorithm trading, *Information Systems Research*, forthcoming.

A. Internet Appendix

This Internet Appendix collects the per-horizon equivalence and power detail behind Table 5, the detail behind the illustrative episodes of Section IV.F, and two checks confirming that the Reflexivity Index summarizes the market's feedback state and does not incrementally forecast extremes.

Per-horizon equivalence and power detail

Table 6 reports, for each design and horizon, the amplification effect in percent of unconditional realised volatility with its 95% confidence interval, the minimum detectable effect at 80% and 90% power, and the two-one-sided-test verdict at each margin. The aggregate design is reported for completeness and read as suggestive.

Table 6: Per-horizon equivalence and power detail

This table reports per-horizon amplification effects, confidence intervals, and minimum detectable effects (MDE), all in percent of unconditional realised volatility. TOST columns report whether the two-one-sided test rejects amplification of at least the stated margin at the 5% level. Sample sizes per horizon: aggregate design 456–467 months; cross-section 244,131–249,093 stock-months; Russell reconstitution boundary panel as in the descriptive statistics.

h	Effect	95% CI		MDE_{80}	MDE_{90}	TOST 5%	TOST 10%	TOST 20%
Panel A. Aggregate local projection (suggestive)								
1	-0.22	[-6.61,	6.17]	9.13	10.57	–	reject	reject
3	0.41	[-12.13,	12.95]	17.92	20.73	–	–	reject
6	14.69	[-4.24,	33.62]	27.06	31.31	–	–	–
12	46.05	[7.72,	84.37]	54.78	63.39	–	–	–
Panel B. Within-date cross-section								
1	-0.11	[-0.93,	0.71]	1.17	1.36	reject	reject	reject
3	-0.28	[-1.07,	0.50]	1.13	1.30	reject	reject	reject
6	0.05	[-0.84,	0.94]	1.27	1.47	reject	reject	reject
12	-0.05	[-0.86,	0.76]	1.16	1.34	reject	reject	reject
Panel C. Russell reconstitution								
1	0.19	[-0.37,	0.75]	0.80	0.93	reject	reject	reject
3	0.42	[-0.13,	0.97]	0.79	0.91	reject	reject	reject
6	-0.22	[-0.96,	0.52]	1.06	1.22	reject	reject	reject
12	-0.37	[-1.07,	0.33]	1.00	1.16	reject	reject	reject

Illustrative episodes: detail

As Kirilenko et al. (2017) document, the 2010 Flash Crash provides an illustrative market-side episode consistent with the reflexive feedback mechanism. Within minutes, interacting execution algorithms withdrew liquidity, amplified order-book imbalances, and generated a self-reinforcing price collapse across equities and futures. The episode is widely interpreted as involving rapid liquidity withdrawal and feedback among automated trading systems, which makes it a useful illustration of the market-reflexivity channel (α), in which endogenous synchronization among trading algorithms produces a transient instability that rapidly unwinds once liquidity returns.

Information-side amplification (δ) implies that fluctuations in the informational environment (news intensity, sentiment, and narrative diffusion) should co-move with realized volatility. As emphasized by Shiller (2017), contagious economic narratives shape expectations and investment behavior. In modern algorithmic markets, these narratives are processed and acted upon by machines

as well as people: sentiment analysis, natural-language processing, and news-ranking algorithms convert qualitative narratives into quantitative trading signals, creating feedback loops between the informational environment and market dynamics. Bybee et al. (2024) show that fluctuations in business-news volume and tone track macroeconomic cycles and forecast aggregate risk premia; Aleti and Bollerslev (2025) demonstrate that variation in the informational environment translates into volatility in asset-pricing kernels; and Hirshleifer et al. (2025) document how news diffuses through social networks and synchronizes investor attention across markets. The informational layer therefore possesses its own propagation dynamics: as narrative intensity and algorithmic amplification rise, informational feedback loops can reinforce price movements even when fundamentals remain unchanged.

As Pedersen (2022) show, the 2021 GameStop short squeeze provides a vivid illustration of information-side reflexivity. A narrative originating in retail discussion on platforms such as Reddit's WallStreetBets was propagated and amplified by the algorithmic information layer into a self-reinforcing feedback loop with market behavior. Viral posts and their amplification across news feeds and trading apps transformed attention itself into a tradable signal. As prices rose, the movement re-entered ranking, recommendation, and sentiment models as fresh input, and the algorithms conditioning on them accelerated the momentum. Recursive amplification between the informational layer and the market layer produced extreme price dynamics despite the absence of corresponding fundamental news; the reflexivity is a property of the algorithmic propagation of the narrative, with the human narrative entering as the input the information layer amplifies.

Three mechanisms link narrative diffusion to volatility on the information side: (i) cross-platform attention cascades, where trending stories are boosted by ranking and recommendation systems; (ii) recommender and news-feed amplification, which increases the persistence and salience of particular narratives; and (iii) model homogeneity in sentiment and NLP pipelines, which aligns informational inputs across institutions and induces correlated expectations. These processes correspond directly to the information-reflexivity parameter δ.

Conditional exceedances

A qualitative implication of the model concerns the clustering of *moderate* volatility-scaled shocks: as the reflexivity parameter approaches the stability boundary, shocks to the volatility state propagate more persistently, so a high-reflexivity month should contain more frequent exceedances of an ex-ante volatility threshold at moderate magnitudes (equation (23)). I assess this with a conditional-exceedance test that normalizes returns by an ex-ante volatility forecast and isolates whatever content implied feedback carries beyond what standard volatility dynamics already anticipate. The test is

predictive (lead–lag), measuring reflexivity at month t and relating it to outcomes realized in month $t + 1$, which avoids mechanical co-movement from overlapping rolling windows. For each month $t + 1$, let $\widehat{\sigma}_{t+1|t}$ denote a forecast of daily volatility formed at month t from a HAR-style specification, and define the exceedance count

$$\mathrm{Exc}_{k,t+1} \equiv \sum_{d \in t+1} \mathbf{1}\{|r_d| > k\, \widehat{\sigma}_{t+1|t}\}. \tag{26}$$

On the sample for which an ex-ante threshold is available, the exceedance content of implied feedback is fully embodied in the volatility forecast: the $k = 2$ coefficient on IF_t is positive but small, and the far-tail ($k = 3$) coefficient is close to zero. Because IF_t is itself a transform of volatility persistence, an ex-ante HAR forecast built from the realized-volatility history already carries its conditional content. The moderate-shock clustering the model implies therefore operates through volatility persistence, which a standard volatility forecast already captures.

Point forecast of extreme returns

Implied feedback also predicts the next month's largest daily return, $X_{t+1} \equiv \max_{d \in t+1} |r_d|$, in univariate form, with a positive and significant slope (Table 7, column 1). The relationship largely overlaps with standard volatility dynamics: adding the lagged level of $\log RV_t$ and HAR-style components (Corsi, 2009) drives the coefficient on IF_t from 0.49 to 0.16 and then to 0.05, losing significance, while the R^2 rises from 0.05 to about 0.38 (columns 1–3). Because a HAR forecast built from the realized-volatility history absorbs most of its predictive content, I treat the Reflexivity Index as a diagnostic of the feedback state, and not as an incremental point forecaster of extreme returns. Its informational content resides in the slow-moving secular drift of the feedback state, so the full-sample relationship reflects the secular co-movement of reflexivity and exceedance propensity over 1980–2024.

Table 8 reports the co-movement of implied feedback with the Kritzman et al. (2011) absorption ratio, the trading-layer commonality check discussed in Section IV.C.

Table 7: Point-forecast horse race for next-month extreme returns

This table reports the point-forecast horse race for the next-month extreme return. The dependent variable is $\log X_{t+1}$, $X_{t+1} = \max_{d \in t+1} |r_d|$. Columns (1)–(3) use the full 1982–2024 sample; column (4) adds the month-t level of $\log VIX_t$ and is therefore estimated on the 1990–2024 subsample. As with the HAR controls, option-implied volatility absorbs most of IF_t's point-forecast content, consistent with treating IF_t as a diagnostic of the feedback state. Newey–West (HAC) standard errors with 12 lags in parentheses.

	(1)	(2)	(3)	(4)
IF_t	0.488 (0.206)	0.162 (0.093)	0.049 (0.086)	0.090 (0.078)
$\log RV_t$		0.577 (0.046)	0.305 (0.080)	0.213 (0.095)
HAR($\log RV$): 3m			0.312 (0.095)	0.108 (0.090)
HAR($\log RV$): 12m			0.102 (0.065)	0.039 (0.071)
$\log VIX_t$				0.514 (0.119)
Constant	-4.143 (0.076)	-2.829 (0.109)	-2.511 (0.100)	-4.752 (0.547)
N	517	517	517	419
R^2	0.050	0.347	0.376	0.445

Table 8: Absorption ratio and implied feedback

This table reports the co-movement of implied feedback IF_t with the Kritzman et al. (2011) absorption ratio, the share of the top-500 large-cap return cross-section explained by its top five principal components on a trailing 24-month daily window, 1982–2024. Levels reflect a shared secular rise; first differences and the linear-detrended correlation isolate common higher-frequency movement, so the link is more than a common trend. As an *outcome* of synchronized trading, the absorption ratio offers suggestive fingerprint evidence on trading-layer commonality (ρ_M), with the magnitude measured in the cross-section of Section IV.E. Newey–West (HAC) standard errors (12 lags) in parentheses; $^{***}/^{**}/^{*}$ denote $p < 0.01/0.05/0.10$.

	Levels	First differences
Correlation with IF_t	0.61	0.21
(linear-detrended)	0.51	
Regression $IF_t = a + b\, AR_t$		
b (HAC s.e.)	1.175*** (0.156)	0.669*** (0.196)
R^2	0.368	0.045
N	518	517