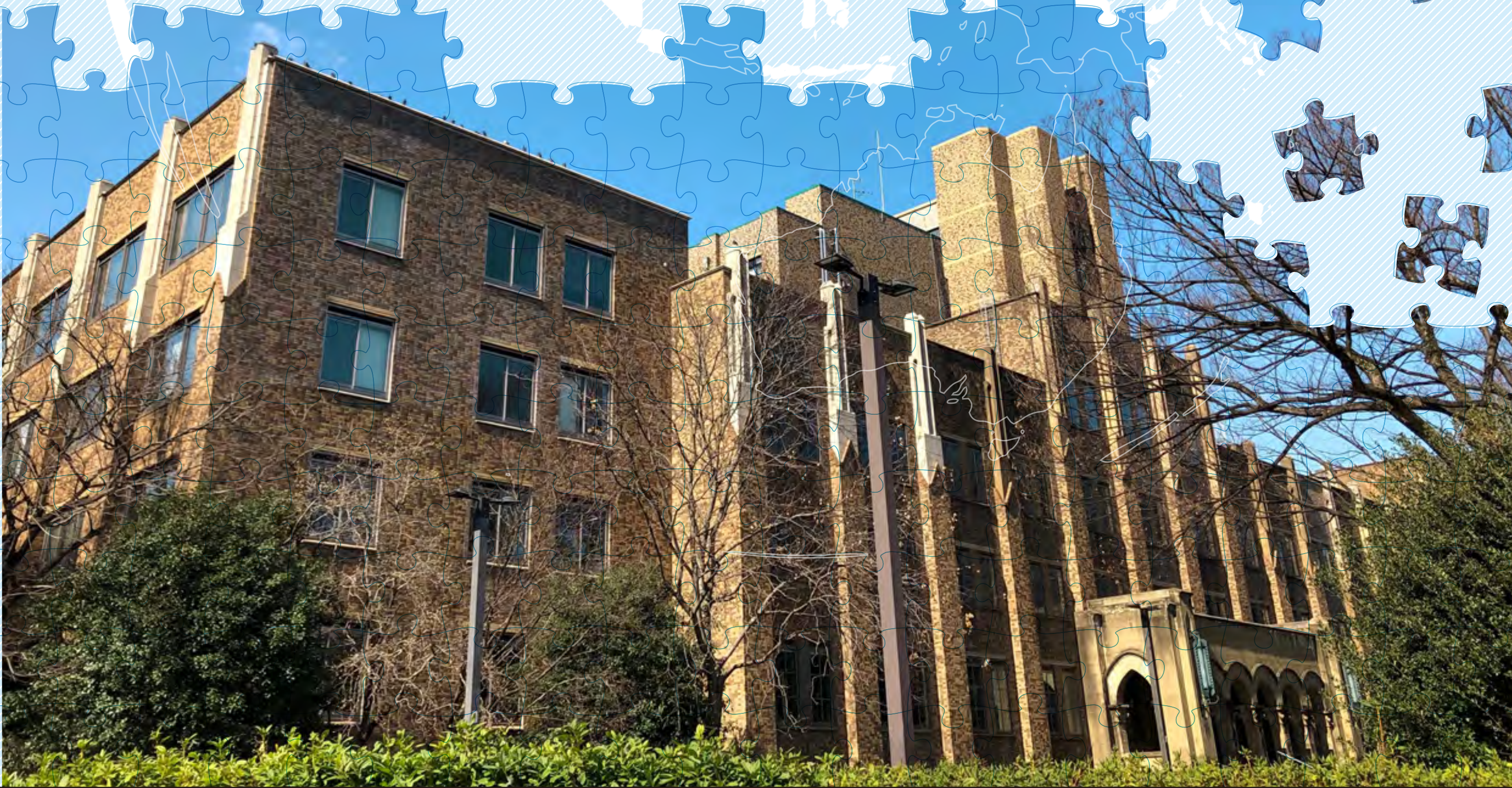
Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo

CSRDA supports the Sustainable Development Goals

CSRDA Discussion Paper

Diagnosing Industry Network Resilience under Financial Uncertainty: An Auditable Public-Data Framework Using Fama–French 49 Industry Portfolios



No.	Date	SDGs
140	September. 2026	8 DECENT WORK AND ECONOMIC GROWTH; 9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

Name Etsusaku Shimada, Hiroto Imai

Diagnosing Industry Network Resilience under Financial Uncertainty

An Auditable Public-Data Framework Using Fama–French 49 Industry Portfolios

Etsusaku Shimada[1] Hiroto Imai[1]

[1]Faculty of Policy Studies, Iwate Prefectural University, 152–52 Sugo, Takizawa-shi, Iwate 020–0693, Japan

E-mail addresses: etsusaku@gmail.com (E. Shimada); imai.hiroto.0411@gmail.com (H. Imai).

Version date: August 19, 2026

Abstract

Financial connectedness is not a sufficient statistic for fragility. Under the pre-specified centrality-based rule, a state-specific graph induces a model-implied allocation of common response, making diagnostic resilience a property of the topology–loading pair rather than of topology alone. We develop an auditable state-conditioned diagnostic using Fama–French 49 industry returns. A fixed 58-month benchmark verifies internal coherence; a recovered and cryptographically hashed 375-month panel supports VIX-state inference, 2,000 moving-block-bootstrap replications, exploratory counterfactual decomposition, and rolling validation. At the primary threshold, High-VIX density rises from 0.253 to 0.849, yet state-specific model-implied loadings yield a larger stability margin, a higher local boundary, and a smaller diffuse cumulative response. Holding the pooled raw-network loading fixed removes the raw resilience gap, whereas market-factor residualization reveals a denser and less resilient High-VIX residual topology under that fixed loading. Within the diagnostic, the reversal is therefore attributed to the more diffuse centrality-based loading induced by the High-VIX graph. Rolling associations with subsequent volatility, industry dispersion, and drawdown magnitude are specification-sensitive, so the evidence does not support a forecasting claim. The contribution is an empirically implementable topology–loading decomposition embedded in a fully traceable public-data architecture: a measurement discipline and auditable design, not a structural contagion model, causal design, or universal forecasting rule.

Keywords: financial uncertainty; industry-return networks; VIX; block bootstrap; stability margin; reproducibility

JEL codes: C32; C38; C45; G10; G17

1 Introduction

Fragility is not an attribute of a network graph in isolation. It is a property of the graph together with the loading rule through which a disturbance enters and the rule by which responses propagate. Under the pre-specified centrality-based rule used here, a state-specific graph induces a state-specific model-implied allocation of common response. A market state can therefore change both the links and the diagnostic loading, allowing density to rise while measured resilience also rises and connectedness alone to misrank stress vulnerability.

The paper turns this conceptual distinction into a sharply delimited empirical question: when financial uncertainty reshapes an observed industry-return network, does its capacity to absorb a normalized common-response stress deteriorate in the way density alone would suggest? The object is diagnostic rather than structural. A correlation network is not a balance-sheet network, and a state comparison is not a causal shock design. We therefore neither interpret links as contractual exposures nor identify causal spillovers or reliance on a particular forecasting technology. Instead, we construct a transparent reduced system in which observed topology and the model-implied concentration of normalized common response can be varied and interpreted separately.

To answer that question without conflating calibration, inference, prediction, and audit, the empirical architecture is deliberately layered. A fixed August 2021–May 2026 snapshot first establishes internal coherence across correlation thresholds, response intensities, central, peripheral, and diffuse shock locations, and alternative centrality rules. A recovered January 1990–March 2021 panel then supports Low- and High-VIX comparisons with 2,000 twelve-month circular moving-block-bootstrap replications. Rolling 120- and 60-month windows test the stronger proposition that the diagnostic contains stable information about risk over the following year. Finally, versioned inputs, code, draws, intermediate objects, and manuscript values close the evidentiary chain from public data to reported claim.

The central result is a density–fragility reversal, not merely a state difference. High-VIX months generate a much denser positive-correlation network: density rises from 0.253 to 0.849. Yet, under state-specific model-implied centrality-based loadings, the High-VIX system has a larger stability margin, a higher model-implied local boundary, and a smaller diffuse cumulative response. The corresponding High-minus-Low differences are 0.037 for the margin, 0.041 for $\beta^\star$, and -15.196 for diffuse-shock CSR. Their 95% moving-block-bootstrap intervals exclude zero, so three pre-specified fragility predictions reverse rather than merely attenuate.

An exploratory counterfactual decomposition makes the reversal interpretable. High uncertainty adds links, but it also makes eigenvector centrality less concentrated: the HHI of q falls from 0.0295 to 0.0214. Fixing the model-implied common-response loading at pooled raw-network values eliminates the apparent Low–High resilience gap; removing the market factor while holding that loading fixed reveals a significantly less resilient High-VIX residual topology. Within the diagnostic, the same market state therefore combines two offsetting channels: a denser and more fragile residual topology under the fixed raw loading, and a more diffuse state-specific centrality-based loading that offsets it in the raw specification.

The paper's contribution is best understood as a measurement and evidentiary advance. Conceptually, it operationalizes the topology–loading pair as the relevant object for cross-state diagnostic rankings and exposes the estimand embedded in those rankings: what is allowed to change, and what is held fixed, when resilience is compared. Empirically, it documents a state-dependent density–fragility reversal in public industry-return data and offers an explicitly exploratory attribution through controlled topology–loading counterfactuals. Econometrically, it propagates sampling uncertainty through the full network-construction pipeline, separates broad market co-movement from residual topology, and treats specification-sensitive predictive results as an evidentiary boundary rather than as a result-selection problem. Computationally, it makes auditability part of the contribution by separating the recent benchmark from the recovered historical extension, preserving executable transformations, and synchronizing reported results to canonical reruns.

The broader significance extends to any environment in which a common signal can align actions across an already connected system. Algorithmic forecasts are one salient example, but the present study does not observe AI adoption, forecast quality, or endogenous deployment. Its claim is deliberately more general: topology records where empirical links are observed, whereas diagnostic loading records the model-implied allocation through which a normalized common response enters the system under the stated rule. Their joint movement, not either object alone, determines measured resilience.

The remainder of the paper proceeds as follows. Section 2 positions the diagnostic relative to structural and empirical network research, uncertainty measurement, common signals, predictive validation, and computational reproducibility. Section 3 explains the deliberately separate samples and the recovered historical provenance. Section 4 defines the network, dynamics, state contrasts, decomposition, and rolling design. Sections 5–8 report the benchmark, state evidence, decomposition, and predictive results. Section 9 documents the audit, scope, and limitations. Section 10 concludes. The appendices provide threshold sensitivity, boundary cases, predictive details, and audit verification.

2 Conceptual and Empirical Positioning

The relevant literatures supply complementary pieces of the problem but do not, by themselves, identify the object measured here. Structural network models explain propagation mechanisms; empirical connectedness reconstructs co-movement; uncertainty measures classify states; common-signal models motivate coordinated response; and reproducibility establishes traceability. The contribution is to combine these ingredients without collapsing their distinct inferential roles.

2.1 Structural networks and empirical co-movement

The first boundary separates structural exposure networks from empirical co-movement graphs. Structural financial-network models show that connectivity can diversify exposures in some regions of the state space and amplify shocks in others (Allen and Gale, 2000; Elliott et al., 2014; Acemoglu et al., 2015; Glasserman and Young, 2016). Related work on endogenously formed supply networks reinforces the broader point that network structure and fragility need not move monotonically together (Elliott et al., 2022). For the present study, the implication is precise: density is not intrinsically stabilizing or destabilizing because its effect depends on the surrounding exposure, leverage, response, and shock environment.

The empirical literature answers a different but complementary question: how does dependence observed in market series evolve over time? Variance-decomposition networks and related econometric measures reveal time-varying relationships among financial entities (Billio et al., 2012; Diebold and Yilmaz, 2014). Our correlation graph belongs to this empirical category. It is observable, reproducible, and easy to audit, but it is not a structural map of claims, funding links, or causal contagion.

For state comparisons, this distinction is not semantic but inferential. Correlations can change mechanically when volatility changes, and tests based on conditional correlations require care under heteroskedasticity (Forbes and Rigobon, 2002). We therefore interpret the High–Low graphs as state-conditioned co-movement summaries, not as evidence that contractual contagion intensified. The narrower diagnostic question is conditional: given the observed graph and a common normalized stress rule, how do topology and loading jointly map into a reduced response?

2.2 Financial uncertainty and state dependence

The second boundary concerns what a VIX state does—and does not—identify. Financial uncertainty is multidimensional, and data-rich measures distinguish latent uncertainty from readily observed proxies (Jurado et al., 2015). VIX is useful for identifying financially salient states because it is option-implied and available at high frequency, but it combines expected-volatility and risk-aversion components rather than measuring a single primitive (Bekaert et al., 2013; Cboe Global Markets, 2026). We therefore use VIX as a transparent market-state variable, not as a sufficient measure of structural uncertainty.

Against this background, the closest empirical comparison is work that constructs dynamic industry uncertainty networks and relates them to the business cycle (Baruník et al., 2024). Our question is complementary rather than substitutive. We do not propose a new uncertainty measure; we hold the diagnostic rule fixed and ask whether observed Low- and High-VIX networks differ only in connectedness or also in the concentration of the response entering those networks.

2.3 Common signals and coordinated response

The third boundary links common information to coordinated response without claiming to estimate the underlying mechanism. Public-information models show that a widely observed signal can coordinate behavior even when its social value differs from its private value (Morris and Shin, 2002). Information design likewise emphasizes that the distribution of signals can shape actions (Kamenica and Gentzkow, 2011), while performative-prediction research formalizes the related idea that a deployed prediction may alter the environment in which it is evaluated (Perdomo et al., 2020).

That motivation disciplines, rather than expands, the interpretation of the model. The rank-one term is normalized and deliberately parsimonious; it does not estimate beliefs, welfare, equilibrium information use, AI adoption, or a deployment mechanism. Its purpose is to expose a measurement problem: if a state changes both the graph and the concentration of common response, a single connectedness statistic cannot reveal which channel drives the resulting stress outcome.

2.4 Validation, negative results, and reproducibility

The final boundary concerns evidentiary strength: internal coherence, sampling uncertainty, decomposition, predictive content, and reproducibility answer different questions. Threshold and centrality checks evaluate construction sensitivity. Moving-block resampling quantifies uncertainty while preserving local temporal dependence (Künsch, 1989). Market-factor residualization asks whether a market-wide component accounts for the apparent network result. Rolling regressions test whether the contemporaneous diagnostic has stable forward-looking content, using HAC inference for overlapping horizons (Newey and West, 1987). Related work shows that only some systemic-risk measures contain robust predictive information for subsequent adverse outcomes (Giglio et al., 2016).

Reproducibility, in turn, is not a substitute for structural validity; it is the condition for knowing exactly what the reduced-form evidence establishes. Reproducible research requires preservation of data transformations, code, computational environments, and the links between outputs and reported claims (Peng, 2011; Sandve et al., 2013). Here, auditability does not convert a correlation graph into a causal network. It makes the diagnostic, the surprising reversal, the decomposition, and the limits independently inspectable.

3 Data, Sample Construction, and Audit Trail

The empirical design is intentionally bifurcated because internal calibration and external state inference require different samples. Treating the recent snapshot and the historical panel as interchangeable would blur their analytical roles and compromise provenance; they are therefore kept separate from raw input to reported claim. The recent sample anchors the benchmark and its replication tests, whereas the longer panel supports state comparison, decomposition, and rolling validation.

Table 1: Empirical layers and audit roles

Layer	Fixed input	Analytical role	Verification
Recent benchmark	58 months × 49 industries, Aug. 2021–May 2026	Threshold, response-intensity, shock-location, centrality, and sub-window sensitivity	SHA-256 manifest; 47/47 integrity checks; 31/31 independent recalculation tests
Historical extension	375 months × 49 industries plus factors and VIX, Jan. 1990–Mar. 2021	State comparison, 2,000-replication bootstrap, decomposition, and rolling validation	Three raw-input hashes; end-to-end rerun; 2,000/2,000 finite draws; no-look-ahead audit
Manuscript layer	Canonical tables, figures, and reconciliation files	Reported estimates and interpretive claims	Automated output mapping and numerical reconciliation

Notes: The recent benchmark and historical extension are versioned separately. The current-official-vintage refresh is an additional sensitivity exercise and does not overwrite the recovered historical inputs underlying the reported analysis.

3.1 Recent benchmark snapshot

The recent snapshot is a calibration and verification sample, not an extension of the historical panel. It uses the Fama–French 49 Industry Portfolios, Average Value Weighted Returns–Monthly, from August 2021 through May 2026 (French, nd). The fixed input contains 58 balanced months for 49 industries, uses percentage returns without standardization, and treats −99.99, −999.0, and −999.99 as missing codes. No month is removed. This file is the canonical input for benchmark network construction, shock analysis, and replication tests.

Its compactness is an analytical feature rather than a limitation. The sample is not appended to the historical state panel; instead, it verifies that the reduced system responds monotonically to the scenario-intensity parameter, remains below the local boundary over the studied range, and produces coherent differences across shock locations and network thresholds before any external state comparison is attempted.

3.2 Historical uncertainty-state sample and provenance recovery

The long panel supplies the temporal variation required for state-conditioned inference and rolling validation. It covers January 1990 through March 2021, yielding 375 months of value-weighted Fama–French 49 industry returns, Fama–French market and risk-free factors, and daily VIX closes aggregated by arithmetic monthly mean (French, nd; Cboe Global Markets, 2026). Industry ordering is checked against the recent benchmark snapshot.

Because the paper's claims depend on a historical data vintage, provenance is part of the evidentiary design. The reported panel is not represented as the current French Data Library vintage; it is the fixed historical input underlying the analysis. During the final audit, the three long-sample inputs were recovered from immutable archival copies of the upstream official public series, restricted to the reported sample, and fixed by SHA-256 hashes. The archival copies serve as transport records, while the Kenneth R. French Data Library and Cboe remain the authoritative upstream providers. The recovered inputs were then rerun end to end, and every manuscript value was reconciled to the canonical outputs.

Version discipline is therefore substantive, not clerical. Portfolio histories can change when upstream databases or portfolio-construction files are revised, so the package separates (i) exact reproduction of the recovered reported vintage from (ii) a current-official-vintage sensitivity script. The latter downloads the current official files in an internet-enabled environment and writes them to a separate versioned directory; it does not enter the canonical recovered-vintage estimates reported below.

3.3 VIX-state construction

The state rule is designed to be transparent, replicable, and intentionally non-structural. Low and High states are the bottom and top quartiles of monthly mean VIX, with cutoffs of 13.954 and 23.232. Each state contains 94 months, while the middle 187 months are excluded from the primary contrast. The cutoffs remain fixed at their original-sample values in every bootstrap replication. This design compares financially calm and highly uncertain months without imposing a latent regime-switching model.

That simplicity fixes the comparison but limits its interpretation. The classification does not assert that every High-VIX month is generated by the same mechanism, nor that VIX isolates a primitive uncertainty shock. Its purpose is to hold the state rule constant while reconstructing the corresponding industry-return networks and response loadings.

4 Diagnostic Framework

The object of interest is a response operator, not a network statistic in isolation. The framework therefore constructs an empirical topology, specifies how a normalized common response loads onto it, and evaluates the resulting local dynamics.

4.1 Network construction

The first step converts public return co-movement into a graph that can be rebuilt exactly. Let $R_t = (R_{1t}, \ldots, R_{nt})^\top$ denote the vector of monthly industry returns, with $n = 49$, and let $\widehat{\rho}_{ij,s}$ be the sample correlation between industries i and j in sample or state s. The recent benchmark uses an absolute-correlation adjacency,

$$\widetilde{W}^{\mathrm{abs}}_{ij,s}(\tau) = \mathbf{1}\{i \neq j\}\mathbf{1}\{|\widehat{\rho}_{ij,s}| > \tau\},$$

whereas the primary state comparison uses a positive-correlation adjacency,

$$\widetilde{W}^{+}_{ij,s}(\tau) = \mathbf{1}\{i \neq j\}\mathbf{1}\{\widehat{\rho}_{ij,s} > \tau\}.$$

At the examined state thresholds, the positive- and absolute-correlation rules coincide because no state-sample pairwise correlation is below -0.30. We nevertheless retain the positive-correlation definition as the primary state specification because negative co-movement is not treated as a common-direction transmission link.

The graph is then translated into two auditable objects: propagation weights and response concentration. Non-isolated rows are normalized to one, while isolated rows remain zero, so

$W_s(\tau)$ is row-normalized and substochastic. Network density is

$$D_s(\tau) = \frac{\sum_{i \neq j} \widetilde{W}_{ij,s}(\tau)}{n(n-1)}.$$

Let $q_s(\tau)$ be the nonnegative leading-eigenvector centrality of the binary adjacency, normalized so that $\mathbf{1}_n^\top q_s = 1$.

4.2 Reduced dynamics, local boundary, and cumulative response

The second step maps the empirical graph into a reduced dynamic system. Define

$$A_s(\tau) = 0.40I + 0.35W_s(\tau), \qquad b_s(\tau) = 0.5\mathbf{1}_n + 0.5nq_s(\tau),$$

so $b_s(\tau)$ is a model-implied centrality-based loading induced by the state-specific graph under the pre-specified rule; it is not an independently observed exposure vector. Define

$$J_s(\beta, \tau) = A_s(\tau) + \beta b_s(\tau) q_s(\tau)^\top.$$

The interpretation of β is deliberately scenario-based. It scales the normalized common-response channel but is not an estimate of observed AI use, coordinated trading, or a structural behavioral coefficient. The stability margin is $1 - \rho\{J_s(\beta, \tau)\}$, where $\rho(\cdot)$ denotes the spectral radius.

The rank-one structure also yields an analytically transparent local boundary. When $I - A_s$ is invertible, the continuation boundary has the closed form

$$\beta_s^\star(\tau) = \left[q_s(\tau)^\top \{I - A_s(\tau)\}^{-1} b_s(\tau)\right]^{-1}.$$

This is a model-implied local boundary under the chosen normalization, not a universal market threshold.

To compare states on a common scale, the primary disturbance is diffuse and unit-normalized:

$$u = \frac{\mathbf{1}_n}{\sqrt{n}},$$

so $\|u\|_2 = 1$ and its magnitude is identical across states. Starting from zero, the state receives u at $t_0 = 10$ and then evolves under J_s through $T = 100$. The finite-horizon cumulative shock response is

$$\mathrm{CSR}_s = \sum_{t=t_0}^{T} \|x_{s,t}\|_2.$$

Convergence time is the first post-shock index at which the state norm falls below 5% of its one-step post-shock norm. If no such index occurs by the horizon, the case is recorded as right-censored rather than treated as ordinarily converged.

4.3 Pre-specified state contrasts and block-bootstrap inference

Before examining any decomposition, the empirical conjecture is encoded in four directional contrasts. At the primary threshold $\tau = 0.40$ and intensity $\beta = 0.18$, the High-minus-Low quantities are density (F01), stability margin (F02), $\beta^{\star}$ (F03), and diffuse-shock CSR (F04), with predicted signs positive, negative, negative, and positive. The first prediction says uncertainty densifies the graph; the remaining three encode the conjecture that the High-VIX system is more fragile.

Sampling uncertainty is propagated through the network-construction pipeline rather than attached only to final coefficients. We use 2,000 twelve-month circular moving blocks, resampling industry returns, VIX, and market factors in common while fixing the state cutoffs. Within each replication, the state-specific specification re-estimates the networks, centralities, model-implied loadings, market residuals, and diagnostic outcomes. The exploratory fixed-loading comparisons condition, by construction, on the pooled raw-network q and b estimated once from the canonical original sample. Their reported intervals are therefore conditional moving-block-bootstrap intervals for the corresponding fixed-loading counterfactuals. Percentile intervals are reported. The procedure preserves local dependence without claiming that a twelve-month block is uniquely correct.

4.4 Topology–loading decomposition

Because the raw state contrast lets topology and loading move together, it cannot reveal which object drives a sign reversal. After observing the pre-specified reversal, we therefore use two explicitly exploratory comparisons to locate its empirical source. The pooled-fixed comparison holds q and b at values from the pooled raw-return network while allowing Low- and High-state topology to differ. The market-residual comparison removes the market factor from each industry's excess return, reconstructs state topology, and continues to hold the pooled raw loading fixed. These exercises attribute the reduced-form difference; they do not causally identify a structural mechanism.

4.5 Rolling predictive validation

The final analytical layer asks whether a valid contemporaneous diagnostic also earns an early-warning interpretation. Networks are reconstructed monthly from 120-month windows, with 60 months as a robustness choice, and the stability margin is related to the following twelve months' equal-weighted realized volatility, mean cross-industry dispersion, and maximum drawdown magnitude. The main predictor is the margin at $\tau = 0.40$ and $\beta = 0.18$. Newey–West standard errors with lag 11 accommodate the overlapping horizons. The exercise is predictive, not causal.

5 Benchmark Diagnostic on the Recent Snapshot

Before the diagnostic is asked to distinguish market states, it must pass an internal coherence test. The recent benchmark shows that higher thresholds reduce density and that, at every examined threshold, increasing β from zero to 0.18 raises $\rho(J)$, narrows the stability margin, increases CSR, and lengthens convergence. All studied cases remain below the local boundary. Eigenvector- and degree-based centrality produce the same qualitative ordering, and overlapping sub-windows preserve the monotone response to β.

Table 2: Recent-snapshot benchmark diagnostic at $\beta = 0.18$

τ	Edges	Density	$\rho(J)$	Margin	Central CSR	Conv.	$\beta^\star$
0.30	979	0.832	0.934	0.066	3.722	31	0.245
0.40	815	0.693	0.939	0.061	4.143	35	0.238
0.50	631	0.537	0.949	0.051	5.110	44	0.227

Notes: The fixed recent snapshot contains 58 monthly observations from August 2021 through May 2026. The central shock targets the leading-eigenvector-centrality industry. All reported cases remain below the local boundary.

Table 3: Benchmark shock-location sensitivity at $\beta = 0.18$

τ	Shock	CSR	Convergence	Target
0.30	Central	3.722	31	Ships
0.30	Peripheral	1.904	6	Coal
0.30	Diffuse	15.008	45	All industries
0.40	Central	4.143	35	Transportation
0.40	Peripheral	1.800	5	Coal
0.40	Diffuse	16.143	49	All industries
0.50	Central	5.110	44	Wholesale
0.50	Peripheral	1.759	5	Gold
0.50	Diffuse	18.598	58	All industries

Notes: Every shock has unit Euclidean norm. The spectral radius and margin depend on the diagnostic matrix, not on shock location; CSR and convergence depend on how the shock aligns with the network.

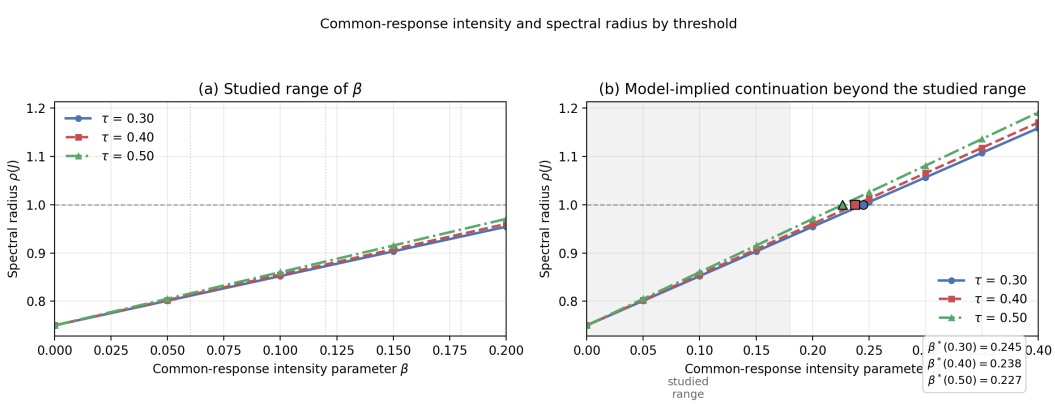


Figure 1: Benchmark sensitivity to common-response intensity. Across all three thresholds, a larger β moves the spectral radius toward the local boundary; line types and markers remain distinguishable in grayscale.

The benchmark delivers three substantive lessons, not merely a calibration table. First, common-response intensity moves the diagnostic in the theoretically coherent direction: the spectral radius approaches one, the margin contracts, cumulative response grows, and convergence slows. Second, thresholding changes the scale of the graph but not the direction of the response to β, so the qualitative conclusion is not tied to a single edge count within the examined range. Third, a unit-norm diffuse shock produces far larger cumulative responses than a central or peripheral shock. Broad exposure can therefore dominate a locally targeted disturbance even when the latter strikes a highly central industry; stress geometry is as consequential as connectedness.

Equally important, the benchmark establishes computational continuity between the recent sample and the historical extension. The same audited network, centrality, dynamic, and boundary functions are carried forward, while agreement across eigenvector- and degree-based centrality and overlapping sub-windows limits the possibility that the pattern is an artifact of one rule or endpoint choice. The package passes 47 integrity checks and 31 independent recalculation tests.

6 State-Conditioned Evidence under Financial Uncertainty

The state comparison is the paper's first out-of-benchmark historical test: does uncertainty alter only the amount of observed connectedness, or also the mapping from network structure to diagnostic resilience? At $\tau = 0.40$, High VIX raises positive-correlation density by 0.596, with a 95% interval of $[0.307, 0.717]$. Because each bootstrap draw resamples returns, VIX, and market factors jointly and reconstructs the networks, the interval reflects uncertainty in the full state-comparison procedure rather than treating the graph as fixed. The result identifies a pronounced state difference in empirical co-movement, not contractual contagion.

The diagnostically consequential finding is not densification itself but the reversal of the three fragility predictions. Despite its greater density, the High-VIX state has a larger margin, a higher $\beta^\star$, and a smaller finite-horizon diffuse CSR, and each reversal is statistically distinguishable from zero. At $\tau = 0.50$, the Low-VIX state is already beyond the local boundary at $\beta = 0.18$, so its large finite-horizon CSR is treated as a boundary result rather than an ordinary robustness estimate.

Table 4: VIX-state networks and state-specific diagnostics ($\tau = 0.40$)

State	Months	Edges	Density	Isolates	q HHI	$\rho(J)$	Margin	$\beta^{\star}$	Diffuse CSR
Low VIX	94	298	0.253	2	0.0295	0.972	0.028	0.203	30.204
High VIX	94	999	0.849	1	0.0214	0.934	0.066	0.244	15.008

Notes: The recovered fixed sample runs from January 1990 through March 2021. Low and High are the bottom and top VIX quartiles. Diagnostics use positive-correlation networks and the state-specific model-implied q and b, with $\beta = 0.18$ and a unit-norm diffuse shock. The Low-VIX diffuse response is right-censored at the simulation horizon; CSR remains a finite-horizon cumulative response.

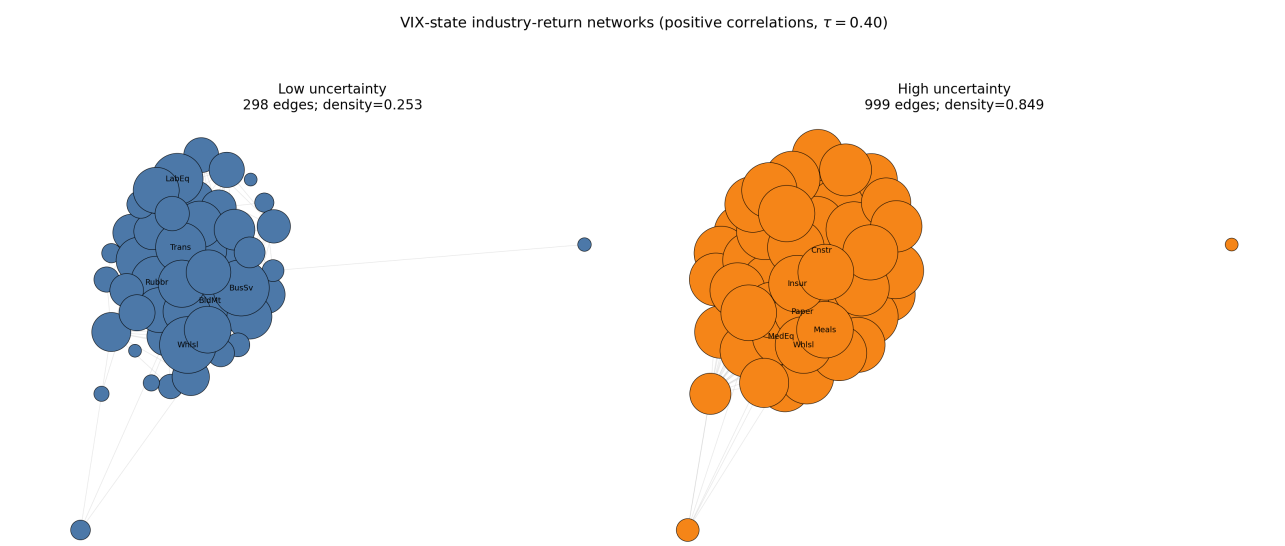


Figure 2: High uncertainty sharply densifies the positive-correlation industry network. Low- and High-VIX graphs use a common node layout; node size reflects eigenvector centrality.

Table 5: Pre-specified VIX-state contrasts and moving-block-bootstrap inference

ID	High-minus-Low quantity	Estimate	Bootstrap SE	95% interval	Expected	Result
F01	Network density	0.596	0.105	[0.307, 0.717]	Positive	Supported
F02	Stability margin at $\beta = 0.18$	0.037	0.010	[0.016, 0.054]	Negative	Reversed
F03	Model-implied boundary $\beta^\star$	0.041	0.010	[0.018, 0.056]	Negative	Reversed
F04	Diffuse-shock CSR at $\beta = 0.18$	-15.196	8.551	[-34.257, -4.360]	Positive	Reversed

Notes: Intervals are percentile intervals from 2,000 twelve-month circular moving-block-bootstrap replications. Holm-adjusted one-sided conclusions support only F01 in the originally specified directions. The sign reversals are reported rather than redesigned away.

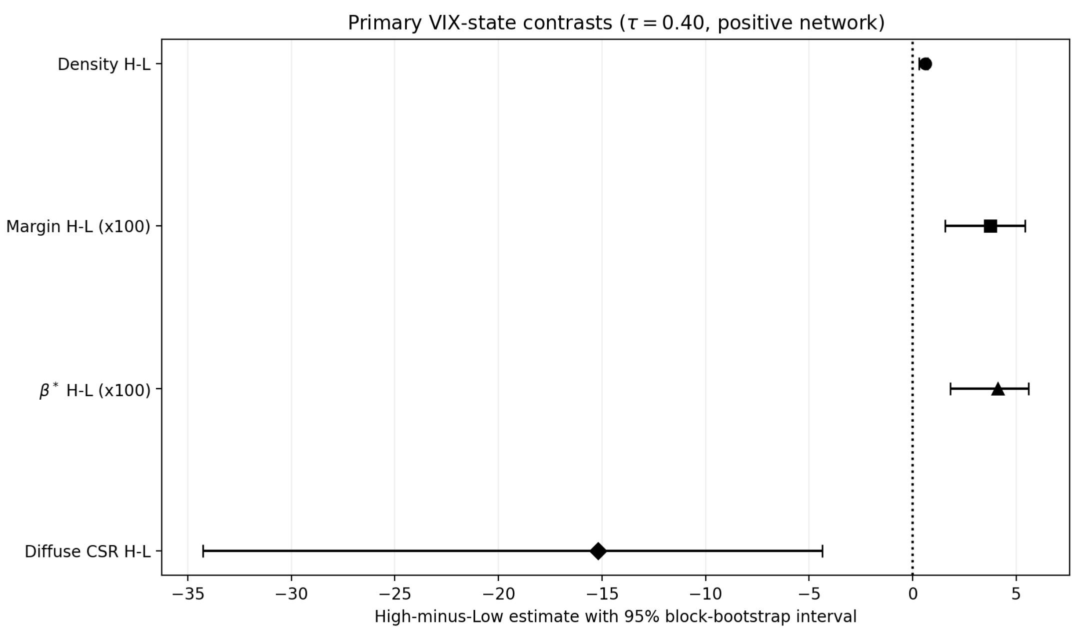


Figure 3: Only the density contrast moves in the pre-specified direction; the three fragility contrasts reverse. Error bars are 95% moving-block-bootstrap intervals. Margin and $\beta^\star$ are multiplied by 100 for display.

The reversal is therefore a diagnostic warning, not evidence that uncertainty stabilizes markets. The Low-VIX system lies close to the local boundary at the primary threshold and its diffuse trajectory is right-censored; more fundamentally, the state-specific model-implied loading changes with the graph. The result calls for decomposition rather than direct causal interpretation. Its empirical value is to reject the shortcut from "more connected" to "more fragile" when the state is allowed to alter both topology and the distribution of common response.

7 Decomposing the Density–Fragility Reversal

The reversal becomes interpretable through an exploratory counterfactual separation of the two state-varying objects. In High-VIX states, the HHI of q falls from 0.0295 to 0.0214, so the model-implied centrality-based loading becomes less concentrated precisely when the graph becomes denser. The raw contrast therefore combines topology and loading channels that move in opposite directions within the diagnostic.

Two counterfactual comparisons provide an empirical attribution of the reversal. Holding q and b fixed at pooled raw-network values removes the raw resilience difference: the margin, boundary, and CSR intervals all include zero. By contrast, after market-factor removal and with the same raw loading held fixed, the High-VIX residual topology has a smaller margin, a lower $\beta^{\star}$, and a larger diffuse response. The pre-specified fragility direction therefore appears in the fixed-loading market-residual comparison.

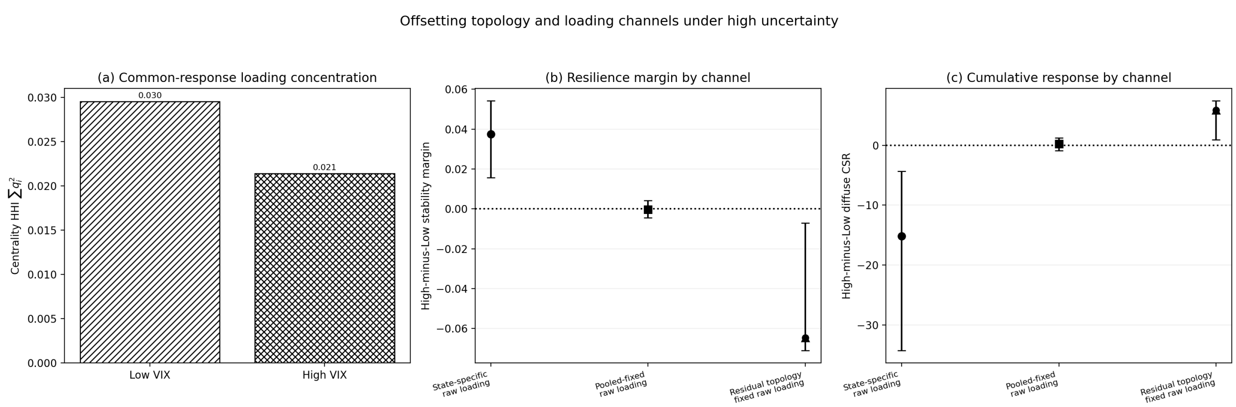


Figure 4: The exploratory counterfactual comparisons show offsetting topology and loading channels. Under the pre-specified rule, High uncertainty induces a more diffuse state-specific loading, while market-residual topology is less resilient under the fixed raw loading. Error bars are 95% moving-block-bootstrap intervals.

Table 6: Offsetting topology and common-response-loading channels

Comparison	Margin H–L	$\beta^\star$ H–L	Diffuse CSR H–L
Raw networks, state-specific loading	0.037 [0.016, 0.054]	0.041 [0.018, 0.056]	-15.196 [-34.257, -4.360]
Raw networks, pooled fixed loading	-0.000 [-0.005, 0.004]	-0.000 [-0.005, 0.005]	0.240 [-0.895, 1.220]
Market-residual topology, fixed raw loading	-0.065 [-0.071, -0.007]	-0.098 [-0.106, -0.010]	5.851 [0.926, 7.394]

Notes: The first row is the pre-specified state-specific specification. The second holds q and b fixed at pooled raw-network values and changes only raw-return topology. The third removes the market factor and holds the same raw loading fixed. Brackets report 95% moving-block-bootstrap intervals. The decomposition was specified after the pre-specified sign reversal and is therefore labelled exploratory. The fixed-loading intervals condition on the pooled raw-network q and b estimated from the canonical original sample; they are conditional counterfactual intervals.

Market-factor removal reveals a topology channel that the raw graph conceals. The residual networks are sparse, especially in Low-VIX months, and recomputing state-specific loadings on them places the diagnostic beyond its local boundary at $\beta = 0.18$. Those responses are therefore not treated as ordinary robustness estimates. The boundary-aware comparison fixes the raw loading; under that design, High VIX raises residual density by 0.073, lowers the margin by 0.065, lowers $\beta^\star$ by 0.098, and raises diffuse CSR by 5.851.

The exploratory counterfactual comparisons yield the paper's central interpretive result. Under the pre-specified loading rule, High uncertainty produces a denser raw-return graph and a more fragile market-residual topology, but the High-VIX graph also induces a more diffuse model-implied centrality-based loading. The pooled-loading comparison shows that raw topology alone does not generate the apparent resilience advantage, while the factor-residual comparison shows that the market-residual topological contrast points toward fragility under the fixed raw loading. The reversal is therefore an offset between diagnostic channels, not evidence that dense systems are inherently resilient.

8 Predictive Relevance and Its Limits

A contemporaneously valid diagnostic need not be a useful forecasting signal. The rolling exercise therefore tests a strictly stronger proposition: whether a lower current margin consistently precedes greater risk over the following year. It is a forward-looking validation check within the historical panel, not a causal interpretation of the state result.

The predictive estimates do not form a stable empirical law. With a 120-month window and both lagged volatility and density controls, a larger margin is associated with lower subsequent drawdown magnitude and marginally lower realized volatility. Removing density substantially attenuates those coefficients; only the dispersion association remains statistically distinguishable from zero. With a 60-month window, including density reverses the margin coefficient, whereas excluding density returns it toward zero. Variance-inflation factors near 67 for margin and density in the 120-month full-control specification indicate severe collinearity.

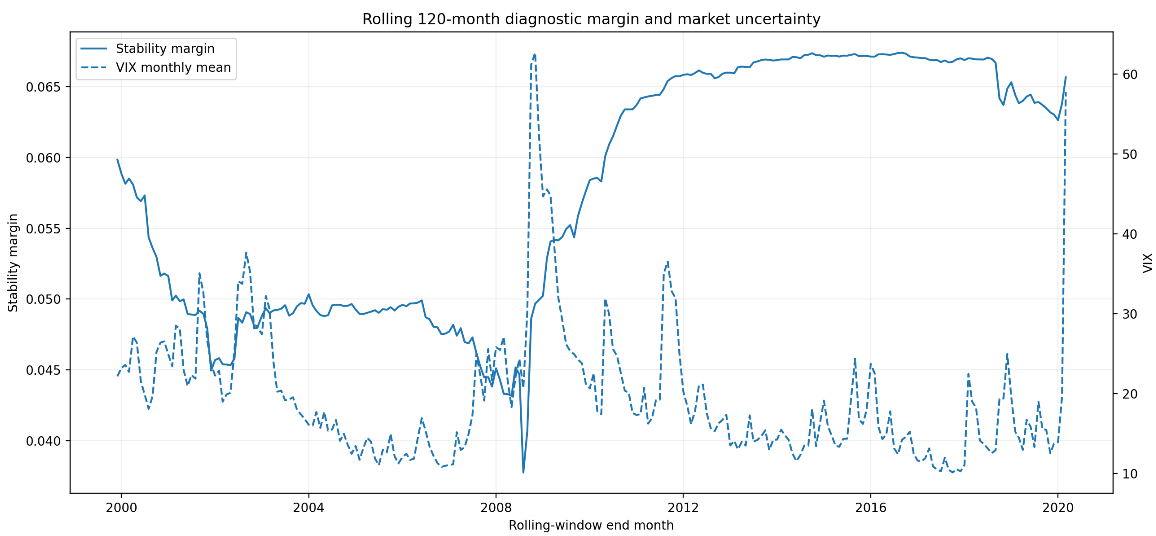


Figure 5: Rolling stability margin and financial uncertainty. The figure is descriptive; overlapping-horizon inference is reported in Table 7.

Table 7: Rolling predictive associations for the pre-specified outcomes and control sets

Window	Outcome	Coefficient	HAC SE	p-value	Controls
120	Realized volatility	-24.415	12.869	0.058	Lagged volatility and density
120	Industry dispersion	-1.239	2.949	0.674	Lagged volatility and density
120	Drawdown magnitude	-40.993	19.088	0.032	Lagged volatility and density
120	Realized volatility	-1.600	1.679	0.341	Lagged volatility only
120	Industry dispersion	-0.494	0.230	0.032	Lagged volatility only
120	Drawdown magnitude	-3.458	2.575	0.179	Lagged volatility only
60	Realized volatility	9.234	2.678	0.001	Lagged volatility and density
60	Industry dispersion	2.938	0.693	< 0.001	Lagged volatility and density
60	Realized volatility	-0.505	1.147	0.660	Lagged volatility only
60	Industry dispersion	-0.330	0.261	0.206	Lagged volatility only

Notes: The predictor is the stability margin, scaled per 0.01-unit increase. All combinations of the two primary outcomes, the 120- and 60-month windows, and the two pre-specified control sets are shown. Maximum drawdown magnitude is a secondary outcome and is reported for the 120-month window only. HAC standard errors use Newey–West lag 11.

The value of the predictive exercise is therefore delimitative. Sign changes across windows and control sets prevent a stable forecasting interpretation: the long-window full-control estimates are consistent with lower margins preceding worse drawdowns, but the relationship attenuates when density is removed, and the short-window full-control coefficients move in the opposite direction. The diagnostic varies meaningfully over time, yet the present design does not establish a robust stand-alone early-warning rule. This negative finding leaves the contemporaneous decomposition intact while preventing a stronger claim that the evidence does not support.

9 Reproducibility, Scope, and Limitations

A contribution built around a diagnostic must establish two things simultaneously: what the statistic reveals and the exact evidentiary conditions under which the claim survives. This section integrates the analytical layers, the audit trail, and the limitations into a single statement of scope.

9.1 What the analytical layers jointly establish

Taken together, the three empirical layers form a nested validation sequence rather than interchangeable robustness checks. The recent benchmark establishes internal coherence: the diagnostic responds monotonically to common-response intensity, remains directionally stable across examined construction choices, and distinguishes localized from diffuse stress. The state comparison establishes out-of-benchmark historical differentiation: financial uncertainty is associated with a sharply denser empirical network, but topology and model-implied response-loading concentration move in offsetting directions. The rolling exercise tests a stronger proposition and does not support it robustly. Together, the layers sustain a diagnostic and decompositional claim while ruling out a causal or general forecasting interpretation.

9.2 Audit trail and reproducibility

Reproducibility is substantive here because each headline result is the endpoint of an ordered transformation chain. The recent benchmark package stores the fixed snapshot, configuration, correlation and adjacency matrices, W, A, q, b, shock vectors, trajectories, tables, figures, logs, and a SHA-256 manifest. It passes 47 integrity checks and 31 independent recalculation tests.

The historical audit closes the same chain for the state-conditioned analysis. It stores the three recovered raw inputs, processed panels, state labels, market-residual returns, bootstrap draws, rolling diagnostics, model outputs, configuration, code, and reconciliation files. The canonical run contains 2,000 valid bootstrap replications with no nonfinite draw, 548 rolling-diagnostic rows—244 from the 120-month design and 304 from the 60-month design—16 stored predictive-model rows, and a no-look-ahead audit. Table 7 reports the 10 outcome–window–control combinations retained in the manuscript. The rerun reproduces the central state point estimates and synchronizes the ancillary intervals, decomposition values, and drawdown-magnitude regressions reported in this version.

Version control is the final element of that chain. The recovered historical vintage is sufficient for exact reproduction of the reported 1990–2021 analysis, while a separate script retrieves current official Kenneth French and Cboe files for a version-sensitivity exercise in an internet-enabled environment. Current-vintage files and outputs are written separately and cannot overwrite the recovered canonical archive.

9.3 Scope and limitations

The same design discipline that sharpens the contribution also narrows it. First, the correlation graph does not identify causal spillovers or contractual exposures. Second, VIX quartiles are not a structural regime model and VIX is not a pure latent-uncertainty measure. Third, the rank-one response term is intentionally parsimonious, and $\beta^\star$ is normalization-dependent. Fourth, the topology–loading decomposition is exploratory because it was motivated by the sign reversal, and its fixed-loading intervals are conditional on the pooled raw-network loading estimated from the canonical original sample. Fifth, state-specific market-residual loadings cross the local boundary at $\beta = 0.18$; the interpretable residual comparison therefore fixes the raw loading. Sixth, the rolling exercise is affected by overlapping outcomes, multicollinearity, and window sensitivity. Finally, a current-official-vintage rerun is a useful additional version check but is conceptually distinct from exact reproduction of the historical vintage used here.

These restrictions are constitutive of the claim rather than appended disclaimers. They mark the line between what an auditable public-data diagnostic can establish and what would require structural exposure data, an identified causal design, endogenous response modeling, or a new current-vintage empirical study.

10 Conclusion

This paper replaces a common shortcut with a sharper empirical object. Fragility is not a property of network density alone; under a specified response rule, diagnostic fragility is a property of the topology–loading pair. In U.S. industry returns, High VIX produces a much denser raw graph, while the pre-specified centrality rule induces a more diffuse state-specific model-implied loading. In explicitly exploratory fixed-loading comparisons, the raw resilience gap disappears and market-factor residualization reveals a less resilient High-VIX topology under the fixed raw loading.

The principal contribution is therefore neither a new connectedness index nor a claim that uncertainty stabilizes markets. The pre-specified state contrasts establish that connectedness and diagnostic fragility can move in opposite directions; the exploratory counterfactual decomposition then exposes the estimand implicit in cross-state network rankings: what is allowed to change, and what is held fixed, when fragility is compared. Together, these results convert an apparent empirical contradiction into a coherent measurement result and support the paper's central proposition—connectedness is an input into fragility, not a sufficient statistic for it.

The benchmark and predictive layers define the evidentiary boundary of that proposition. The benchmark shows that response intensity, thresholding, and shock geometry govern distinct dimensions of the diagnostic; the rolling analysis shows that contemporaneous differentiation does not automatically translate into a stable early-warning rule. The predictive non-result is therefore part of the evidentiary boundary: it separates a stable contemporaneous diagnostic result from a forecasting claim that the present evidence cannot sustain.

The broader academic implication is a discipline of joint reporting. Empirical financial-network research should pair network architecture with the stated loading rule and the resulting model-implied concentration of common response, especially when networks are reconstructed across states. Environments shaped by public narratives, common signals, or algorithmic outputs make this separation particularly consequential, although no specific coordination mechanism is required by the diagnostic. For reproducible social science, the corresponding lesson is methodological: an unexpected reversal becomes cumulative knowledge only when inputs, transformations, decompositions, negative results, corrections, and version histories remain auditable.

Author Contributions

Etsusaku Shimada conceived the study, developed the empirical strategy, implemented and executed the VIX-state, moving-block-bootstrap, market-factor-residual, topology–loading, and rolling-window analyses, conducted the long-sample provenance recovery and numerical reconciliation, and wrote and revised the manuscript. Hiroto Imai developed and audited the Fama–French data pipeline, benchmark network diagnostics, threshold and shock-location analyses, and the replication framework that supplied the computational foundation for the subsequent empirical work. Both authors reviewed the analytical specification and results.

Conflict of Interest

None declared.

Data Availability

The underlying industry-return and factor series are publicly available from the Kenneth R. French Data Library, and VIX data are publicly available from Cboe. The authors retain a manuscript-specific replication and provenance archive containing the fixed inputs, processed data, executable code, 2,000 bootstrap draws, rolling outputs, figures, and audit records underlying the reported results. The archive is available from the authors upon reasonable request. A separate current-official-vintage refresh script is also retained for versioned sensitivity analysis and writes current-vintage inputs and results separately from the recovered canonical archive.

A State Threshold and Edge-Rule Sensitivity

Robustness across thresholds is informative only when boundary crossings are distinguished from ordinary stable-system responses. Table 8 reports state-specific point estimates at $\tau \in \{0.30, 0.40, 0.50\}$. The density–loading tension is clearest at the primary threshold. At $\tau = 0.30$, High VIX remains denser and has a larger margin and $\beta^\star$, although the diffuse-CSR difference is smaller. At $\tau = 0.50$, the Low-VIX system is beyond the local boundary at $\beta = 0.18$, so its finite-horizon CSR is a boundary outcome rather than an ordinary robustness estimate. Positive- and absolute-correlation rules produce identical state graphs at these thresholds in the recovered sample.

Table 8: State-specific threshold sensitivity at $\beta = 0.18$

τ	State	Density	Margin	$\beta^\star$	Diffuse CSR	Boundary status
0.30	Low	0.488	0.052	0.227	18.533	Below
0.30	High	0.919	0.068	0.247	14.662	Below
0.40	Low	0.253	0.028	0.203	30.204	Below, censored
0.40	High	0.849	0.066	0.244	15.008	Below
0.50	Low	0.098	-0.044	0.153	831.527	Beyond
0.50	High	0.714	0.060	0.237	16.189	Below

Notes: Boundary status is evaluated at $\beta = 0.18$. "Censored" means the convergence criterion is not reached by the finite simulation horizon. The $\tau = 0.50$ Low-VIX CSR should not be interpreted as an ordinary stable-system response.

B Market-Residual Boundary Cases

Residualization sharpens the topology channel but also creates an admissibility problem. When q and b are recomputed from the very sparse residual networks, both Low- and High-state matrices cross the local boundary at $\beta = 0.18$. Those large finite-horizon responses are not used as ordinary robustness estimates. The fixed-raw-loading decomposition in Table 6 is the boundary-aware comparison used for interpretation.

C Predictive Design and Boundary Conditions

The predictive exercise is governed by two constraints: strict timing and local-boundary admissibility. The 120-month and 60-month specifications use 244 and 304 overlapping forecast origins, respectively. All raw-return rolling matrices remain below the local boundary at $\beta = 0.18$, whereas state-specific residual rolling matrices do not; residual predictive coefficients are therefore excluded from the central interpretation. The reported table includes every combination of the two primary outcomes, two windows, and two pre-specified control sets. Drawdown magnitude is secondary and is reported for the 120-month window.

Timing discipline is explicit. Each network uses returns available only through the window end, and each outcome is constructed from the subsequent twelve months. The overlap in outcomes motivates Newey–West lag 11. These choices address timing and serial correlation but do not eliminate multicollinearity between rolling density and margin.

D Audit and Manuscript-Output Verification

The audit closes the chain from raw input to published claim. It verifies not only that code executes, but also that sample dimensions, transformations, simulated draws, canonical outputs, and manuscript values remain mutually consistent.

The final evidentiary record is therefore both computational and documentary. The three recovered long-sample raw-file hashes are recorded in the provenance ledger and SHA-256 manifest, while the package preserves canonical output CSVs, all 2,000 bootstrap draws, clean-run logs, and the mapping from every manuscript table and figure to its source files.

Table 9: Selected reproducibility and integrity checks

Check	Result	Interpretation
Recent benchmark integrity checks	47/47 PASS	Sample, matrices, centrality, dynamics, files, and CSV round trips
Independent benchmark recalculation tests	31/31 PASS	Results recomputed from the fixed snapshot rather than copied from output CSVs
Historical raw-input hashes	PASS	FF49 returns, factors, and VIX match the provenance ledger
Historical sample dimensions	375 months, 49 industries	January 1990–March 2021 continuous panel
Bootstrap replications	2,000/2,000 valid	No nonfinite bootstrap output
Rolling no-look-ahead audit	PASS	Predictors use information through the window end; outcomes use subsequent months
Canonical state point estimates	PASS	Low/High state values regenerated from recovered inputs
Manuscript reconciliation	PASS after synchronization	Tables and figures use canonical rerun values in this version
Current-official-vintage refresh	Script provided	Separate version-sensitivity exercise; not part of recovered canonical estimates

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Baruník, J., Bevilacqua, M., and Faff, R. (2024). Dynamic industry uncertainty networks and the business cycle. *Journal of Economic Dynamics and Control*, 159:104793.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559.

Cboe Global Markets (2026). Cboe volatility index (VIX). Cboe Global Markets. Index overview and historical-data resources; accessed August 8, 2026.

Diebold, F. X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Elliott, M., Golub, B., and Jackson, M. O. (2014). Financial networks and contagion. *American Economic Review*, 104(10):3115–3153.

Elliott, M., Golub, B., and Leduc, M. V. (2022). Supply network formation and fragility. *American Economic Review*, 112(8):2701–2747.

Forbes, K. J. and Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance*, 57(5):2223–2261.

French, K. R. (n.d.). 49 industry portfolios. Kenneth R. French Data Library, Tuck School of Business, Dartmouth College. Accessed August 2, 2026.

Giglio, S., Kelly, B., and Pruitt, S. (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics*, 119(3):457–471.

Glasserman, P. and Young, H. P. (2016). Contagion in financial networks. *Journal of Economic Literature*, 54(3):779–831.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3):1217–1241.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060):1226–1227.

Perdomo, J. C., Zrnic, T., Mendler-Dünner, C., and Hardt, M. (2020). Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pages 7599–7609. PMLR.

Sandve, G. K., Nekrutenko, A., Taylor, J., and Hovig, E. (2013). Ten simple rules for reproducible computational research. *PLoS Computational Biology*, 9(10):e1003285.