Cross-Asset Lead-Lag Dynamics in the Tails

Evidence from Quantile Granger Causality

Kensley Blaise* Ibrahim Raheem† Agboola Yussuf‡

January 2026

Abstract

This study asks whether the risk linkages across U.S. equity sectors behave differently during market crashes and rallies than during normal trading. Using quantile Granger causality tests on 15 assets spanning all major sectors and key benchmarks, we find that cross-sector spillovers are overwhelmingly a tail phenomenon. Predictive linkages are strong during extreme market movements but virtually absent under normal conditions, producing a robust U-shaped pattern across the return distribution. The nature of these spillovers is regime-dependent. Crash-driven predictive linkages dominate during monetary-tightening periods, while symmetric tail spillovers emerge during crisis-recovery episodes. These results suggest that conventional correlation-based risk metrics, which rely on full-sample averages, may understate cross-sector vulnerability during extreme events.

Keywords: Quantile Granger causality, lead-lag relationships, tail risk, regime dependence
JEL codes: G12, G14, C32

*Corresponding author: k.blaise@uea.ac.uk, School of Economics, University of East Anglia (UEA), UK.
†School of Business, Southern Alberta Institute of Technology, Calgary, AB. Canada.
‡Department of Economics, University of Ilorin, Nigeria.

Introduction

Understanding how information spreads across equity sectors is a key focus of portfolio risk management, systemic risk monitoring and asset allocation strategy. Research documenting return predictability and liquidity driven price adjustments demonstrates that lead-lag relationships are common across assets and industries, which suggests that events in one sector can influence others, causing ripple effects (Lo and Mackinlay, 1990; Chordia and Swaminathan, 2000). This concept is supported by the gradual information diffusion hypothesis proposed by Hong and Stein (1999). Moreover, sectoral interdependence tends to intensify during periods of financial stress, reducing diversification benefits and increasing vulnerability (Diebold and Yilmaz, 2012; Rapach et al., 2013).

The standard approach to measuring lead-lag relationships between assets is Granger causality, which tests whether lagged returns of one asset help predict the conditional mean of another. While computationally convenient and well understood statistically, this methodology implicitly assumes homogeneous dynamics across the entire return distribution. As a result, it ignores critical features of financial markets, such as nonlinearities, volatility clustering, heterogeneity, asymmetry, and tail dependence (Raheem et al., 2024). Mean-based Granger causality cannot capture important dependence structures in the tails of the return distribution. Studies have shown that causal relationships vary significantly across quantiles, implying that observed relationships during normal times might not hold in extreme market conditions (Jeong et al., 2012; Balcilar et al., 2016). The absence of quantile-based causality analysis, therefore, represents a significant gap in the literature.

This study suggests using Quantile-based Granger Causality (QGC) to address this gap, as it allows testing for predictive relationships within particular parts of the return distribution. This enables a direct assessment of whether lead-lag relationships between sectors vary across bearish, median, and bullish return regimes. The application of quantile regression to the US sector Exchange Traded Funds (ETFs) dynamics is limited. Sector ETFs are the main tools for institutional and retail investors to express sector views, and their lead-lag relationships affect tactical allocation and hedging.

The objective of the study is to examine the lead-lag relationship among 15 U.S. assets, including 11 GICS sector ETFs and 4 cross-asset benchmarks, using a QGC framework. In addition, we seek to assess how sectoral spillover dynamics evolve across distinct market regimes spanning 2018-2023, including the pre-pandemic expansion, the COVID-19 crisis, the post-pandemic bull markets, and the period of monetary tightening. These episodes generate heterogeneous shocks across sectors; thus providing an ideal setting to examine how sectoral information transmission evolves across changing macro-financial environments.

We make two contributions to the literature. First, we extend the sectoral lead-lag literature by adopting a QGC framework. This allows for the identification of nonlinear and asymmetric transmission that are not captured by the conventional mean-based Granger causality framework. Second, we examine the sectoral evolution linkages across multiple market regimes (tranquil and turbulent eras), as previous studies usually focus on a particular regime/period (Bouri et al., 2021; Shahzad et al., 2021). The remainder of the paper is structured as follows. Section 2 discusses the data and methodology. Empirical results are presented in Section 3. Section 4 offers the conclusion and policy recommendations.

2. Data and Methodology

2.1 Sample construction

The empirical analysis uses daily returns for 15 assets across all 11 GICS equity sectors and 4 cross-asset benchmarks from January 2018 to December 2023. This sample period is chosen deliberately to encompass four distinct market regimes: the low-volatility expansion of 2018–2019, the COVID-19 crisis and recovery of 2020, the post-pandemic bull market of 2021, and the Federal Reserve tightening cycle of 2022–2023. The diversity of market conditions allows us to test whether tail causality patterns are regime-dependent. With 15 assets, we have 210 directed pairs (15×14), each tested at 19 quantiles ($\tau \in \{0.05, 0.10, \ldots, 0.95\}$) across 4 regimes, yielding 15,960 individual QGC tests. Table 1 describes the asset universe.

Table 1: Asset Universe: GICS Sector ETFs and Cross-Asset Benchmarks

Ticker	Description	Asset Class	Economic Role
Panel A: GICS Sector ETFs			
XLK	Technology Select Sector	Technology	Growth, innovation, rate sensitivity
XLF	Financial Select Sector	Financials	Credit conditions, yield curve
XLE	Energy Select Sector	Energy	Commodity prices, inflation
XLV	Health Care Select Sector	Health Care	Defensive, demographic trends
XLY	Consumer Discretionary Sector	Cons. Discr.	Consumer sentiment, cyclicality
XLP	Consumer Staples Sector	Cons. Staples	Defensive, inflation pass-through
XLI	Industrial Select Sector	Industrials	Capex, global trade, PMI
XLB	Materials Select Sector	Materials	Commodity demand, construction
XLRE	Real Estate Select Sector	Real Estate	Rates, credit, housing
XLU	Utilities Select Sector	Utilities	Bond proxy, rate sensitivity
XLC	Communication Services Sector	Comm. Services	Ad spending, consumer engagement
Panel B: Cross-Asset Benchmarks			
SPY	SPDR S&P 500 ETF	Equity Index	Broad market benchmark
VIX	CBOE Volatility Index	Volatility	Fear gauge, tail risk indicator
GLD	SPDR Gold Shares	Precious Metals	Safe-haven asset, inflation hedge
TLT	iShares 20+ Year Treasury	Long-Term Bonds	Duration risk, flight-to-quality

Source: Authors' computation.

Panel A of Table 1 shows the ETFs. Using sector-specific ETFs rather than individual stocks, we capture sector-level dynamics while avoiding idiosyncratic noise. Panel B adds four cross-asset benchmarks. SPY serves as the aggregate market, VIX for the volatility channel (Whaley, 2009), GLD for safe-haven dynamics (Baur and Lucey, 2010), and TLT for interest rate sensitivity and flight-to-quality effects.

Table 2 reports summary statistics for daily returns. All equity assets exhibit negative skewness and substantial excess kurtosis, confirming the presence of fat tails that motivate a quantile-based approach. VIX displays positive skewness and the highest volatility, consistent with its role as a tail-risk indicator. TLT is the only asset with slightly positive skewness, reflecting flight-to-quality dynamics.

Table 2: Summary Statistics of Daily Returns

Ticker	Mean	Std Dev	Skewness	Kurtosis
Panel A: GICS Sector ETFs				
XLK	0.0008	0.0171	−0.43	8.26
XLF	0.0003	0.0163	−0.57	14.02
XLE	0.0003	0.0225	−0.89	13.62
XLV	0.0004	0.0117	−0.39	10.67
XLY	0.0004	0.0159	−0.75	7.57
XLP	0.0004	0.0106	−0.49	15.75
XLI	0.0004	0.0146	−0.59	12.30
XLB	0.0004	0.0150	−0.49	9.05
XLRE	0.0003	0.0152	−1.10	17.22
XLU	0.0003	0.0140	−0.21	15.37
XLC	0.0003	0.0153	−0.53	6.23
Panel B: Cross-Asset Benchmarks				
SPY	0.0005	0.0131	−0.77	12.14
GLD	0.0003	0.0091	−0.27	3.25
TLT	−0.0001	0.0107	0.11	4.83
VIX	−0.0001	0.0771	1.13	4.25

Note: Daily returns over January 2018 – December 2023 (1,392 trading days). Kurtosis is reported as excess kurtosis (normal = 0). All equity assets exhibit negative skewness and excess kurtosis, indicating fat left tails.

2.2 Market Regime Classification

The sample is partitioned into four regimes based on macroeconomic conditions and key policy events. Regime boundaries are anchored to widely recognised dates: the onset of pandemic-related market stress in late February 2020, the beginning of the post-crisis rally in January 2021, and the start of the Federal Reserve tightening cycle in early 2022. We emphasise that these boundaries are defined retrospectively for analytical purposes, following standard practice in the regime-dependent spillover literature (Diebold and Yilmaz, 2012; Bouri et al., 2021), rather than as a real-time classification scheme. The average VIX values reported in Table 3 serve as descriptive context and play no role in the regime selection.

Table 3: Market Regime Definitions and Characteristics

Regime	Period	Days	Avg VIX	Characterization
Pre-COVID	Jan 2018 – Feb 2020	426	15.8	Low volatility expansion
COVID Crisis	Mar 2020 – Dec 2020	213	29.2	Extreme volatility, V-recovery
Recovery	Jan 2021 – Dec 2021	252	19.7	Strong rally, declining vol
Inflation Shock	Jan 2022 – Dec 2023	501	23.4	Fed tightening, rate uncertainty

Source: Authors' computation

2.3 Methodology

Standard Granger causality asks whether past returns of one asset help predict the average future return of another. However, average-based tests may miss important relationships that only appear during market crashes or rallies. To overcome this limitation, we adopt the QGC framework motivated by Jeong et al. (2012), which tests for predictability at specific points of the return distribution rather than just the mean.

The intuition is straightforward. For each directed pair $(X \to Y)$, we ask whether adding the past five days of X's returns improves predictions of Y's return at a given quantile τ, beyond what Y's own five lags can explain. The quantile τ refers to the distribution of the effect variable Y. At $\tau = 0.05$, the test targets Y's worst trading days. At $\tau = 0.95$, it targets Y's best. By scanning from $\tau = 0.05$ to $\tau = 0.95$, we map how predictive relationships change across bearish, normal, and bullish conditions.

We implement the test via quantile regression (Koenker and Bassett, 1978) with five lags, comparing a restricted model (own lags only) against an unrestricted model (own lags plus the cause variable's lags). An F-type test determines whether the additional lags are jointly significant. The QGC coefficient reported in Tables 5 and 6 is $\sum_{i=1}^{5} |\hat{\beta}_i|$, measuring the aggregate predictive content of X's recent history for a given quantile of Y. This captures predictive magnitude rather than a marginal effect.

The test is applied to all 210 directed asset pairs (15 assets $\times$ 14 potential targets) at 19 quantile levels ($\tau \in \{0.05, 0.10, \ldots, 0.95\}$), separately for each of the four market regimes. Because this design produces a large number of simultaneous tests, we control for multiple comparisons using the Benjamini and Hochberg (1995) False Discovery Rate (FDR) procedure at the 5% level, applied within each quantile-regime combination. This ensures that within each quantile-regime cell, the reported significant pairs are not driven by chance discoveries. We acknowledge that this local FDR control does not directly govern cross-quantile comparisons such as the U-shape pattern. Three complementary lines of evidence support the U-shape as a genuine feature rather than an artefact. First, under the global null of no true causality, false rejections should be approximately uniformly distributed across quantiles ($\approx$10.5 per cell). The observed pattern, with 50–100+ rejections at $\tau = 0.05$ and $\tau = 0.95$ but exactly zero at the median after FDR correction, is inconsistent with uniform false positive allocation. Second, the merged Crisis-Recovery robustness check (Section 3) reproduces the U-shape in an independent regime partition with a larger effective sample. Third, the systematic bootstrap equality tests in Table 6 directly compare tail and median coefficients across all four regimes.

We note three additional caveats. First, at extreme quantiles ($\tau = 0.05, 0.95$), the effective number of observations defining the tail is small, particularly in the shorter regimes (e.g., approximately 11 observations at $\tau = 0.05$ in the 213-day COVID Crisis window). Results at extreme quantiles in shorter regimes should therefore be interpreted with appropriate caution. To address this concern, we conduct a robustness check in which the COVID Crisis and Recovery regimes are merged into a single "Crisis-Recovery" window (March 2020 – December 2021, 465 trading days). This combined regime yields approximately 23 observations at the 5th percentile, substantially improving the reliability of tail estimates while preserving the ability to compare crisis-era dynamics against the stable Pre-COVID and Inflation Shock periods. Second, the 210 directed pairs are not fully independent, as they share common assets. Our network-level analysis (leader-follower scores, heatmaps) aggregates across pairs and is therefore less susceptible to this concern than individual pair-level inference. Third, as with all bivariate Granger causality tests, significant results may partly reflect common exposure to unobserved macroeconomic shocks (e.g., FOMC announcements, oil price movements) rather than purely directional transmission from one asset to another. The regime-specific design mitigates this concern by conditioning on broadly homogeneous macro environments, but cannot fully eliminate it.

3. Empirical Results

Figure 1 presents the full result. An overview of Figure 1A shows that cross-asset causality follows a pronounced U-shaped cross-return distribution. This validates the study's quest to use QGC. The U-shape is not symmetric. The left tail consistently dominates the right tail during the Pre-COVID and Inflation Shock periods, reflecting the well-documented asymmetry in financial contagion, whereby crashes generate stronger cross-asset linkages than rallies (Bouri et al., 2021). The magnitudes of the coefficients are shown in Figure 1b, which indicates that causality is not only prevalent but also economically meaningful in the tails. The COVID Crisis period shows the strongest coefficients, with the left-tail averaging 1.74 and the right-tail 1.21, indicating substantial predictability between assets during extreme market conditions.

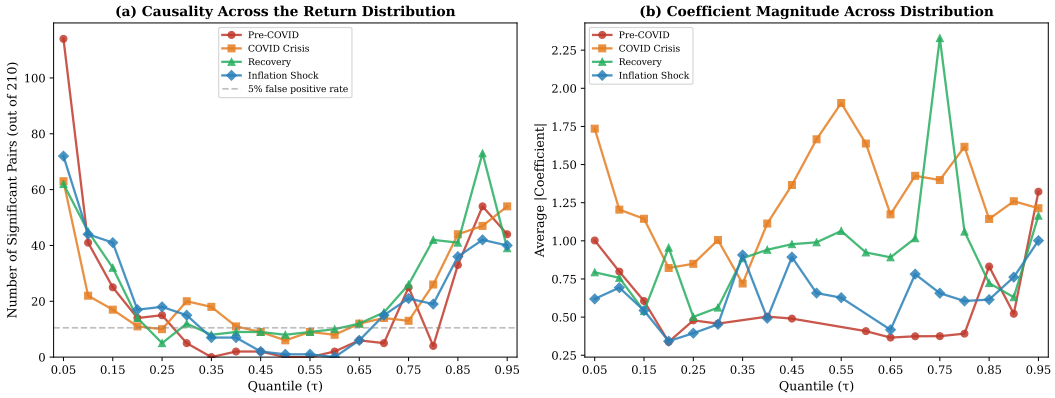

Figure 1: Quantile Process of Cross-Asset Causality. Panel (a) displays the number of significant causality pairs (out of 210) at each quantile $\tau \in \{0.05, 0.10, \ldots, 0.95\}$, based on unadjusted p-values at the 5% level. Panel (b) displays the average coefficient magnitude across significant pairs, confirming that tail causality is not only more prevalent but also economically larger.

Table 4 presents the summary of the significant causality pairs across the entire distribution. The table shows that the U-shape is remarkably consistent across regimes, with the pre-COVID era leading with 104 significant relationships at $\tau = 0.05$. The concentration of causality in the left tail during stable periods reflects the asymmetric nature of market stress. During the pre-pandemic expansion, markets experienced occasional sharp selloffs (December 2018, August 2019) that triggered coordinated movements across assets. These crash events, though infrequent, are precisely when cross-asset predictive linkages matter most. The left tail captures these episodes, while the median captures the many unremarkable trading days when assets moved independently.

The COVID Crisis regime exhibits a flatter U-shape compared to other regimes, with some significance persisting at intermediate quantiles ($\tau = 0.10$ retains 5 significant pairs after Full Distribution Return (FDR) correction). This reflects the extraordinary volatility of the pandemic period, during which even moderate return movements contained cross-asset information. The convergence of left- and right-tail causality reflects the unique market dynamics of the pandemic period, in which both crashes and rallies occurred with unusual frequency.

The recovery era saw synchronized gains across asset classes, including tech, cyclicals, commodities, and distressed sectors, as risk appetite returned. This created strong right-tail causality, as positive shocks to one asset were associated with extreme positive returns in others. The Federal Reserve's commitment to unlimited quantitative easing, including unprecedented corporate bond purchases, reduced downside risk and suppressed the risk of a crash. The liquidity injection boosted upside co-movement, with investors bidding up risk assets. The COVID-era right-tail dominance

stems from this policy response.

The Inflation Shock period of 2022–2023 produced a stark disparity between tail and median significance. After FDR correction, the left tail exhibited 58 significant causality pairs while the median exhibited zero. This extreme concentration of causality in the tails reflects the binary nature of markets during aggressive monetary policy shifts.

Table 4: Significant Causality Pairs Across the Full Return Distribution

Quantile (τ)	Pre-COVID	COVID Crisis	Recovery	Inflation Shock
0.05	**104**	58	36	58
0.10	23	5	19	21
0.15	1	1	3	18
0.20	0	0	0	2
0.25	0	0	0	0
0.30	0	0	0	0
0.35	0	0	0	0
0.40	0	0	0	0
0.45	0	0	0	0
0.50	0	0	0	0
0.55	0	0	0	0
0.60	0	0	0	0
0.65	0	0	0	0
0.70	0	0	0	0
0.75	0	0	0	2
0.80	0	6	1	0
0.85	3	14	13	3
0.90	18	30	39	18
0.95	16	**45**	10	15

Note: Each cell reports the number of directed asset pairs (out of 210) where the QGC test rejects the null of no causality after applying the Benjamini-Hochberg (1995) False Discovery Rate correction at the 5% level within each quantile-regime combination. The U-shaped pattern of tail concentration is preserved across all four regimes: significance is concentrated in the extreme quantiles while no pairs survive FDR correction at the median. Shaded rows highlight the left tail ($\tau = 0.05$, red), median ($\tau = 0.50$, green), and right tail ($\tau = 0.95$, blue).

Sectoral Lead-Lag Network

The results of the sectoral lead-lag dynamics is presented in Figure 2[1]. Figure 2 displays the directed causality networks across all 15 assets at the left tail ($\tau = 0.05$) for each regime. Each cell (i, j) shows the QGC coefficient from asset i to asset j when statistically. The network structure reveals important heterogeneity across regimes. During Pre-COVID regime, the network was dense and broadly distributed, with most sectors both leading and following others. During the Inflation Shock, the network becomes sparser but more concentrated around specific predictive channels, consistent with the macro/micro day dichotomy, thereby revealing the sector-level architecture of tail-risk linkages. This implies that cross-sectional diversification benefits weaken during stressful market episodes.

[1]For brevity, we present only the left-tail results ($\tau = 0.05$). Results for the full distribution are available upon request.

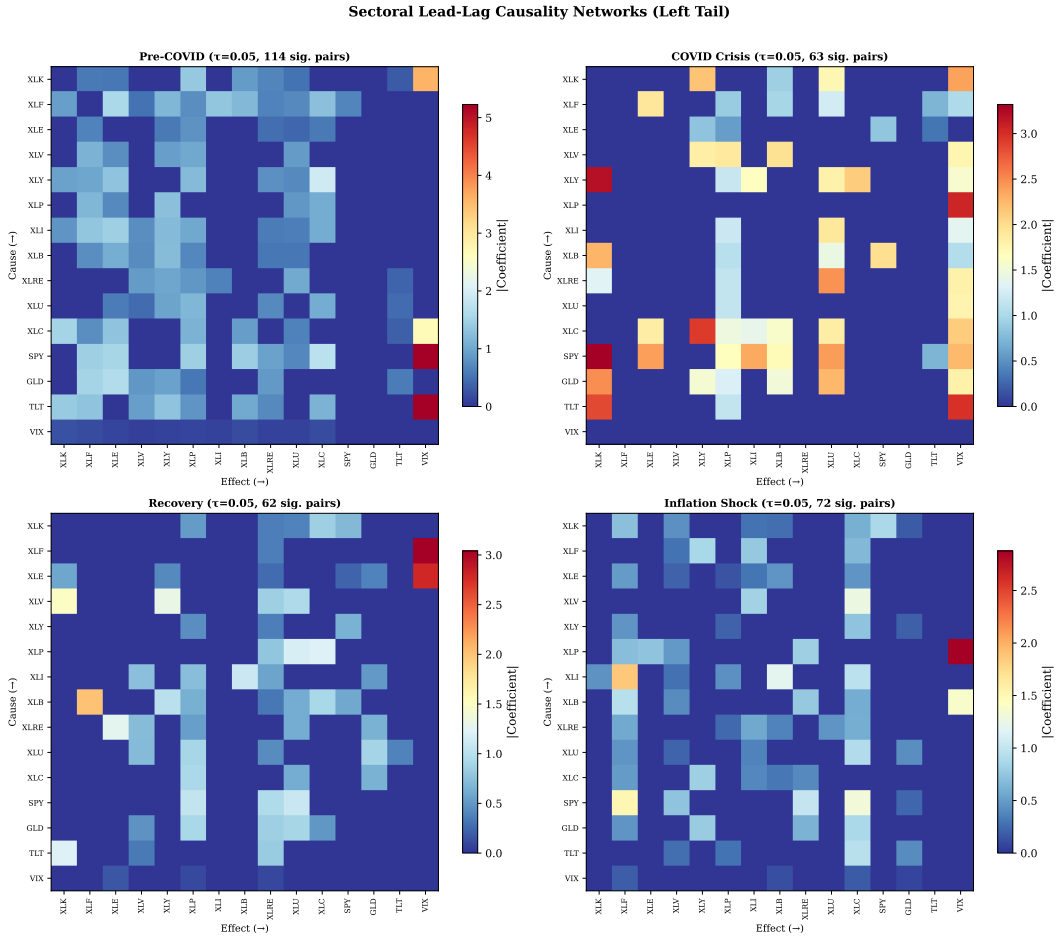

Figure 2: Sectoral Lead-Lag Causality Networks at the Left Tail ($\tau = 0.05$). Each panel shows a 15×15 matrix of directed causality coefficients for significant pairs. Rows indicate the "cause" (leading) asset and columns indicate the "effect" (lagging) asset. Darker cells indicate stronger causality. The Pre-COVID network is the densest, with 104 significant links, while regime-specific patterns reveal shifting predictive channels.

Leader-Follower Dynamics

Figure 3 decomposes the sectoral networks into leader scores (out-degree: how many sectors an asset's extreme returns predict) and follower scores (in-degree: how many sectors predict an asset's extreme returns). The leader rankings reveal economically interpretable patterns. During Pre-COVID, Financials (XLF) and VIX share the top position with out-degree 11 (predicting 11 of 14 other assets in their left tails), followed by Industrials (XLI) and Treasuries (TLT). The dominance of Financials as a systemic leader is consistent with the sector's role as a credit conditions barometer: stress in bank stocks signals deteriorating lending conditions that affect all sectors. VIX's leadership reflects the volatility channel whereby spikes in implied volatility predict subsequent tail realisations across asset classes. During COVID, SPY and Communication Services (XLC) led, while VIX dropped to zero out-degree, consistent with the policy-suppressed volatility regime. During the Inflation Shock, Technology (XLK) rises to the top, reflecting the sector's rate sensitivity as the dominant driver of macro-day selloffs.

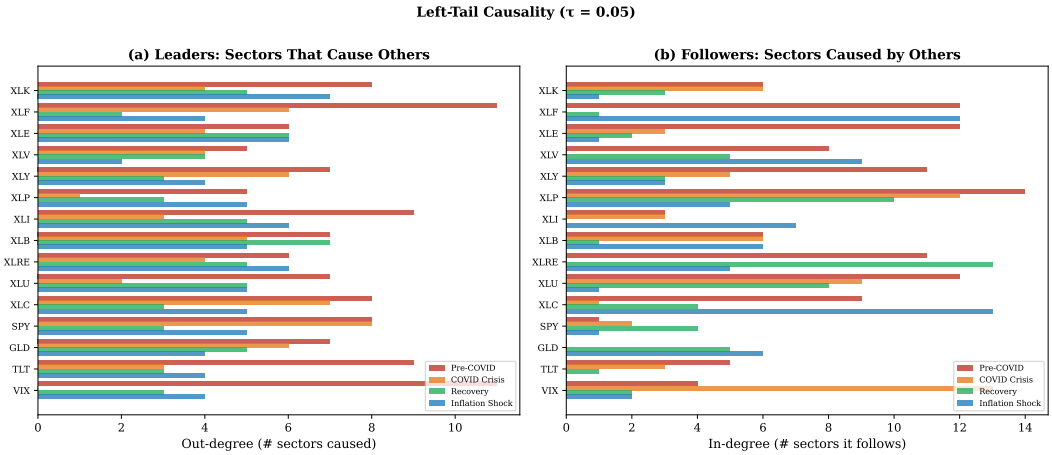

Figure 3: Leader and Follower Scores at the Left Tail ($\tau = 0.05$). Panel (a) shows the out-degree (number of other assets whose tail returns are predicted by each asset). Panel (b) shows the in-degree (number of assets that predict each asset's tail returns). Financials (XLF), VIX, and Industrials (XLI) emerge as the strongest leaders during stable markets.

Table 5 presents the strongest pairwise causality relationships from the QGC regressions. A striking feature is that VIX appears as the "effect" variable in all nine of the top pairs, suggesting it is the most frequently predicted variable in tail-risk networks. The largest coefficient is XLY→VIX during the Recovery regime at $\tau = 0.95$ (coefficient 12.84), indicating that lagged Consumer Discretionary returns carry substantial predictive power for the upper tail of the VIX distribution. More broadly, the dominance of the VIX channel is consistent with the leverage effect operating in both tails, as extreme equity movements, whether crashes or rallies, carry predictive information for implied volatility (Bouri et al., 2021; Shahzad et al., 2021).

Table 5: Strongest Pairwise Causality Relationships

From	To	Coefficient	95% CI	p-value	Regime	Quantile
XLY	VIX	**12.84**	[4.29, 18.70]	0.0143	Recovery	Right Tail (0.95)
SPY	VIX	**10.45**	[3.84, 16.27]	0.0217	Recovery	0.75
XLP	VIX	10.05	[4.05, 16.82]	0.0004	Pre-COVID	Right Tail (0.95)
XLC	VIX	8.20	[3.63, 13.38]	0.0260	Pre-COVID	Right Tail (0.95)
GLD	VIX	8.16	[4.19, 15.04]	0.0131	Pre-COVID	Right Tail (0.95)
XLP	VIX	8.15	[3.05, 10.72]	0.0434	Recovery	0.75
XLU	VIX	7.95	[4.35, 14.98]	0.0002	Pre-COVID	Right Tail (0.95)
SPY	VIX	6.98	[3.56, 14.38]	0.0489	Recovery	0.20
XLC	VIX	6.90	[2.06, 8.15]	0.0297	Recovery	0.75

Note: Top 9 pairwise causality coefficients from 15,960 QGC tests, ranked by coefficient magnitude. The p-values and 95% confidence intervals (block bootstrap, 500 replications, block size = 10) refer to pairwise unadjusted significance and are not FDR-corrected. Some entries (e.g., Recovery at $\tau = 0.20$ and $\tau = 0.75$) do not survive the BH-FDR correction applied in Table 4, which controls for multiplicity within each quantile-regime combination. These pairs are included to illustrate the strongest individual predictive relationships and the dominance of VIX as the effect variable.

To formally assess whether tail coefficients differ from median coefficients, we conduct a systematic bootstrap equality test across all four regimes. Rather than selecting pairs based on

exploratory point estimates, we draw a random sample of 10 pairs from the set of FDR-significant pairs at $\tau = 0.05$ in each regime, then test H_0: $\beta(0.05) = \beta(0.50)$ using 200 block bootstrap replications. Table 6 reports the 95% confidence interval for the *difference* $\beta(0.05) - \beta(0.50)$, which directly addresses whether the tail coefficient exceeds the median.

The results reveal a clear regime gradient. In the Pre-COVID period (426 days), all 10 randomly sampled pairs reject equality at the 5% level, with difference CIs comfortably excluding zero. In the COVID Crisis (213 days), 5 of 10 reject. In the Recovery (252 days) and Inflation Shock (501 days) regimes, rejection rates drop to 10% and 0%, respectively. This pattern reflects the well-known difficulty of estimating extreme-quantile coefficients precisely in shorter or more volatile samples, where the effective number of tail observations is small. The Pre-COVID result, based on the longest stable regime, provides the strongest evidence that tail predictability genuinely exceeds median predictability. The weaker results in other regimes should be interpreted as reflecting low statistical power rather than the absence of tail effects, given that the U-shape in rejection counts (Table 4) is robust across all regimes and the merged-regime check.

Table 6: Systematic Test of Tail vs. Median Coefficient Equality

Regime	**From → To**	$\hat{\beta}(0.05)$	$\hat{\beta}(0.50)$	**Diff**	**95% CI (Diff)**	p
Pre-COVID	XLY → XLE	1.31	0.27	1.04	[0.13, 1.27]	0.020**
Pre-COVID	XLC → XLF	0.77	0.24	0.53	[0.12, 1.11]	0.010**
Pre-COVID	XLF → XLY	1.18	0.20	0.97	[0.17, 1.43]	0.020**
COVID	XLE → XLY	0.84	0.15	0.69	[0.04, 1.27]	0.040**
COVID	GLD → XLB	1.51	0.34	1.18	[−0.18, 2.45]	0.070
COVID	SPY → VIX	2.23	3.20	−0.97	[−2.39, 2.28]	0.970
Recovery	XLE → GLD	0.38	0.11	0.27	[0.03, 0.59]	0.030**
Recovery	XLV → XLY	1.31	0.55	0.76	[−0.26, 1.64]	0.130
Inflation	XLP → VIX	3.60	1.50	2.09	[−1.24, 3.09]	0.350
Inflation	XLU → GLD	0.42	0.08	0.33	[−0.06, 0.42]	0.120

Note: For each regime, 10 pairs were drawn at random from the set of FDR-significant pairs at $\tau = 0.05$ (representative subset shown). The "Diff" column reports $\hat{\beta}(0.05) - \hat{\beta}(0.50)$, with 95% confidence intervals and p-values from 200 block bootstrap replications (block size = 10). ** denotes rejection at 5%. Rejection rates across the full random samples are 100% (Pre-COVID), 50% (COVID Crisis), 10% (Recovery), and 0% (Inflation Shock). The Pre-COVID regime, with the largest effective sample at the 5th percentile (≈21 observations), provides the strongest evidence that tail predictability exceeds median predictability.

4. Conclusion and Policy Implications

This study examines directional sectoral spillovers in the US equity markets using the QGC test. The results demonstrate that sectoral lead-lag transmission is predominantly tail-driven. Spillover remains weak under normal market conditions. Further results confirm regime heterogeneity, with crash-driven spillovers dominating during monetary-tightening periods and symmetric tail spillovers emerging during crisis-recovery episodes. The results of the leader-follower dynamics is sensitive to the regime under consideration.

Results presented thus far have important policy implications. The concentration of spillovers in extreme returns suggests that conventional correlation-based risk metrics may underestimate systemic vulnerability during crisis periods. The regulatory framework, primarily based on full-sample dependence structures, may fail to capture cross-sector tail dependence during extreme events. Furthermore, the regime-dependent nature of spillovers suggests that predictive linkages

evolve alongside macroeconomic and monetary policy conditions. During tightening cycles, markets appear more sensitive to common macroeconomic shocks, resulting in synchronised sectoral declines. Risk managers should employ tail-specific correlation matrices for stress testing rather than relying on single-regime estimates. The regime dependence of tail causality also complicates hedging. Static strategies calibrated to crash protection may fail during coordinated rallies (as in COVID), while strategies designed for normal conditions will be overwhelmed during tail events.

These findings have implications for risk management and financial stability. Value-at-Risk models should incorporate tail-specific correlations. Stress tests should use sector-level causality networks rather than aggregate correlation matrices. Macroprudential monitoring should track the evolving leader scores of key sectors as early warning indicators. We acknowledge that bivariate Granger causality cannot fully distinguish directional transmission from common exposure to unobserved macroeconomic shocks, and the inclusion of broad aggregates (SPY, VIX) alongside sector ETFs may amplify this concern. Future research should extend this analysis to conditional or factor-orthogonalised specifications, higher frequency data, and international markets, where cross-border tail dependence networks may exhibit similar U-shaped patterns.

References

[1] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.

[2] Balcilar, M., Gupta, R., & Pierdzioch, C. (2016). Does uncertainty move the gold price? New evidence from a nonparametric causality-in-quantiles test. *Resources Policy*, 49, 74–80.

[3] Bouri, E., Cepni, O., Gabauer, D., & Gupta, R. (2021). Return connectedness across asset classes around the COVID-19 outbreak. *International Review of Financial Analysis*, 73, 101646.

[4] Chordia, T., & Swaminathan, B. (2000). Trading volume and cross-autocorrelations in stock returns. *Journal of Finance*, 55(2), 913–935.

[5] Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive. *International Journal of Forecasting*, 28(1), 57–66.

[6] Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54(6), 2143–2184.

[7] Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.

[8] Jeong, K., Härdle, W. K., & Song, S. (2012). A consistent nonparametric test for causality in quantile. *Econometric Theory*, 28(4), 861–887.

[9] Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3(2), 175–205.

[10] Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: What is the role of the United States? *Journal of Finance*, 68(4), 1633–1662.

[11] Raheem, I. D., Ur Rehman, M., & le Roux, S. (2023). Oil shocks and the Islamic financial market: Evidence from a causality-in-quantile approach. *Unpublished manuscript*.

[12] Shahzad, S. J. H., Bouri, E., Kristoufek, L., & Saeed, T. (2021). Impact of the COVID-19 outbreak on the US equity sectors: Evidence from quantile return spillovers. *Financial Innovation*, 7(1), 1–23.