Dipartimento
di Economia
Working Papers

n. 294
2026

Eduardo Montuori,
Francesco Benedetto
and Loretta Mastroeni

SYSTEMIC RISK SYNCHRONIZATION ACROSS EUROPEAN BANKING AND INSURANCE SECTORS: A TIME-WARPING AND ENTROPY APPROACH

ISSN 2279-6916 Working Papers

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Dipartimento
di Economia
Working Papers

WORKING PAPERS
Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
workpapers.economia@uniroma3.it
https://economia.uniroma3.it/

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Systemic Risk Synchronization Across European Banking and Insurance Sectors: A Time-Warping and Entropy Approach*

Eduardo Montuori[a,b], Francesco Benedetto[c] and Loretta Mastroeni[c]

[a]*Sapienza University of Rome, Department of Social and Economic Sciences, Piazzale Aldo Moro 5, Rome, 00185, Italy*
[b]*Bank of Italy, Via Nazionale 91, Rome, 00184, Italy*
[c]*Roma Tre University, Department of Economics, Via Silvio D'Amico 77, Rome, 00145, Italy*

ARTICLE INFO

Keywords:
Systemic Risk
Financial Interconnectedness
Tail Risk
Banking and Insurance Sector
Dynamic Time Warping
Transfer Entropy
JEL classification: G010, G170, G210, G220.

Abstract

This paper studies how systemic risk builds up and propagates among European banks and insurers, two central sectors to economic growth and to the stability of the whole financial system. Understanding their joint reaction to shocks is relevant because stronger short-run synchronization can amplify system-wide distress. Using daily prices and risk measures—volatility, Value-at-Risk, and Expected Shortfall—for the STOXX Europe 600 Banks and Insurance indices, we assess whether cross-sector linkages manifest through price co-movements or the alignment of downside risk. A dynamic time warping approach captures non-synchronous responses to common shocks. Results show that price co-movements are episodic and concentrated around crises, whereas volatility and downside risk remain persistently aligned, with Expected Shortfall displaying the strongest alignment. Granger causality and Effective Transfer Entropy confirm the existence of directional spillovers. A key implication is that macroprudential policy frameworks should adopt a joint and coordinated approach to the banking and insurance sectors.

1. Introduction

Financial stability is a precondition for the real economy to provide jobs, credit and growth. The 2008 global financial crisis highlighted the need of anticipating adverse macro-prudential developments and preventing the build-up of excessive risks within the financial system, going beyond a purely micro-prudential approach (see, e.g., Basel Committee on Banking Supervision, 2011). Preserving the broader financial stability requires understanding how systemic risk emerges and propagates across the entire financial system and assessing the contribution of each financial sector (see, e.g., Bernal et al., 2014). A key challenge is that stress does not necessarily propagate through contemporaneous co-movements alone: linkages may be asynchronous, state-dependent, and directionally biased, reflecting heterogeneous balance-sheet structures, regulatory constraints and market-specific dynamics across sectors (see, e.g., Diebold, F. X. and Yilmaz, K. 2015). This is particularly relevant for banks and insurance companies, considering their essential role for the real economy. Disruptions in either sector can create spillover effects and potentially impair investment and growth for the broader economy. The economic centrality and systemic relevance of banks and insurers is widely recognized by the EU legislators, which defined a comprehensive regulatory environment for those entities in order to ensure their sound prudential management and their orderly failure. Moreover, the insurance crisis-management toolkit has been reinforced by the newly-introduced Insurance Recovery and Resolution Directive (IRRD).

Historically, financial regulation has focused primarily on the banking sector, reflecting its pivotal role in the economy. In contrast, the debate on the role of insurers in financial stability has become central only in recent years, both in academic research and at the institutional level. At Academic level, a growing strand of research has investigated the systemic relevance of the insurance sector and its links with the broader financial system (see, e.g., Weiß and Mühlnickel, 2014 and Mühlnickel and Weiß, 2015), showing that insurers can become systemically relevant in light of their size. Regarding the institutional level, the Financial Stability Board (FSB) and the International Association of

*The views expressed in the articles are those of the Authors and do not involve the responsibility of the Bank of Italy.

We thank Prof. Maria Carannante for serving as referee for this working paper in the Roma Tre University Department of Economics Working Paper Series, and for her helpful comments.

edu rdo.montuori@unirom 1.it (E. Montuori); fr ncesco.benedetto@unirom 3.it (F. Benedetto); lorett .m stroeni@unirom 3.it (L. Mastroeni)
ORCID(s):

Insurance Supervisors (IAIS) jointly published the first list of global systemically important insurance institutions in 2013, now incorporated in the Holistic Framework for the assessment and mitigation of systemic risk in the insurance sector (see IAIS, 2019; Financial Stability Board, 2019). At the European level, the European Systemic Risk Board (ESRB) has increasingly emphasized the need for macroprudential oversight of insurers (see, e.g., ESRB, 2015), including, inter alia, its 2018 report on macroprudential provisions, measures, and instruments for insurance, as well as more recent annual reports and monitoring publications (see, e.g., ESRB 2024 and 2025).

Addressing some of the afore-mentioned issues, this paper focuses on the *evolution* of risk dynamics in the European banking and insurance sectors and on the mechanisms through which stress is transmitted across them, proposing a synchronization framework with complementary directional-dependence (causal-diagnostic) tools.

Traditionally, much of the existing literature analyzes systemic risk through well-established statistical and econometric frameworks, including volatility models and (conditional) tail-based measures (Bollerslev, 1986; Engle, 2002; Idier, J., Lamé, G., & Mésonnier, J.-S., 2014; Banulescu, G.-D., & Dumitrescu, E.-I., 2015; Adrian and Brunnermeier, 2016; Brownlees and Engle, 2017). These approaches have proven to be highly informative in quantifying the magnitude of risk and, to some extent, its transmission mechanism between financial institutions and sectors (Acharya et al., 2017). By construction, however, they rely on predefined functional forms, which may be unable to capture non-linear temporal distortions and evolving lead–lag relationships in the propagation of financial stress. Our approach complements these cornerstone methodologies by providing a flexible, model-free perspective on the timing and synchronization of risk dynamics. In particular, we apply the Dynamic Time Warping (DTW) methodology to daily prices, volatility, Value at Risk (VaR), and Expected Shortfall (ES) for the STOXX Europe 600 Banks and Insurance indices, over a long sample period (from 2000 up to december 2025). By comparing results across these measures, we disentangle different layers of financial risk integration, ranging from valuation dynamics to extreme tail risk, which could be particularly informative during periods of market wide turmoil. Our results are then enriched by an information-flow analysis, to capture directional dependence exploiting (i) bivariate Granger causality tests - a workhorse framework for linear predictive content under minimal identifying assumptions - and (ii) Effective Transfer Entropy (ETE), a nonparametric information-theoretic measure designed to detect nonlinear, tail-driven directional information flow. Both tools address predictive directionality rather than structural causality: Granger tests whether the lags of a source series add incremental forecasting power for the target conditional on the target's own lags, while ETE quantifies directed information transfer after correcting for spurious dependence via shuffled-source surrogates. Overall, the information-flow analysis results provide stronger support for the patterns identified through the DTW framework. These findings highlights that cross-sector linkages are not primarily reflected in the returns, but materialize through the risk channel, and most clearly through tail risk. In stress episodes, volatility and extreme losses synchronize and spill over across sectors, even if the peaks arrive with delays. The rest of this paper is organized as follows. Section 2 describes the state-of-art solutions published so far, highlighting both the significant advances and critical limitations in terms of identification and measurement of the systemic risk in the insurance and banking sector, especially from a joint perspective. Section 3 introduces our methodology for detecting the co-evolution of risk measures across sectors, providing a direct, model-free way to detect shared stress patterns under temporal misalignment complemented by directional-dependence (causal-diagnostic) tools. Section 4 illustrates our set-up in terms of data. Section 5 provides an in-depth analysis of the results of our approach, while robustness analysis as well as comparison with state-of-the-art methods are reported in Section 6. Our conclusions are briefly stated in Section 7, while a discussion about policy implications is illustrated in Section 8.

2. Literature review

The measurement of interconnectedness and systemic risk has evolved from early approaches based on linear co-movements of asset prices toward frameworks that explicitly quantify the transmission of stress across institutions and sectors. A central insight is that the magnitude of the relation is often state-dependent: linkages that look moderate in tranquil periods tend to strengthen during turmoil. This aspect has motivated empirical designs that go beyond simple correlations and focus on time-varying connectedness and tail-risk dynamics. A summary of relevant studies on systemic risk among different financial institutions, with a focus on DTW approaches is provided in Table 1. A first strand of research develops econometric measures of connectedness to jointly monitor multiple segments of the financial system. Billio, Getmansky, Lo and Pelizzon (2012) propose econometric measures of connectedness and systemic risk and apply them to returns from several financial sectors, including banks and insurance companies. Their approach allows the identification of institutions and sectors that are central to the system and contribute to a higher extent to system-wide fragility. A related line of inquiry evaluates how different types of intermediaries contribute to systemic risk. Bernal, Gnabo and Guilmin (2014) assess, applying the traditional Delta CoVaR, the contribution of banks, insurance companies and other financial services to systemic risk, contributing to the debate on whether systemic relevance is confined to traditional banking. Their evidence supports the view that insurers may contribute materially to systemic fragility.

In a similar vein, Zhou, Liu and Wang (2020) provide a comparative analysis of CoVaR, MES and SRISK across banks, insurers and securities firms over 2007-2018 in China and show that these metrics convey distinct information and may rank institutions differently. Coleman, LaPlante and Rubtsov (2018) analyze the SRISK measure and its application to banking and insurance industries, emphasizing the relevance of expected capital shortfall as a monitoring concept. Extended versions of these measures have been recently applied in analyzing in particular the Chinese systemic risk: Wang et al. (2022) use a Double CoVaR; Hao (2025) employs a trivariate CoVaR based on vine copulas and Liu (2025) develops a composite systemic risk model using a financial knowledge graph for risk measurement. Similarly, focusing directly on the bank-insurer nexus, Chen, Cummins, Viswanathan and Weiss (2014) analyze interconnectedness between banks and insurers using daily market value data on credit default swaps (CDS) and equities. They document that the strength of the relationship is time-varying and found that the impact of banks on insurers is stronger and of longer duration than the impact of insurers on banks. CDS-based measures are also used by Rodríguez-Moreno, M., & Peña, J. I. (2013).

Beyond market prices, the literature emphasizes structural channels that can spread distress across those intermediaries. In this regard, Barucca, Mahmood and Silvestri (2021) study common asset holdings as a source of systemic vulnerability across multiple types of financial institutions. Chang (2018) investigates the linkage between insurance and banking activities through an analysis centered on insurance-sector assets and their interaction with banking and financial system indicators. Song, Xing and Li (2023) analyze the role of insurers' institutional shareholding and banking systemic risk within a high-dimensional network setting, pointing to ownership-related channels as additional mechanisms through which cross-sector risk can propagate.

Collectively, these contributions denote a comprehensive analysis of the two sectors and their interactions through various established risk measures. Methodologically, the dominant paradigm when contagion is estimated via market-based information relies on advanced econometric tools (VAR-based connectedness, Granger-causality networks and spillover indices) all characterized by the assumption of a shared time grid and fixed temporal alignment. Yet, stress transmission across sectors is plausibly asynchronous because banks and insurers differ in balance-sheet structure, losses materialization, liability duration, hedging and reinsurance practices and regulatory constraints. Fixed-alignment methods may therefore understate the similarity when comparable stress patterns unfold at different speeds. DTW offers a conceptually different perspective by allowing a flexible alignment of two time series, matching patterns that occur at different speeds or with sector-specific delays. DTW has gained traction in finance for similarity analysis and clustering when the shape of the trajectory is the object of interest. D'Urso, De Giovanni and Massari (2019) develop trimmed fuzzy clustering of financial time series based on DTW; Gassouma, Benhamed and El Montasser apply DTW to study similarities between Islamic and conventional banks. Lim and Ng (2023) use shape based unsupervised learning for analyzing international markets and downside-event regimes, consistent with the idea that trajectory shape contains information about stress synchronization. Mastroeni et al. (2021) apply a DTW framework to energy markets and show that Brent and crude oil prices display strong underlying similarity once temporal distortions are accounted for, even when standard correlation measures suggest weaker integration. Their analysis emphasizes that DTW does not merely measure comovement but rather reconciles temporally displaced responses to common shocks, offering a dynamic interpretation of market integration. Subsequent applications show that DTW can uncover contagion and synchronization patterns that remain hidden under standard approaches, especially during periods of stress (Denkowska and Wanat, 2021; Guo et al., 2022). Grzejszczak et al. (2022) apply the DTW algorithm to stock market data from the Warsaw Stock Exchange, combining it with statistical techniques such as Fourier transform and random permutation tests to detect and verify similarity patterns between financial time series, demonstrating the usefulness of DTW for analyzing similarity in financial time series.

Building on this intuition, this study aims to contribute to the literature by extending the DTW framework to the analysis of systemic risk across the banking and insurance sectors, jointly considering prices and multiple risk measures, including volatility, Value-at-Risk and Expected Shortfall. By explicitly accounting for non-synchronous risk dynamics, the analysis complements existing econometric approaches and provides a dynamic perspective on how common sources of financial stress propagate across the two core pillars of the financial system. Despite these methodological advances, DTW has not been systematically employed to study bank-insurance linkages. Moreover,

Table 1
Summary of recent studies on systemic risk among financial institutions

Study	Scope / data	Methodologies
Billio et al. (2012)	Monthly returns data for hedge funds, broker/dealers, banks, and insurers	Principal Component nalysis and Granger Causality Networks
Bernal et al. (2014)	S&P 500 ex-Financial index (United States); STOXX Europe 600 (Eurozone)	Delta CoVaR, extended with the Kolmogorov-Smirnov test
Chen et al. (2014)	Daily equity and Credit Default Swaps (CDS) market value data (global banks and insurers)	Systemic risk measure applied on insurance companies according to Huang et al. (2009) (based on Probability of Default derived from CDS and correlation); Granger causality tests
Denkowska & Wanat (2021)	Market data on EU insurance companies	Dynamic Time Warping, tail dependence via multivariate G RCH measures
Grzejszczak et al. (2022)	Warsaw Stock Exchange companies	Dynamic Time Warping, Fourier transform and Random permutation
Wang et al. (2022)	sia, listed insurance companies	Double CoVaR (as Fang et al., 2021)
Hao (2025)	The China Shenwan (SWS) Industry Index on Banking, Securities and Insurance	Trivariate CoVaR based on vine copulas
Liu (2025)	Financial companies (banks, insurance, fintech and securities) equity prices	Composite systemic risk model using the financial knowledge graph for risk measurement

from a general point of view, it has been not applied on risk measures such as volatility, VaR and ES as the primary objects of comparison. According to the authors' knowledgde, this is the first paper that brings DTW to the bank-insurance systemic-risk domain, by shifting the empirical focus from synchronous price co-movement to the co-evolution of risk measures. By computing DTW-based similarities across multiple objects (prices, volatility, VaR and ES), we complement the econometric connectedness literature by providing a direct, model-free way to detect shared stress patterns under temporal misalignment.

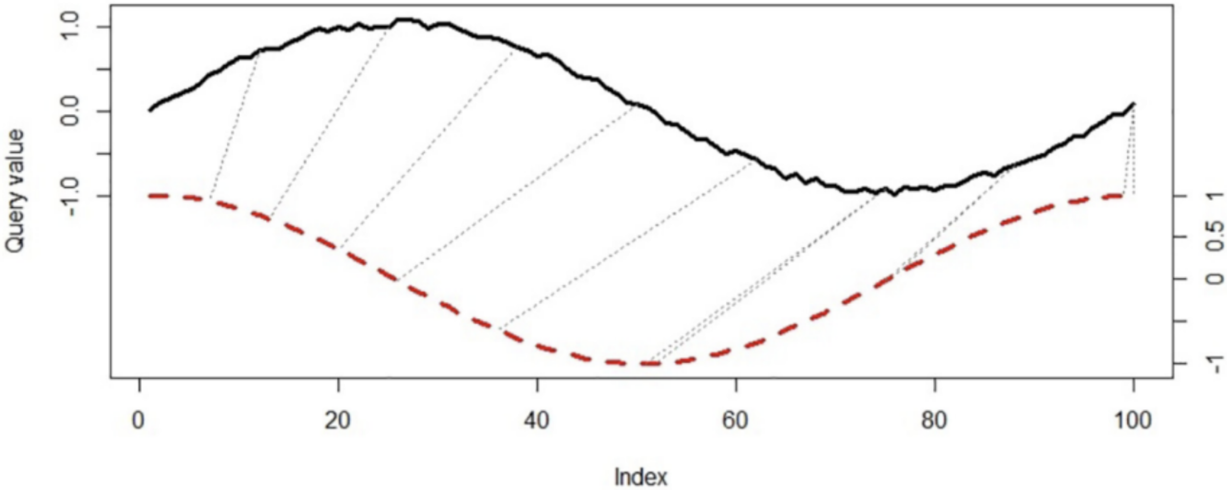
3. Methodology

3.1. Dynamic Time Warping

DTW is the name of a class of algorithms aimed at reveal how much a time series is connected with another one. The rationale behind DTW is to stretch or compress two time series locally, in order to make their shape as similar as possible (see Fig. 1). The distance between them is computed by summing the distances of individual aligned elements. After being introduced around 1970 in the context of speech recognition, it has been employed for classification and clustering aims in a plenty domains, Giorgino (2009). As recalled by Keogh and Pazzani (2000), the reason why Euclidean distance may fail to produce an intuitively correct measure of similarity between the two indexes under investigation is because it is very sensitive to small distortions in the time axis. In order to better illustrate how DTW works, let us consider Fig. 1: the two simulated time series have approximately the same overall shape, but are not aligned on the time axis. The nonlinear alignment (shown in red in the same picture) would allow a more sophisticated distance measure to be calculated.

Let us define two time series as $x = (x_1 \ \dots \ x_N)$ and $y = (y_1 \ \dots \ y_M)$, respectively, where N and M are their respective lengths (by the way, in our problem $N = M$). They are the so-called query and the reference series in the DTW jargon. The total distance between x and y is computed via the so-called warping path, which ensures that each data point of the query is compared to the closest data point of the reference. Alignment is a global property of the time series, and, therefore it does not depend only on the value at a particular time t but also on those in the subsequent and previous ones because it is the global optimum on all possible paths.

Since there is a one-to-one correspondence in the sampling times of both the time series, the DTW algorithm searches for the time of the series x optimally matching to the time of the series y. That is, if it finds that day 10 of x matches day 10 of y, day 11 of x matches day 11 of y, and so on. This results in a diagonal warping path. This time correspondence makes the warping path very close to the true diagonal. Based on the process that produces the



igure 1: Dynamic time warping distance. simple example of aligning two time series. query (solid) and a reference (dashed).

warping path, Zoumpoulaki et al. (2015) suggest that its distance from the diagonal provides a reliable measure of the direction of latency differences between two time series. A positive distance, which results from the warping path being below the diagonal, indicates that the reference time series (in our application, Insurance index), used for alignment, precedes the query time series (in our application, Banking index), while a negative one indicates that the reference time series follows the query. This simple fact measures whether a pattern appears earlier in the reference than in the query and can be interpreted as a preliminary indicator of a potential causal relation (e.g., if the reference precedes the query, then it can also be a driver of the query). In our baseline implementation we use the DTW methodology as described above, as our primary interest is in investigating global alignment properties and time-varying lead-lag patterns without imposing an ex-ante bound on allowable warping. As a robustness analysis, we then replicate our DTW-based results under a Sakoe-Chiba band constraint using a restrictive window of 60 trading days, confirming that the main conclusions are not driven by excessively large temporal warpings.

DTW can be successfully used to investigate the potential non-synchronous responses to common shocks by banking and insurance sectors. It enables a flexible temporal alignment of time series and captures time-varying lead lag relationships, revealing whether temporally misaligned risk episodes could reflect the same underlying systemic stress. We apply our framework to two time series (STOXX Europe 600 Banks and Insurace Indexes) that are fully comparable due to the data recorded on the same day. In addition, most of the literature highlighted that the two sectors of banking and insurances have historically contributed to the build up of systemic risk and have been largely exposed to common shocks, even if in some case have responded with a different timespan. Thanks to the previous considerations, DTW: i) can be considered a method to detect the alignment points between the two time series when dynamic similarity and synchronization patterns appear; ii) is suitable for investigating whether a pattern appears first in one time series rather than the other, thus providing a measure of the latency difference between two time series; iii) is a non-parametric technique that works directly on price and risk measures derived from data; iv) is able to detect synchronization patterns also in the presence of non-linear variations between the two series.

3.2. Information flow analysis

The DTW does not provide relevant insights on the information flow or potential causal effects among the time series under investigation. Therefore, in order to complement our analysis, we employ two notions of directed dependence: (i) linear, with the Granger causality and (ii) non–linear, measured by the Transfer Entropy (TE). Brought from information theory, TE describes the degree to which a random process reduces uncertainty about the future value of another random process beyond the degree to which the former reduces uncertainty about its own future. On the other hand, Granger causality is described in terms of prediction based on a vector autoregression approach (see e.g. Barnett et al. (2009)). Furthermore, Granger causality needs a parametric model for the data, while transfer entropy is non-parametric and does not require assumptions on the joint probability distribution of the data (see e.g. Barnett et al. (2009)). Both Granger causality and transfer entropy are most commonly implemented and interpreted under (weak) stationarity, or within frameworks that render the working series approximately stationary. In particular, standard Granger tests rely on stable autoregressive representations for effective inference, while transfer entropy is estimated from empirical joint and conditional distributions sufficiently stable over the estimation window. In this regard, prior to applying Granger causality and Effective Transfer Entropy, we apply standard stationarity-inducing transformations commonly adopted in empirical finance. Specifically, we work with daily log-returns for prices and with first differences for the rolling risk measures (volatility, VaR and ES). We formally confirm stationarity of these transformed four measures for the two series using Augmented Dickey–Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) unit root tests. Results are provided in Table 3, in the subsequent section 4 (Data and preliminary statistics)

Bivariate Granger causality To test whether the Insurance sector index contains incremental predictive information for the Banking sector (and vise versa), we conduct bivariate Granger causality tests on each transformed measure. Let $x = (x_1 \ \dots \ x_N)$ denote the target series and $y = (y_1 \ \dots \ y_M)$ the source series, where N and M are their respective lengths (and $N = M = T$ in our case). We estimate the single-equation Autoregressive Distributed Lag (ARDL($p \ q$)) model:

$$x_t = c + \sum_{i=1}^{p} \alpha_i \, x_{t-i} + \sum_{j=1}^{q} \beta_j \, y_{t-j} + \varepsilon_t \qquad t = \max\{p \ q\} + 1 \ \dots \ T \tag{1}$$

where c is a constant, $p \geq 1$ and $q \geq 1$ are, respectively, the numbers of lags included for the target and source series, and ε_t is an error term. We test the null hypothesis of no Granger-causality from y to x,

$$H_0 : \ \beta_1 = \cdots = \beta_q = 0$$

using a Wald test with Heteroskedasticity and Autocorrelation Consistent (HAC) robust covariance estimation (see, e.g., Granger, 1969; Newey and West, 1987). The analysis is performed in both directions (Banks $\leftrightarrow$ Insurance) and across the four objects of interest (prices, volatility, VaR and ES).

Transfer Entropy and Effective Transfer Entropy (ETE). Transfer entropy was first proposed by Schreiber (2000) as a method to detect asymmetric information transfer in non linear systems. To introduce the concept, let us consider a source X that can transmit N symbols. Then the Information or Shannon entropy is given by the formula

$$E(X) = -\sum_{i=0}^{N-1} p_i \log_2 p_i \tag{2}$$

where p_i is the probability of the observation of the symbol i in the stream of characters of the message. Eq. (2) provides a way to estimate the average minimum number of bits needed to encode a string of symbols (emitted by a source X), based on the frequencies of the respective symbols. Symbols which are more probable than others produce observations (i.e. receptions) which are less informative for the receiver. On the other hand, rarer symbols provide more information when observed. The entropy is zero when an outcome is certain: in practical terms, the information is the removal of uncertainty. If Shannon entropy quantifies the uncertainty associated with a single process, transfer entropy quantifies the reduction in that uncertainty when past information from another process is included. Therefore, TE can be interpreted as a directional and dynamic extension of the entropy concept introduced above.
TE is particularly appealing in finance because, unlike linear Granger causality, it captures nonlinear and non-Gaussian dependence; moreover, it is worth noting that TE and Granger causality are equivalent only under Gaussian assumptions. In line with Benedetto et al. (2020), in order to measure the information flow between two time series, we assume that the underlying processes evolve over time according to a Markov process. Each series is *symbolically encoded* into three states using the empirical 5% and 95% quantiles: observations in the lower tail (below the 5th percentile), in the central region, and in the upper tail (above the 95th percentile). Let the encoded source and target processes be denoted by $\tilde{x}_t$ and $\tilde{y}_t$. TE from x to y is then computed as the Kullback–Leibler divergence between transition probabilities of the target with and without conditioning on the source. Let $p(\ )$ denote the joint

and conditional probability density functions of the discrete-valued Markov processes $(\tilde{x}_t \ \ \tilde{y}_t)$:

$$TE_{x \quad y} = \sum p(\tilde{y}_{t+1} \ \ \tilde{y}_t \ \ \tilde{x}_t) \log \frac{p(\tilde{y}_{t+1} \mid \tilde{y}_t \ \ \tilde{x}_t)}{p(\tilde{y}_{t+1} \mid \tilde{y}_t)}. \tag{3}$$

In this regard, we adopt Markov orders $k = l = 1$, consistent with the computational feasibility in long daily samples and consistent with the empirical applications of the methodology. A well-known issue is that finite-sample bias and noise can inflate TE estimates. Therefore, following the approach of Benedetto et al. (2020), we compute *Effective Transfer Entropy* by subtracting Random Transfer Entropy (RTE):

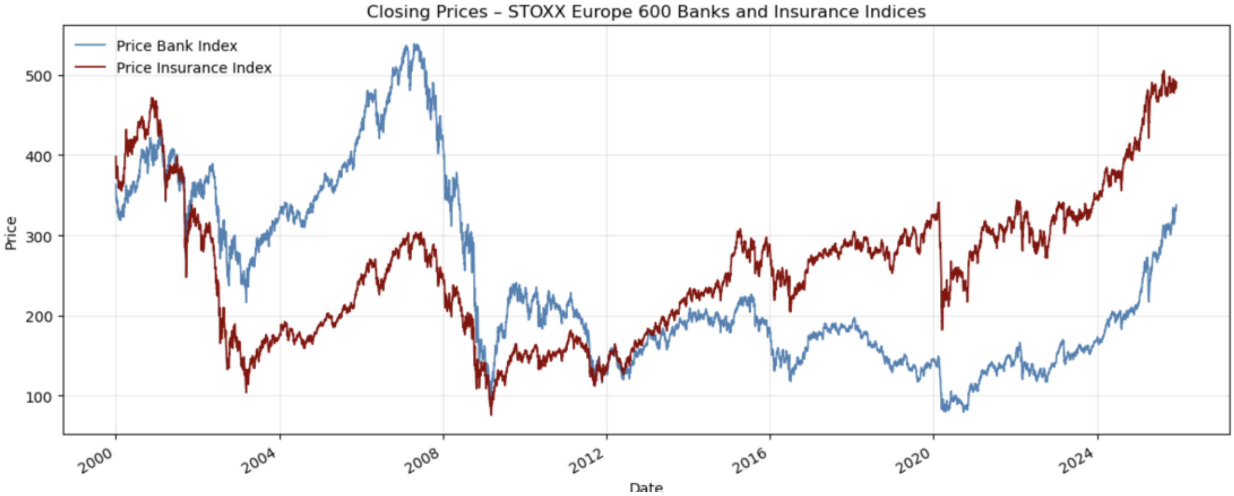
$$ETE_{x \quad y} = TE_{x \quad y} - RTE_{x \quad y} \qquad RTE_{x \quad y} = \frac{1}{N} \sum_{n=1}^{N} TE_{x^{(n)} \quad y} \tag{4}$$

where $x^{(n)}$ is obtained by i.i.d. shuffling the source series (thus destroying directed dependence while preserving the marginal distribution), and N denotes the number of shuffled replications used to approximate RTE. We set $N = 100$ shuffles, use 300 bootstrap replications, and adopt a burn-in of 50 observations. To describe the time variation of nonlinear directionality, we compute dynamic ETE using both an expanding (growing) window and a rolling windows of one year (in terms of trading days). Following Benedetto et al. (2020), we remove the first 500 dynamic estimates, since ETE computed on short early samples tends to be dominated by sampling noise. Robustness analysis to our baseline set-up in terms of quantile selection and Markov orders are reported in the subsequent dedicated robustness chapter.

4. Data and preliminary statistics

The analysis, which relies on the application of the above-mentioned methodologies (DTW, Granger Causality and TE) is based on daily data from the STOXX Europe 600 Banks Index and the STOXX Europe 600 Insurance Index, which are widely used benchmarks for the European banking and insurance sectors, respectively. The STOXX Europe 600 universe comprises major EU-wide listed firms from 17 European countries, selected on the basis of market capitalization and liquidity. The sample spans a long time horizon, from january 2000 until december 2025 and, therefore, covers several major episodes of financial stress, including the Global Financial Crisis, the European sovereign debt crisis and the COVID-19 shock. All series are aligned on common trading days to avoid distortions due to asynchronous observations. To facilitate comparability across sectors and measures and to focus on relative dynamics rather than scale effects, the series used in the DTW analysis are standardized to have zero mean and unit variance. This transformation does not alter the temporal structure of the data, but ensures that differences in magnitude do not mechanically drive similarity measures.

Table 2 reports descriptive statistics for the price indices of the European banking and insurance sectors. Both series cover the same sample length and display comparable magnitudes, although the insurance index exhibits a higher average level over the period. Price dispersion, measured by the standard deviation, is larger for the banking index, suggesting greater variability in bank equity valuations. The distribution of prices further indicates that both indices experienced wide ranges over the sample, reflecting the presence of pronounced market cycles and periods of stress, while maintaining distinct distributional characteristics across sectors. Turning to higher-order moments, both series display positive skewness (more pronounced for the banking index). This indicates that the distributions are not symmetric around their mean. Kurtosis is slightly below the Gaussian reference benchmark of three, suggesting that departures from normality in levels are driven more by asymmetry than by extreme tail thickness alone. This interpretation is confirmed by the Jarque–Bera (JB) test, whose null hypothesis is joint normality, namely zero skewness and kurtosis equal to three (see Jarque and Bera, 1980). In both cases the null is strongly rejected at the 1% significance level, implying that the unconditional distribution of the original price indices is not Gaussian. In parallel, the Ljung–Box (LB) test, whose null hypothesis is the absence of serial correlation up to the selected lag length, is also strongly significant for both series (see Ljung and Box, 1978). This points to substantial persistence in the original price levels, which is common in financial time series. Figure 2 illustrates the evolution of the price indices. A first visual inspection already suggests a substantial degree of co-movement over the analyzed period, with pronounced joint fluctuations during major market episodes, while sector-specific dynamics appears to be the drivers of the heterogenous movements outside crisis periods. To carry out our analysis, for each index we consider daily closing prices and construct a set



igure 2: STOXX Europe 600 Banks and Insurance Index – Daily prices

Table 2
Descriptive statistics of the STOXX Europe 600 Indexes

Series	Mean	Std. Dev.	Min	Max	Skewness	Kurtosis	JB	LB
Banks	239.4906	110.04154	79.27	538.80	0.869789	2.727567	857.763***	131233.540***
Insurance	258.5232	88.40044	75.54	505.45	0.515730	2.843997	301.081***	128898.436***

ote: ***, ** and * denote rejection of the null hypothesis at the 1%, 5% and 10% significance levels, respectively.

of risk measures designed to capture different dimensions of financial vulnerability.In particular, we analyze price dynamics, realized volatility, VaR and ES.

Returns are computed as log-differences of daily prices. Volatility is measured at the daily frequency, while VaR and ES are estimated at the 95% confidence level using a historical approach, a standard choice in empirical studies that balances sensitivity to downside risk with statistical reliability at the daily frequency. We focus on these three measures because they capture complementary dimensions of financial risk. Volatility provides a standard proxy for market uncertainty, whereas VaR and ES focus on the lower tail of the return distribution within a loss-based quantile framework. VaR identifies the loss threshold associated with a given confidence level, while ES captures the expected magnitude of losses conditional on that threshold being breached. Examining the joint dynamics of these measures is therefore particularly suitable for detecting cross-sector risk synchronization, especially during systemic stress episodes when tail risks and spillover effects tend to intensify. In order to apply the DTW, Granger Causality and TE algorithms to capture the dynamic evolution of these risk measures, we employ a standard practice and use a rolling window approach. Despite simplicity, a rolling window allows the risk estimates to reflect recent market conditions while providing sufficient data points for statistical reliability of the subsequent DTW application. Following the literature, we adopt a rolling window of 60 trading days, which corresponds approximately to three months of market activity. This choice is consistent with recommendations by Jorion (2007), who highlight that for daily VaR, it is common to use a historical window of 60 to 90 trading days. Alexander (2008) similarly suggests that intermediate windows of 2 to 3 months balance the trade-off between responsiveness to market shocks and statistical stability.

Since the original price levels and the raw rolling risk measures are highly persistent, the subsequent analysis is based on transformed series: in highly persistent price levels, apparent dependence may simply reflect strong internal serial dynamics rather than genuine cross-series interaction, giving rise to the classic problem of spurious relationships (see Granger and Newbold, 1974). In this regard, causality analysis is typically based on stationary autoregressive representations (see Hamilton, 1994; Lütkepohl, 2005) and entropy-based measures are more meaningfully estimated

Table 3
Unit root and stationarity tests for the transformed series

Series	DF constant	DF trend	KPSS drift	KPSS trend
Banks log-returns	-31.4818***	-31.5013***	0.2166	0.0910
Insurance log-returns	-26.4789***	-26.5243***	0.2654	0.0364
Banks first diff. Volatility	-17.1279***	-17.1273***	0.0249	0.0214
Insurance first diff. Volatility	-16.2239***	-16.2230***	0.0227	0.0209
Banks first diff. VaR	-31.0278***	-31.0258***	0.0162	0.0141
Insurance first diff. VaR	-20.1424***	-20.1412***	0.0150	0.0143
Banks first diff. ES	-22.0841***	-22.0828***	0.0108	0.0100
Insurance first diff. ES	-23.5150***	-23.5133***	0.0096	0.0095

otes: For the DF test, ***, ** and * denote rejection of the null hypothesis of a unit root at the 1%, 5% and 10% significance levels, respectively. For the KPSS test, the null hypothesis is stationarity; hence, the absence of significance stars indicates that stationarity is not rejected.

when the relevant joint and conditional distributions are sufficiently stable (see Schreiber, 2000; Kaiser and Schreiber, 2002). Therefore, prices are converted into daily log-returns, while volatility, VaR and ES are expressed in first differences. The stationarity properties of these transformed variables are then formally assessed in Table 3, that confirms that all transformed series are stationary (the ADF test strongly rejects the null of a unit root in every case, while the KPSS test does not reject the null of stationarity)

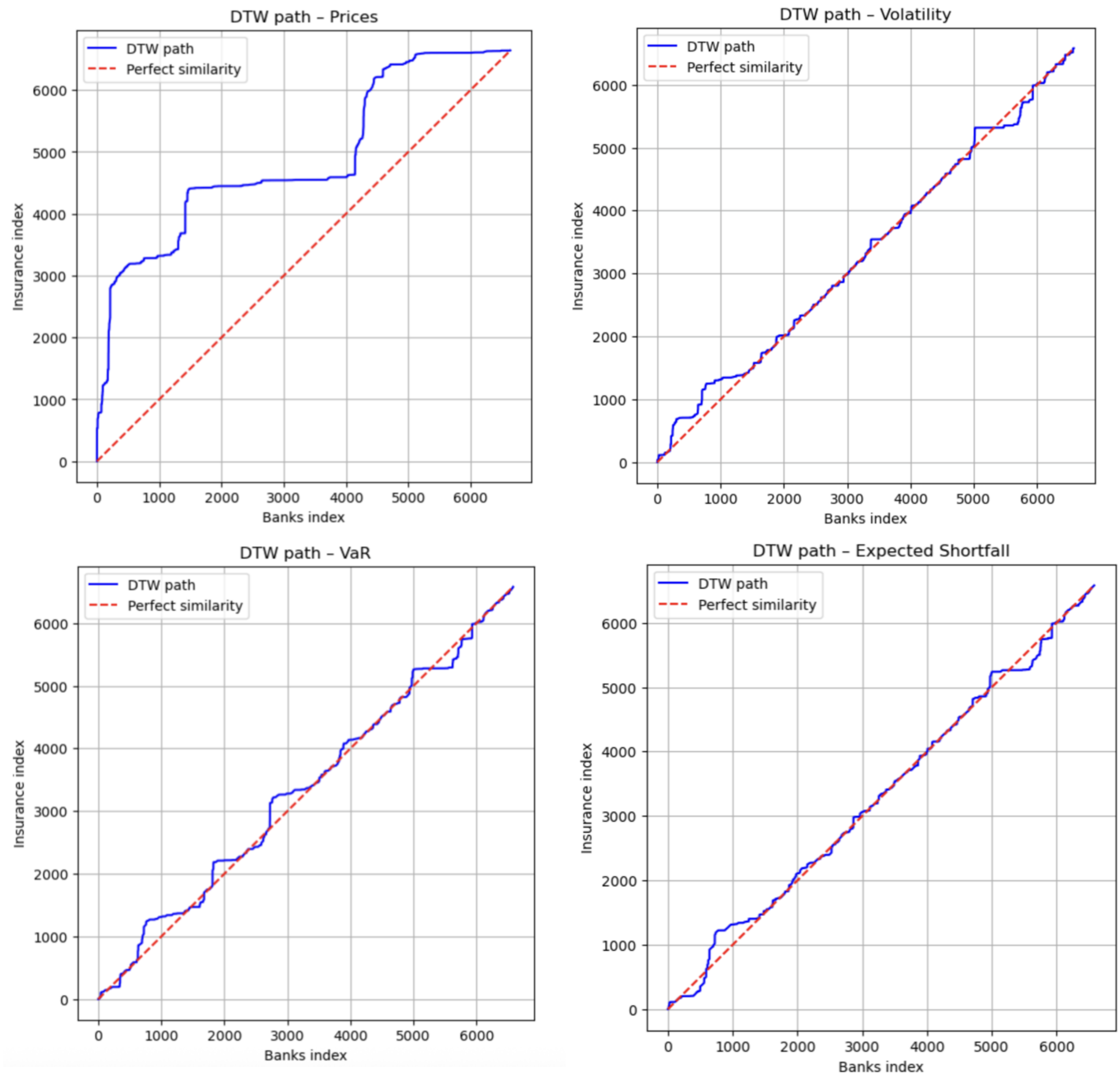
5. Results

5.1. Dynamic Time Warping

This section presents the empirical results of the analysis applied to prices and risk measures for the European banking and insurance sectors. The evidence points to a distinction between price dynamics and risk dynamics: while prices exhibit only episodic synchronization, risk measures display a markedly stronger and more persistent degree of temporal alignment across the two sectors. This pattern becomes increasingly pronounced when moving from volatility to downside risk (VaR) and, more clearly, to tail risk as captured by ES.

For each variable, we focus on the DTW distance from the diagonal (Fig. 3) as the main indicator of dynamic synchronization, as it provides a time-resolved measure of alignment that is directly interpretable in economic terms. Moreover, for the sake of completeness and in order to provide a more complete picture of the phenomenon under consideration, we report the temporal optimal alignment path from the main diagonal (Fig. 4).

A clear distinction emerges across the four panels (see Fig. 3). For volatility, VaR, and ES, the DTW paths remain very close to the diagonal over most of the sample period, indicating a high degree of temporal alignment between banks and insurance companies. Minor local deviations from the diagonal are observed, but these are short-lived and quickly reabsorbed, suggesting that risk dynamics tend to adjust quickly across the two sectors. By contrast, the DTW path for prices exhibits pronounced and persistent deviations from the diagonal, with extended horizontal and vertical segments. These features reflect sustained lead lag effects in price formation, indicating that valuation adjustments in one sector are often followed by delayed responses in the other. This pattern is consistent with the presence of informational frictions and sector-specific price discovery mechanisms. Figure 4 reports the time-varying distance of the DTW optimal alignment path from the main diagonal for prices, volatility, VaR and ES. This measure captures the instantaneous degree of temporal misalignment between the Banks and Insurance indices, where values close to zero indicate near-synchronous dynamics and larger values signal the presence of lead–lag relationships. Across all four panels, periods of systemic stress are associated with a marked reduction in the distance from the diagonal. During the Global Financial Crisis (2008 and 2009), the European Sovereign Debt Crisis (2010 to 2012), and the Covid 19 episode, the distance collapses towards zero, indicating a strong synchronization of both price and risk dynamics across the two sectors. This behavior reflects the pervasive nature of systemic shocks, which simultaneously affect financial institutions regardless of their specific business models. In contrast, outside crisis periods the distance from the diagonal exhibits larger and more persistent fluctuations, particularly for prices. Risk-based measures display a



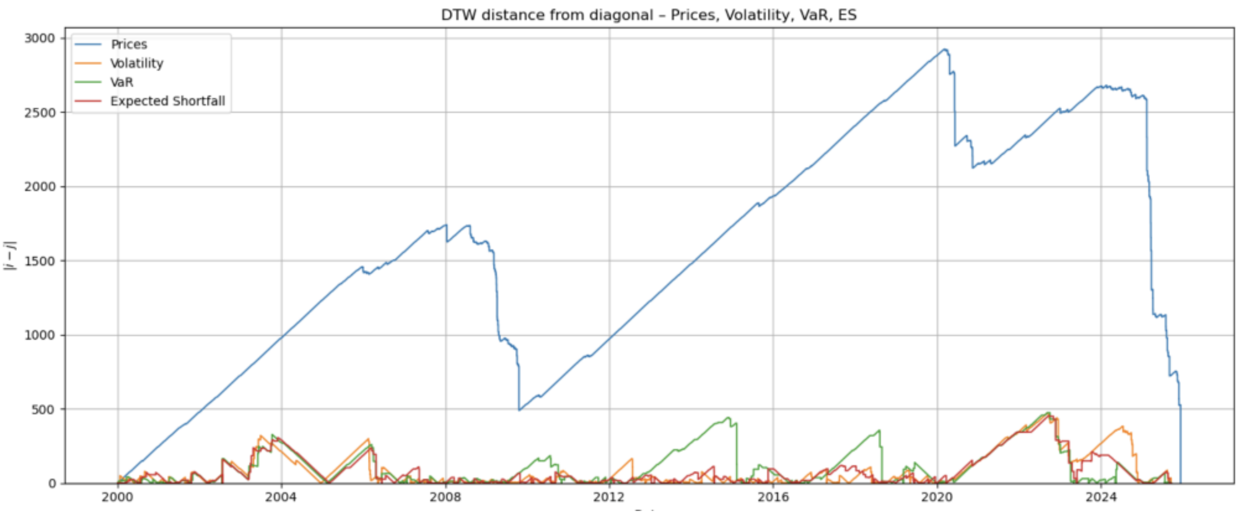
igure 3: Dynamic Time Warping Optimal lignment Path between the STOXX Europe 600 Banks and Insurance indices for prices, volatility, Value at Risk (VaR), and Expected Shortfall (ES)

different pattern: although sharp deviations from the diagonal occasionally emerge, they tend to be short-lived and are rapidly reabsorbed, indicating a faster synchronization of risk.

5.2. Information flow analysis: Granger causality and Transfer Entropy

Granger causality This section complements the DTW-based synchronization evidence with directional diagnostics based on the Granger causality and Effective Transfer Entropy, as described in the Methodology section.

Table 4 reports bivariate Granger causality tests, conducted employing HAC-robust Wald tests. For each direction and measure, the lag orders $(p\ q)$ are selected by minimizing the Bayesian Information Criterion (BIC). Our results are coherent with the dynamics highlighted by the DTW: for price log-returns, we find no evidence of linear directional predictability in either direction, suggesting that sectoral price movements do not exhibit a stable leader–follower structure over the full sample. In contrast, directional dependence emerges in risk dynamics: volatility displays strong



igure 4: Time-Varying Distance of the DTW Warping Path from the Diagonal between the STOXX Europe 600 Banks and Insurance indices for prices, volatility, Value at Risk (VaR), and Expected Shortfall (ES)

Table 4
Bivariate Granger causality tests (H C-robust Wald). Null hypothesis: $H_0 : \beta_1 = \cdots = \beta_q = 0$ (no Granger causality from source to target): $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Measure	Direction	p-value	p	q
Prices (log-returns)	Insurance Banks	0.1235	1	1
Prices (log-returns)	Banks Insurance	0.1027*	1	1
Volatility (Δ 60d)	Insurance Banks	$< 10^{-16***}$	1	3
Volatility (Δ 60d)	Banks Insurance	0.1122	9	1
VaR 95% (Δ 60d)	Insurance Banks	0.8484	1	1
VaR 95% (Δ 60d)	Banks Insurance	0.6567	1	1
ES 95% (Δ 60d)	Insurance Banks	0.0014***	1	1
ES 95% (Δ 60d)	Banks Insurance	0.0065***	2	1

predictability from Insurance to Banks (whereas the reverse direction appears to be not statistically significant). Tail-risk, proxied by ES, exhibits statistically significant bidirectional predictability, suggesting feedback effects in the propagation of extreme-risk conditions across the two sectors, confirming the strong temporal alignment path found by DTW. On the other hand, VaR dynamics do not show robust Granger causality in either direction.

Effective transfer entropy (ETE) In table 5 we report both static and dynamic ETE, considering that directed dependence in financial systems is typically state-dependent and we have a strong interest in investigating their evolution dynamics. On one hand, static estimates summarize average information flow in the entire timespan, while, on the other hand, dynamic estimates allow for time variation and therefore are able to identify episodic linkages around systemic events and large–scale shocks to the economy. Table 5 reports the TE and ETE estimates computed in the symbolic three-state setting (5% and 95% empirical quantiles) with Markov orders $k = l = 1$ and effective correction based on randomized source shuffles, following Benedetto et al. (2020). Also the TEC, as defined in the Methodology section, is reported. Positive TEC indicates net information flow from Insurance to Banks, negative values indicate the opposite. Obviously, $\mathrm{TEC}_{X\ \ Y} = -\mathrm{TEC}_{Y\ \ X}$. The static estimates are computed using $S = 100$ shuffles and $B = 300$ bootstrap replications. Bootstrap resampling is used to approximate the sampling distribution of the transfer entropy estimator and to obtain standard errors and p-values. In constructing the bootstrapped Markov chain, we adopt a burn-in period of 50 observations, i.e. the first 50 simulated transitions are discarded to mitigate initialization effects (see, e.g., Behrendt et al. (2019b) and Dimpfl and Peter (2018)). Finally, the Markov orders are

Table 5
Static Transfer Entropy (TE) and Effective Transfer Entropy (ETE) estimates.

Measure	Direction	TE	ETE	Std. error	p-value	sig	TEC
Prices (log-returns)	Insurance Banks	0.006146	0.004711	0.000599	0.000000	***	0.483514
Prices (log-returns)	Banks Insurance	0.003007	0.001640	0.000570	0.013333	**	-0.483514
Volatility (Δ 60d)	Insurance Banks	0.002105	0.000683	0.000537	0.130000		-0.446008
Volatility (Δ 60d)	Banks Insurance	0.003226	0.001783	0.000486	0.000000	***	0.446008
VaR 95% (Δ 60d)	Insurance Banks	0.001483	0.000036	0.000464	0.486667		-0.772744
VaR 95% (Δ 60d)	Banks Insurance	0.001682	0.000278	0.000515	0.250000		0.772744
ES 95% (Δ 60d)	Insurance Banks	0.002791	0.001316	0.000495	0.013333	**	-0.477225
ES 95% (Δ 60d)	Banks Insurance	0.005185	0.003718	0.000499	0.000000	***	0.477225

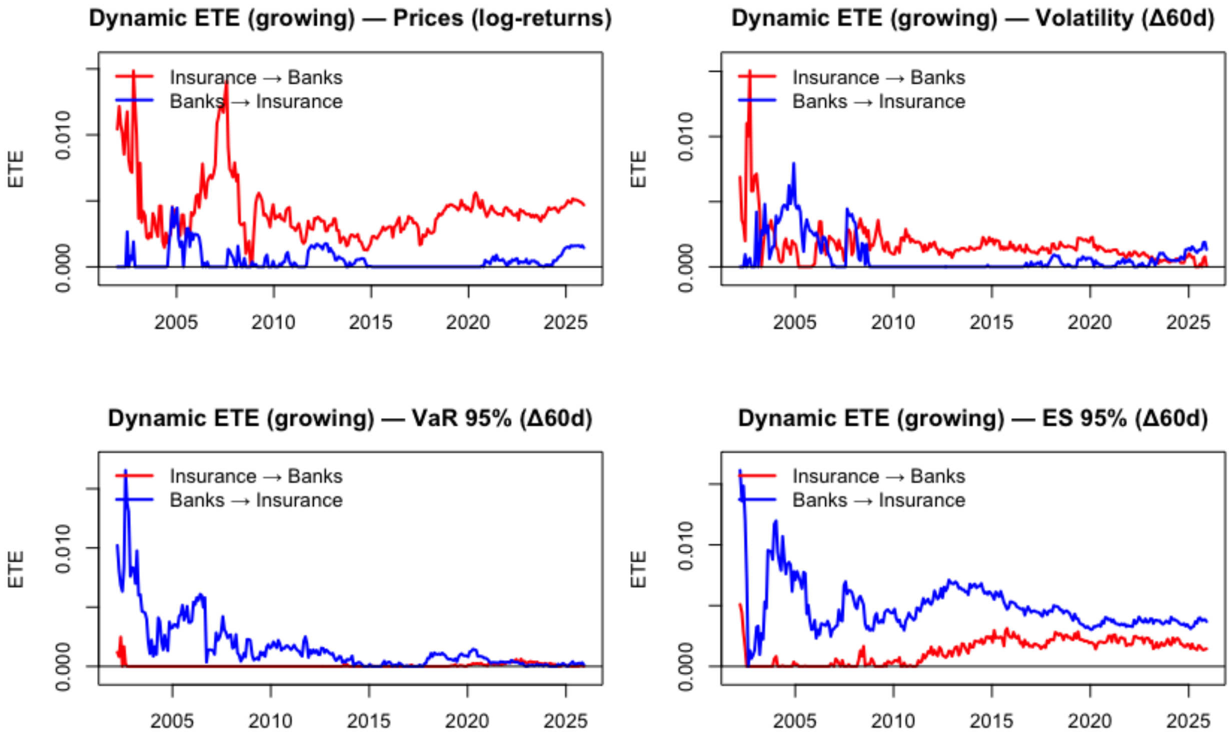
set to $k = l = 1$, which is both the most common baseline choice in the applied literature and computationally convenient. In full-sample (static) estimates, ETE delivers sharper evidence of nonlinear directionality than standard linear predictability tests, because the symbolic three-state encoding based on the 5%–95% quantiles explicitly focuses on tail-driven state transitions. Standard errors and p-values are obtained from the bootstrap distribution. Overall, ETE results are coherent and superseede the Granger Causality ones, confirming that nonlinear information transfer is not uniformly present across all measures: it is statistically significant in log-returns for both directions (with stronger magnitude from Insurance to Banks), absent for volatility innovations from Insurance to Banks but significant in the opposite direction, not significant for VaR innovations in either direction, and strongly significant for ES innovations in both directions.

Static (full-sample) ETE provides an average measure of nonlinear directionality over the entire period and may therefore not represent linkages that are concentrated in specific market phases. For this reason, we also consider a *dynamic* version of ETE designed to track how directed information flow evolves over time. We compute dynamic ETE using an expanding (growing) window updated at a monthly frequency (step size equal to 21 trading days), where at each update ETE is estimated on the subsample from the beginning of the sample up to t. Following Benedetto et al. (2020), we start reporting dynamic estimates only once at least 500 observations are available, since frequency-based transition estimates can be noisy in short initial samples. Consistent with the static calibration, we retain the same symbolic encoding (5%–95%) and Markov orders ($k = l = 1$).

The dynamic profiles, as observed in Figure 5, provide a coherent, economically meaningful reconciliation of DTW synchronization, linear Granger predictability, and nonlinear information-flow evidence (static TE analysis). The growing-window ETE for prices remains relatively small and gradually stabilizes, consistent with the Granger finding of no robust linear causality in log-returns over the full sample; nevertheless, the persistent gap in favor of Insurance Banks aligns with the static result, suggesting that return linkages—when present—operate through nonlinear, state-dependent channels rather than stable linear forecasting relationships. Second, the risk-based panels are markedly more informative and strongly time-varying. Volatility ETE exhibits pronounced early-sample spikes and then attenuates, indicating that nonlinear directional dependence is episodic and concentrated in particular regimes; importantly, the direction highlighted by static ETE (Banks Insurance) is also the direction that tends to dominate during the most informative phases, pointing to bank-sector uncertainty shocks as an upstream driver of insurer volatility regimes. Third, ES-based ETE is the most persistent and economically interpretable: dynamic ETE intensifies around well-known stress periods and remains elevated relative to other measures, consistent with DTW evidence that comovement tightens during systemic events. The bidirectional nature of ES spillovers echoes the HAC-robust Granger results (significance in both directions), while the systematically higher Banks Insurance ETE in both static and dynamic estimates supports an asymmetric narrative in which banking-sector tail-risk states transmit and help shape insurer tail-risk conditions, alongside feedback effects that become visible during severe stress. Finally, the weak and mostly non-significant VaR patterns corroborate the static table: cross-sector information transfer is more salient in extreme-tail dynamics (ES) than at the 95% VaR threshold, reinforcing the interpretation that inter-sector linkages primarily materialize through rare, high-impact states rather than moderate tail fluctuations.

6. Robustness Analysis

In this section we provide additional information in order to evaluate the robustness of our approach, focusing on DTW implementation and Transfer Entropy Analysis



igure 5: Time-varying Effective Transfer Entropy between banks and insurers (expanding window)

6.1. Dynamic Time Warping

Permutation inference via block shuffling To assess whether the DTW-based similarity between the Banks and Insurance series could arise by chance, we complement the descriptive DTW analysis with a permutation (randomization) test based on *block shuffling*. The core idea is to construct an empirical *null distribution* for the DTW distance under the hypothesis of no meaningful cross-series temporal alignment. Concretely, we keep the Banks series fixed and repeatedly generate surrogate versions of the Insurance series by randomly permuting contiguous time blocks of length L (here $L = 20$ trading days). Block shuffling preserves the within-series dependence structure over short horizons (e.g., persistence and clustering typical of financial time series) while breaking the original cross-series timing relationship. For each block-shuffled surrogate of the Insurance series we recompute the unconstrained DTW distance, obtaining B values $\{D_b\}_{b=1}^{B}$. The observed distance on the original data is denoted D^*.

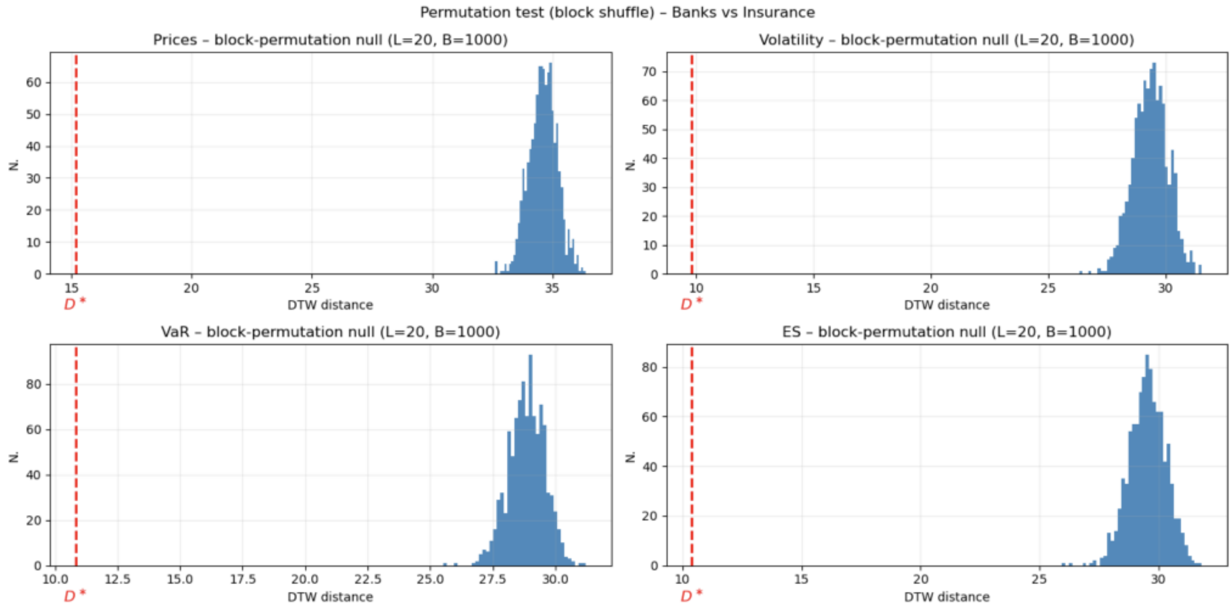
Because DTW is a dissimilarity measure, smaller distances indicate stronger alignment. We therefore use a one-sided (left-tail) Monte Carlo permutation p-value.

$$p = \frac{1 + \sum_{b=1}^{B} \mathbb{1}\left(D_b \le D^*\right)}{B+1}. \tag{5}$$

We use the standard Monte Carlo permutation p-value with finite-sample adjustment $(r+1)/(B+1)$, which yields a valid one-sided p-value when permutations are randomly sampled and prevents zero p-values at finite B (see North et al., 2002; Phipson and Smyth, 2010; Hemerik and Goeman, 2018). To preserve short-horizon serial dependence while destroying long-horizon temporal alignment, surrogates are generated by permuting contiguous blocks of length $L = 20$ trading days (block shuffle). For computational feasibility, the permutation test is carried out on downsampled series (one observation every 5 trading days); both D^* and $\{D_b\}$ are computed on the same downsampled data, and the block length is scaled accordingly ($L' = \mathrm{round}(L/5)$).

Table 6
Block-shuffle permutation test for unconstrained DTW distance (Banks vs. Insurance).

Series	D^*	p-value
Prices	15.175	0.001
Volatility	9.806	0.001
$\text{VaR}_{0.95}$	10.824	0.001
$\text{ES}_{0.95}$	10.391	0.001



igure 6: Permutation test (block shuffle)

Table 6 reports the observed DTW distances D^* and the associated one-sided Monte Carlo permutation p-values from the block-shuffle null. Across prices and all three risk measures, D^* is consistently much smaller than the typical distances obtained with the shuffled series (as observed also in Figure 6), and none of the $B = 1000$ surrogates achieves a distance as small as the observed one. Consequently, all p-values provide strong evidence that the observed DTW alignment is extremely unlikely to arise by chance.
Figure 6 visualizes the permutation distributions of the DTW distances obtained by shuffling (histograms) together with the actual distance D^* (red dashed line) computed for the real series. In all panels, the histogram mass is concentrated far to the right of D^*: none of the $B = 1000$ block-shuffled surrogates obtain a distance as small as the observed one. This graphical separation directly shows that the DTW alignment between the original banking and insurance series is substantially stronger than those obtained by shuffling.

Sakoe-Chiba band Let

$$p_t = (p_{t\,x} \;\; p_{t\,y})$$

where $p_{t\,x} \in \{1 \;\ldots\;\; N\}$ and $p_{t\,y} \in \{1 \;\ldots\;\; M\}$ be the optimal warping path and let $p_{t\,x}, p_{t\,y}$ be their components along the x- and y-axes.

A window is a global constraint that explicitly forbids warping curves to enter some region of the $(i \;\; j)$ plane. For example, according to the well-known Sakoe-Chiba band (Sakoe and Chiba (1978)), the displacement between warping

Table 7
Dynamic Time Warping (DTW) distance under unconstrained alignment versus a Sakoe–Chiba band constraint ($w = 60$ trading days).

Series	DTW (unconstrained)	DTW (Sakoe–Chiba, $w = 60$)	Loss (%)
Prices	32.53	98.57	202.97
Volatility	19.77	27.28	37.95
VaR	21.92	27.91	27.36
ES	21.17	26.67	25.96

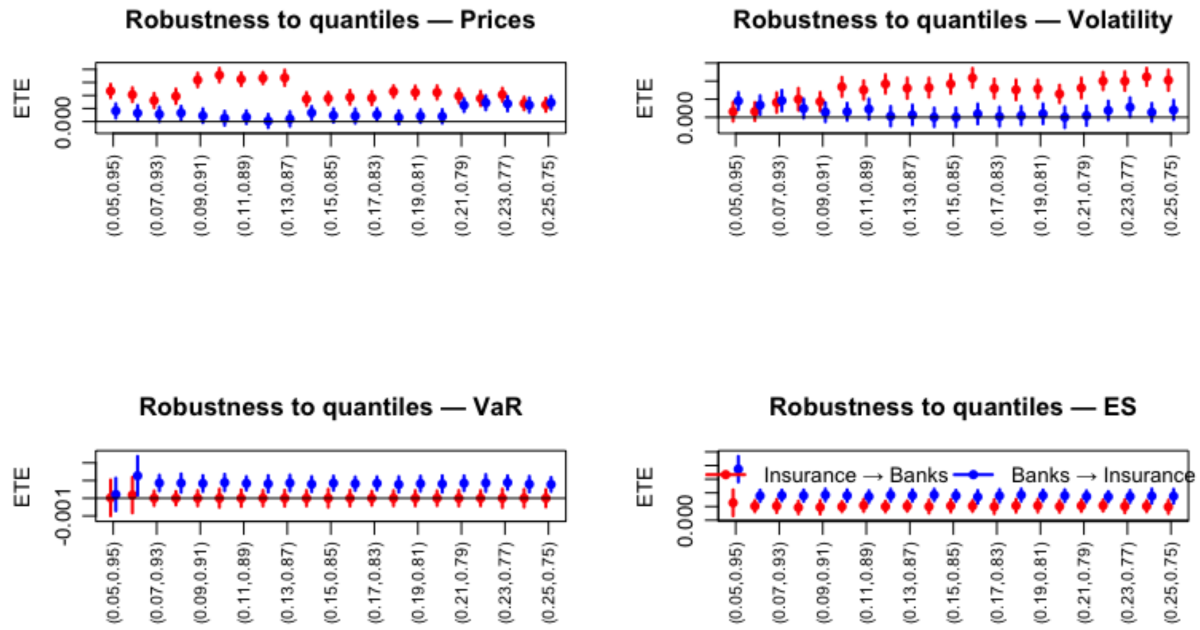
path components is bounded above by a maximum permitted time deviation, i.e.

$$|p_{t\,x} - p_{t\,y}| \quad R$$

where R is the window size. It results in an a-priori knowledge that time distortion is limited see e.g. Giorgino (2009). To quantify the distortion induced by the Sakoe–Chiba constraint, we report in Table 6 the percentage increase in DTW distance relative to the unconstrained alignment, defined as Loss( ) $= 100 \times (\mathrm{DTW}_{SC} - \mathrm{DTW}_{free})/\mathrm{DTW}_{free}$. Higher values indicate that the constraint prevents alignments that substantially reduce the DTW distance, yielding a poorer match.
Imposing a Sakoe–Chiba band of ± 60 trading days, as observable in Table 7, leads to a very large deterioration in fit for prices (loss $\approx$ 203), indicating that unconstrained DTW aligns price levels mainly through long-range warpings across distant regimes. In contrast, the increase in DTW distance for risk measures is moderate (loss $\approx$ 38 for volatility, 27 for VaR, and 26 for ES), supporting tight temporal alignment of risk dynamics within a two-month horizon, relevant for systemic risk analysis purposed.

6.2. Effective transfer entropy ETE)

To assess whether our ETE results depend on a specific discretization of the state space, we conduct a robustness exercise in the spirit of Benedetto et al. (2020) by varying the quantile thresholds used for the three-state encoding. Specifically, for each pair of symmetric tail bins ($\alpha \;\; 1 - \alpha$) with $\alpha \in [0.05 \;\; 0.25]$ (step 0.01), we re-encode each series into *low*, *middle*, and *high* states defined by the empirical α and $1 - \alpha$ quantiles, and we re-estimate static TE and ETE under the same baseline calibration (Shannon entropy, Markov orders $k = l = 1$). Effective transfer entropy is obtained via shuffle correction, i.e. subtracting the average TE computed from i.i.d. randomized source surrogates; uncertainty is quantified through bootstrap replications, yielding 95% confidence intervals and p-values. The robustness profiles confirm that the *magnitude* of ETE is sensitive to the choice of α, as expected in a state-based framework: moving from extreme tails (e.g. 5 –95) toward more central thresholds (e.g. 15 –85 or 25 –75) changes the frequency of state transitions and therefore the strength of the estimated information flow. Importantly, however, the *direction* of the information transfer remains stable across the entire grid of quantile choices for the economically relevant risk measures, most notably for ES innovations. In other words, while the estimated intensity of nonlinear dependence varies with the encoding (highlighting that the transmission mechanism can be more pronounced in tail-driven or in more central regimes), the dominant directional pattern is preserved, and statistical significance is not confined to a single, ad hoc quantile specification. Overall, these results support the interpretation that our ETE findings reflect a genuine and robust directional information-flow channel rather than an artifact of the baseline (5 95) encoding. In figure 7, each panel reports bidirectional ETE estimates computed after symbolically encoding the series into three states using symmetric quantile thresholds ($\alpha \;\; 1 - \alpha$) with $\alpha \in [0.05 \;\; 0.25]$ (step 0.01). Points denote ETE and vertical bars indicate 95% bootstrap confidence intervals; red (blue) corresponds to Insurance Banks (Banks Insurance). ETE is computed under Shannon entropy with Markov orders $k = l = 1$ and is corrected for small-sample bias via shuffle-based surrogates. While the magnitude varies with the quantile choice, the dominant direction of information flow is stable across the grid, supporting robustness to the discretization scheme.
A potential concern in symbolic transfer-entropy settings is that the magnitude of the ETE may depend on the assumed memory length of the underlying Markov process. To assess if, and to which degree, our conclusions could depend on the baseline choice $k = l = 1$, we conduct a dedicated robustness exercise by re-estimating ETE under increasing Markov orders, imposing a symmetric specification $k = l = m$ with $m \in \{1 \; \ldots \; M\}$. This follows the practice in Benedetto et al. (2020) and related work, where the order of the Markov process is varied to verify that information-flow estimates do not arise from an arbitrary parameterization of the state-transition structure.



igure 7: Robustness of static Effective Transfer Entropy (ETE) to quantile encoding

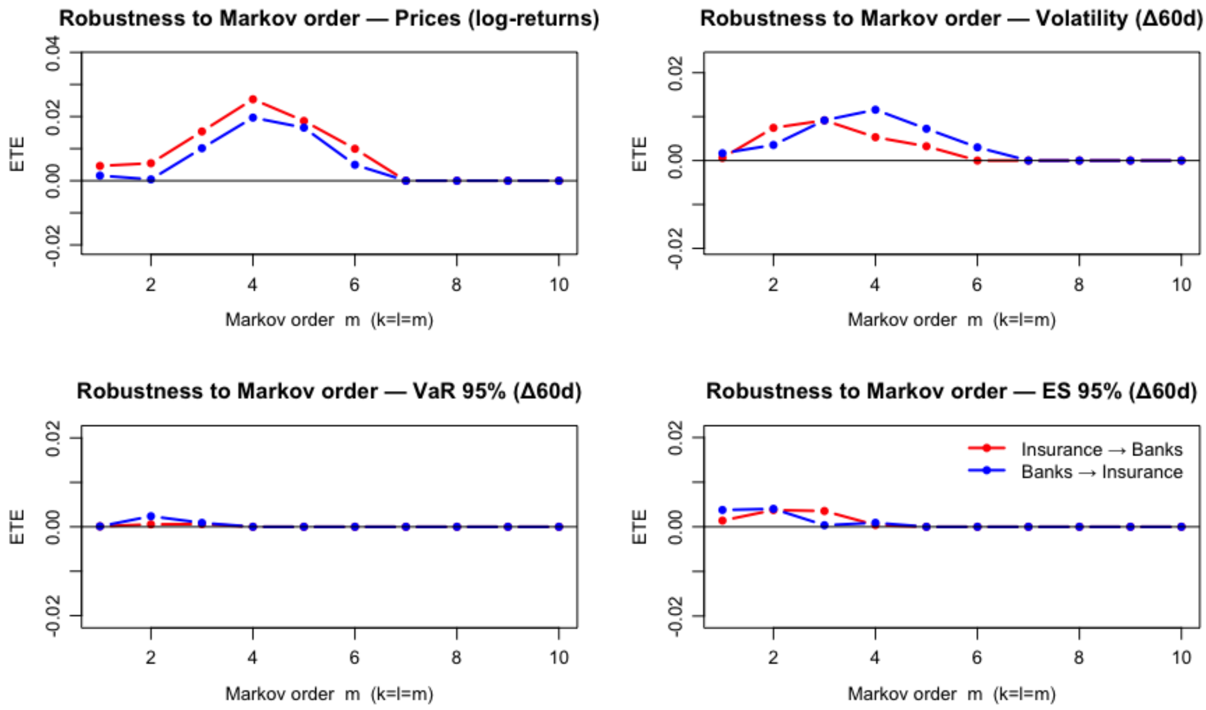
For each measure (log-returns, Δ60d volatility, Δ60d VaR, and Δ60d ES), and for each direction (Insurance Banks and Banks Insurance), we recompute the static ETE using the same three-state encoding based on symmetric quantile bins ($\alpha \;\; 1-\alpha$) and the same effective correction based on shuffled surrogates. Specifically, for each m we estimate the transfer entropy $TE^{(m)}$ with Markov orders $(k \;\; l) = (m \;\; m)$ and then construct the effective quantity

$$ETE^{(m)} = TE^{(m)} - \left[TE^{(m)}_{\text{shuffled}}\right] \tag{6}$$

where the expectation is approximated by averaging transfer-entropy estimates computed on i.i.d. shuffled versions of the source series (holding the target fixed). For each configuration, we report the point estimate of $ETE^{(m)}$ together with 95% uncertainty bands derived from the bootstrap distribution of the estimator.

This robustness check serves two purposes. First, it verifies whether the *direction* of information flow is stable once longer state histories are allowed, thereby reducing the risk that apparent asymmetries are artifacts of an overly short memory specification. Second, it documents how the *magnitude* of ETE evolves with m: in many applications, ETE typically decreases as m increases.

The pattern emerged in figure 8 is interesting and informative. First, the vanishing of ETE for $m \geq 6$ is consistent with the well-known "curse of dimensionality" in symbolic transition estimation: higher-order Markov embeddings require estimating a much larger state-space of joint histories, which quickly produces sparse transition counts and drives the effective, shuffle-corrected information flow towards zero. Hence, large-m estimates are not a more stringent test of dependence, but a regime where the estimator becomes conservative due to data fragmentation. Second, the substantive conclusions are unaffected in the relevant range. For the risk measures - and especially for ES - the directional signal is essentially a *short-memory* phenomenon: the economically meaningful information transfer is



igure 8: Robustness to Markov order

carried by a small number of recent tail-state transitions, and it is precisely at low orders ($m = 1$–3) that ETE is non-negligible and directionally interpretable. VaR remains close to zero for all m, reinforcing the earlier result that cross-sector predictability is more salient in volatility and in the extreme tail than at the 95% VaR threshold. Prices exhibit a mild hump at intermediate orders, but the effect is not persistent and disappears as soon as the embedding dimension increases, supporting the view that return-level linkages are episodic and not indicative of a stable nonlinear information-transfer mechanism.

Overall, the key economic message - that nonlinear directional dependence is concentrated in the risk channel (particularly ES), while VaR and returns provide limited evidence of sustained information transfer - remains confirmed within the low-order region where the estimator is reliable.

6.3. Comparison with state-of-the-art systemic risk measures

It is now interesting to assess the results of the proposed framework with evidence obtained by widely adopted methodologies. Established econometric approaches (e.g., CoVaR-based measures, MES/SRISK indicators, and VAR-based connectedness and spillover indices) typically rely on a fixed temporal alignment and focus either on contemporaneous dependence or on predictive linkages estimated on a common time grid. When applied to the banking–insurance nexus, these approaches generally detect heightened interconnectedness primarily during major crisis episodes, while providing weaker or episodic evidence of linkages outside stress periods. Moreover, by construction, they implicitly assume that comparable stress signals materialize contemporaneously across sectors; as a result, delayed or asynchronous responses may be partially obscured.

Our DTW-based analysis complements this literature by explicitly relaxing the assumption of synchronous adjustment. Even in periods where standard correlation or VAR-based measures would suggest moderate interdependence, DTW reveals a tight alignment in the *shape* of risk dynamics—particularly for volatility and ES—once temporal distortions are accounted for. This indicates that cross-sector systemic vulnerability may manifest through similar stress trajectories unfolding at different speeds, rather than through contemporaneous co-movements alone.

Table 8
Comparison of systemic risk detection capabilities across methodologies.

Methodology	Temporal alignment	synchronous stress	Tail-risk focus	Directionality	Nonlinearity
Correlation / Euclidean distance	Fixed	×	×	×	×
V R-based spillover indices	Fixed	×	∘		×
Granger causality (linear)	Fixed	×	×		×
CoVaR / ΔCoVaR	Fixed	×		∘	×
MES / SRISK	Fixed	×		×	∘
CDS / equity-based contagion	Fixed	×	∘		∘
DTW on prices (this paper)	Flexible		×	×	
DTW on risk measures (this paper)	Flexible			×	
DTW + Granger + ETE (this paper)	Flexible				

Legend: = addressed; ∘ = limited/indirect; × = not addressed.

Importantly, integrating DTW with Granger causality and ETE further clarifies how similarity translates into directional dependence. Table 8 summarizes how the proposed framework compares with standard systemic risk methodologies along key analytical dimensions. While conventional linear methods identify limited predictability in price returns, our results show that risk-based measures, especially tail-risk metrics, exhibit economically meaningful feedback effects that persist even when nonlinear and tail-driven information flows are isolated. In this sense, the proposed framework does not replace existing systemic risk indicators; rather, it extends them by uncovering latent synchronization and spillover patterns that remain hidden under fixed-alignment methodologies. Overall, the comparison suggests that standard approaches may underestimate the degree of bank–insurance interconnectedness when stress transmission is asynchronous, whereas the DTW-based perspective provides a more nuanced and temporally flexible representation of cross-sector systemic risk.

7. Conclusions

This paper has examined the temporal synchronization of risk between the European banking and insurance sectors using a Dynamic Time Warping (DTW) framework applied to prices, volatility, VaR and ES. The analysis provides a dynamic perspective on how risk co-movements between these two core components of the financial system evolve over time and across different layers of risk. The obtained results reveal a clear asymmetry between price dynamics and risk dynamics: systemic risk synchronization emerges primarily in risk measures rather than in price movements. Price synchronization is largely episodic and concentrated around major systemic events, such as the Global Financial Crisis, the European sovereign debt crisis, and the Covid-19 shock. Outside these periods, price movements tend to diverge, reflecting sector-specific factors and heterogeneous adjustment paths. This suggests that price-based co-movements alone may offer only a partial and, at times, delayed picture of systemic vulnerability.
In contrast, measures of risk display a markedly stronger and more persistent degree of temporal alignment. Volatility moves closely across the banking and insurance sectors, indicating that uncertainty shocks tend to propagate rapidly and broadly across financial intermediaries. A similar pattern emerges for VaR, where episodes of sectoral misalignment are typically short-lived and followed by rapid re-synchronization during periods of stress, suggesting that downside risk quickly acquires a system-wide dimension. The strongest degree of alignment is observed for ES, where extreme losses appear closely linked across sectors, with recurrent convergence, even outside the most acute crisis phases. These findings reinforce the view that cross-sector vulnerability materializes mainly through the joint dynamics of risk measures rather than through persistent price co-movements. To complement the DTW evidence, we explicitly assessed whether co-movements also entail *directional predictability* in the sense of incremental information content. For price log-returns Granger tests do not reject the null in either direction, and static ETE indicates only modest, albeit statistically detectable, nonlinear information flow. In other words, sectoral return dynamics do not exhibit a stable leader–follower structure and strong predictive content. The picture changes decisively for risk dynamics. First, Granger causality points to strong directional predictability in volatility and ES innovations. Secondly - and crucially - ETE does not merely echo Granger: it corroborates that the dependence is not an artifact of linear forecasting in a large sample. Static ETE is *strongly significant* for volatility and expected shortfall in both directions (while remaining

negligible for VaR) and indicates that cross-sector linkages operate through *state transitions in risk regimes* rather than through average return propagation. This is exactly the channel DTW is designed to highlight: synchronization is weak and unstable in prices, but it tightens sharply in risk measures when markets enter stress phases.

Taken together, the results overturn the benign interpretation that DTW co-movements are just contemporaneous correlation. The conclusion points out the existence of a interdependence among sectors, with directional information flow: risk conditions in one sector contain predictive content for subsequent risk reconfigurations in the other, especially in the tail-sensitive ES metric. This provides a coherent triad of evidence: DTW documents strong synchronization, Granger detects linear directional predictability in risk innovations, and ETE confirms that a non-trivial component of this predictability reflects genuine (nonlinear) information transfer concentrated in risk-state dynamics rather than in prices. Although DTW–Granger–ETE evidence coherently points to tight cross-sector coupling in volatility and tail-risk (especially ES), it is worth noting that these are market measures of vulnerability and uncertainty. In fact, they capture the magnitude of losses in adverse scenarios but they do not *per se* necessarily translate into realized balance-sheet losses and subsequent defaults.

Lastly, being our analysis conducted on indices, the results capture aggregate synchronization and spillovers but not reflect firm-specific heterogeneity and dynamics, useful to investigate which entities could be more exposed among the sector. From a macroprudential perspective, these findings support the importance of monitoring banking and insurance sectors jointly, as systemic vulnerability may emerge through synchronized risk dynamics even when price signals remain weak or fragmented.

8. Policy implications

In the midst of the 2008/2009 financial crisis, and from subsequent analyses of its causes, it became clear that better cooperation between Authorities for financial supervision would be key to prevent and mitigate similar threats for European and global economies arising in the future and effectively address potential spill-over effects among financial sectors. Our results confirm and support the importance of this approach, providing additional evidence to investigate the linkage between banking and insurance sectors. The evidence in this paper suggests that, for financial-stability surveillance, a relevant object to track is not primarily the co-movement of prices, but the co-evolution of *risk states*. Across methods, synchronization is materially stronger in risk measures — especially tail risk — than in price returns. This could imply that prices may fail to provide a timely and reliable early-warning signal of cross-sector stress. By contrast, ES emerges as the most informative metric of interdependence: it targets the severity of losses in the far tail and is therefore naturally aligned with the macroprudential objective of evaluating the system resilience under adverse (and unexpected) scenarios. In this sense, the strongest and most robust dependence observed in ES suggests that banks and insurers tend to converge toward a common *tail-risk regime*. A key implication is that macro-surveillance frameworks should explicitly treat banking and insurance as a joint risk perimeter, when monitoring extreme adverse conditions. The results support the use of integrated indicators that detect whether the two sectors are converging toward the same tail risk regime. The DTW framework provides a concrete tool for real-time surveillance because it relaxes the often unrealistic assumption that stress must arrive simultaneously across sectors.

DTW is able to (i) accommodate time-varying lead–lag patterns, (ii) detect common stress even when it is temporally misaligned, and (iii) deliver an interpretable measure of *stress proximity* that can be updated as new observations arrive. This makes DTW particularly suitable for narrative and scenario-based stress monitoring: in adverse environments, employing this methodology, supervisors and macroprudential authorities can appreciate how the timing of stress transmission is evolving across sectors, even if unfolds at different speeds. Finally, causal evidence should be interpreted with caution. Granger causality and ETE both indicate statistically significant *directional dependence* in risk dynamics with a particular focus on volatility and ES, while not claiming structural causality in the policy-counterfactual sense. The appropriate interpretation is therefore the information-theory one of predictive spillovers and information feedbacks, not a claim of invariant causal primitives. This not weakens the message: even without structural identification, directional predictability in tail risk is a powerful tool and message for the macro-surveillance: when ES-based indicators signal that banks and insurers are converging toward a shared tail-risk state, cross-sector monitoring should intensify, as the underlying vulnerability is likely to become system-wide rather than remain sector-specific.

References

[1] Acharya, V.V., Engle, R.F., Richardson, M., 2012. Capital shortfall: a new approach to ranking and regulating systemic risks. Am. Econ. Rev. 102(3), 59–64. https://doi.org/10.1257/aer.102.3.59.

[2] Adrian, T., Brunnermeier, M.K., 2016. CoVaR. Am. Econ. Rev. 106(7), 1705–1741. https://doi.org/10.1257/aer.20120555.

[3] Alexander, C., 2008. *Market Risk Analysis, Volume IV: Value-at-Risk Models*. Wiley.

[4] Barnett, L., Barrett, A.B., Seth, A.K., 2009. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys. Rev. Lett. 103(23), 238701. https://doi.org/10.1103/PhysRevLett.103.238701.

[5] Basel Committee on Banking Supervision, 2010. The Basel Committee's response to the financial crisis: report to the G20 (BCBS Publication No. 179). Bank for International Settlements. Available at: https://www.bis.org/publ/bcbs179.htm (accessed 15 February 2026).

[6] Basel Committee on Banking Supervision, 2011. Basel III: a global regulatory framework for more resilient banks and banking systems (December 2010, rev. June 2011) (BCBS Publication No. 189). Bank for International Settlements. Available at: https://www.bis.org/publ/bcbs189.htm (accessed 15 February 2026).

[7] Behrendt, S., Dimpfl, T., Peter, F.J., Zimmermann, D.J., 2019. RTransferEntropy — Quantifying information flow between different time series using Effective Transfer Entropy. SoftwareX 10, 100265. https://doi.org/10.1016/j.softx.2019.100265.

[8] Benedetto, F., Mastroeni, L., Quaresima, G., Vellucci, P., 2020. Does OVX affect WTI and Brent oil spot variance? Evidence from an entropy analysis. Energy Econ. 89, 104815. https://doi.org/10.1016/j.eneco.2020.104815.

[9] Bernal, O., Gnabo, J.-Y., Guilmin, G., 2014. Assessing the contribution of banks, insurance and other financial services to systemic risk. J. Bank. Financ. 47, 270–287. https://doi.org/10.1016/j.jbankfin.2014.05.030.

[10] Berndt, D.J., Clifford, J., 1994. Using dynamic time warping to find patterns in time series. In: *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*.

[11] Bierth, C., Irresberger, F., Weiß, G.N.F., 2015. Systemic risk of insurers around the globe. J. Bank. Financ. 55, 232–245. https://doi.org/10.1016/j.jbankfin.2015.02.014.

[12] Billio, M., Getmansky, M., Lo, A.W., Pelizzon, L., 2012. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. J. Financ. Econ. 104(3), 535–559. https://doi.org/10.1016/j.jfineco.2011.12.010.

[13] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1.

[14] Brownlees, C.T., Engle, R.F., 2017. SRISK: a conditional capital shortfall measure of systemic risk. Rev. Financ. Stud. 30(1), 48–79. https://doi.org/10.1093/rfs/hhw060.

[15] Banulescu, G.-D., Dumitrescu, E.-I., 2015. Which are the SIFIs? A Component Expected Shortfall approach to systemic risk. J. Bank. Financ. 50, 575–588. https://doi.org/10.1016/j.jbankfin.2014.01.037.

[16] Chen, R., Cummins, J.D., Viswanathan, K.S., Weiss, M.A., 2014. Systemic risk and the interconnectedness between banks and insurers: an econometric analysis. J. Risk Insur. 81(3), 623–652. https://doi.org/10.1111/j.1539-6975.2012.01503.x.

[17] Cummins, J.D., Weiss, M.A., 2014. Systemic risk and the U.S. insurance sector. J. Risk Insur. 81(3), 489–527. https://doi.org/10.1111/jori.12039.

[18] Denkowska, D., Wanat, S., 2021. Dynamic time warping algorithm in modeling systemic risk in the European insurance sector. Entropy 23(8), 1022. https://doi.org/10.3390/e23081022.

[19] Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. J. Am. Stat. Assoc. 74(366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531.

[20] Diebold, F.X., Yilmaz, K., 2015. *Financial and Macroeconomic Connectedness: A Network Approach to Measurement and Monitoring*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338290.001.0001.

[21] D'Urso, P., De Giovanni, L., Massari, R., 2021. Trimmed fuzzy clustering of financial time series based on dynamic time warping. Ann. Oper. Res. 299, 1379–1395. https://doi.org/10.1007/s10479-019-03284-1.

[22] Eling, M., Pankoke, D.A., 2016. Systemic risk in the insurance sector: a review and directions for future research. Risk Manag. Insur. Rev. 19(2), 249–284. https://doi.org/10.1111/rmir.12062.

[23] Engle, R.F., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. 20(3), 339–350. https://doi.org/10.1198/073500102288618487.

[24] European Systemic Risk Board, 2015. Report on systemic risks in the EU insurance sector. Available at: https://doi.org/10.2866/85741.

[25] European Systemic Risk Board, 2018. Macroprudential provisions, measures and instruments for insurance. Available at: https://doi.org/10.2849/816957.

[26] European Systemic Risk Board, 2025. ESRB Annual Report 2024. Available at: https://doi.org/10.2849/1040452.

[27] European Systemic Risk Board, 2025. EU Non-bank Financial Intermediation Risk Monitor 2025 (No 10 / September 2025). Available at: https://doi.org/10.2849/2248943.

[28] Financial Stability Board, 2019. Finalisation of IAIS holistic framework for the assessment and mitigation of systemic risk in the insurance sector. Available at: https://www.fsb.org/uploads/P141119-1.pdf.

[29] Franses, P.H., Wiemann, T., 2020. Intertemporal similarity of economic time series: an application of dynamic time warping. Comput. Econ. 56, 1127–1153. https://doi.org/10.1007/s10614-020-10030-2.

[30] Gassouma El M., 2022. Investigating similarities between Islamic and conventional banks in GCC countries: a dynamic time warping approach. Int. J. Islam. Middle East. Financ. Manag. 16(1). https://doi.org/10.1108/IMEFM-11-2020-0565.

[31] Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37(3), 424–438. https://doi.org/10.2307/1912791.

[32] Granger, C.W.J., Newbold, P., 1974. Spurious regressions in econometrics. J. Econom. 2(2), 111–120. https://doi.org/10.1016/0304-4076(74)90034-7.

[33] Grzejszczak, T., Probierz, E., Galuszca, A., Stuglik, K., Jedrasiak, K., Wisniewski, T., 2022. Dynamic time warping in financial data — modification of algorithm in context of stock market similarity analysis. Eur. Res. Stud. J. 25(1), 967–979.
[34] Hamilton, J.D., 1994. *Time Series Analysis*. Princeton University Press. https://doi.org/10.1515/9780691218632.
[35] Hao, X., Zhou, Q., Liu, J., Chen, Z., 2025. Systemic risk among China's financial sectors: Novel evidence from trivariate CoVaR based on vine copulas. Res. Int. Bus. Financ. 78, 102968. https://doi.org/10.1016/j.ribaf.2025.102968.
[36] Harrington, S.E., 2009. The financial crisis, systemic risk, and the future of insurance regulation. J. Risk Insur. 76(4), 785–819. https://doi.org/10.1111/j.1539-6975.2009.01330.x.
[37] Hemerik, J., Goeman, J., 2018. Exact testing with random permutations. TEST 27(4), 811–825. https://doi.org/10.1007/s11749-017-0571-1.
[38] Idier, J., Lamé, G., Mésonnier, J.-S., 2014. How useful is the Marginal Expected Shortfall for the measurement of systemic exposure? A practical assessment. J. Bank. Financ. 47, 134–146. https://doi.org/10.1016/j.jbankfin.2014.06.022.
[39] Jarque, C.M., Bera, A.K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Econ. Lett. 6(3), 255–259. https://doi.org/10.1016/0165-1765(80)90024-5.
[40] Jorion, P., 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*, 3rd ed. McGraw-Hill.
[41] Kaiser, A., Schreiber, T., 2002. Information transfer in continuous processes. Physica D 166(1–2), 43–62. https://doi.org/10.1016/S0167-2789(02)00432-3.
[42] Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? J. Econom. 54(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y.
[43] Lim, T., Ng, H., 2022. Financial portfolio management based on shape-based unsupervised machine learning: a dynamic time warping barycenter averaging approach to international markets and periods of downside event risks. J. Investing 32(1), 74–96. https://doi.org/10.3905/joi.2022.1.251.
[44] Liu, Z., Zhang, Y., Chen, H., 2025. Research on systemic risk measurement based on temporal financial knowledge graph: evidence from China. Pac.-Basin Financ. J. 94, 102924. https://doi.org/10.1016/j.pacfin.2025.102924.
[45] Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series models. Biometrika 65(2), 297–303. https://doi.org/10.1093/biomet/65.2.297.
[46] López-Espinosa, G., Moreno, A., Rubia, A., & Valderrama, L. (2015). Systemic risk and asymmetric responses in the financial industry. *Journal of Banking Finance*, 58, 471–485. doi:10.1016/j.jbankfin.2015.05.004
[47] Lütkepohl, H., 2005. *New Introduction to Multiple Time Series Analysis*. Springer. https://doi.org/10.1007/978-3-540-27752-1.
[48] Mastroeni, L., Mazzoccoli, A., Quaresima, G., Vellucci, P., 2021. Decoupling and recoupling in the crude oil price benchmarks: an investigation of similarity patterns. Energy Econ. 94, 105036. https://doi.org/10.1016/j.eneco.2020.105036.
[49] Mühlnickel, J., & Weiß, G. N. F. (2015). Consolidation and systemic risk in the international insurance industry. *Journal of Financial Stability*, 18, 187–202. https://doi.org/10.1016/j.jfs.2015.04.005
[50] Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55(3), 703–708. https://doi.org/10.2307/1913610.
[51] North, B.V., Curtis, D., Sham, P.C., 2002. A note on the calculation of empirical P values from Monte Carlo procedures. Am. J. Hum. Genet. 71(2), 439–441. https://doi.org/10.1086/341527.
[52] Phipson, B., Smyth, G.K., 2010. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. Stat. Appl. Genet. Mol. Biol. 9(1), Article 39. https://doi.org/10.2202/1544-6115.1585.
[53] Rodríguez-Moreno, M., Peña, J. I., 2013. Systemic risk measures: The simpler the better? J. Bank. Financ. 37(6), 1817–1831. https://doi.org/10.1016/j.jbankfin.2012.07.010.
[54] Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71(3), 599–607. https://doi.org/10.1093/biomet/71.3.599.
[55] Schreiber, T., 2000. Measuring information transfer. Phys. Rev. Lett. 85(2), 461–464. https://doi.org/10.1103/PhysRevLett.85.461.
[56] Wang, L., Hao, Y., Jing, Z., Zhang, J., 2024. Spillover effects of Asian banking sector on systemic risk in the insurance sector: based on a double-CoVaR model. Singap. Econ. Rev. 69(8), 2363–2390. https://doi.org/10.1142/S0217590822500795.
[57] Weiß, G. N. F., & Mühlnickel, J. (2014). Why do some insurers become systemically relevant? *Journal of Financial Stability*, 13, 95–117. https://doi.org/10.1016/j.jfs.2014.05.001

ROMA TRE
UNIVERSITÀ DEGLI STUDI