Resources Policy 112 (2026) 105820

Contents lists available at ScienceDirect

Resources Policy

journal homepage: www.elsevier.com/locate/resourpol

Investigating commodity price interdependence with Granger causality networks

Roberto Esposti

Department of Economics and Social Sciences, Università Politecnica delle Marche Ancona, Italy

ARTICLE INFO

JEL classification:
C32
Q02
O13

Keywords:
Commodity prices
Price interdependence
Granger causality
Network analysis
Sparse VAR models

ABSTRACT

This paper investigates the interdependence among commodity prices. Although the literature on this topic is extensive, it often struggles with the issue of dimensionality. To address this challenge, the paper proposes a solution based on network theory and a sparse estimation approach. The analysis relies on a large dataset of about 50 monthly commodity prices (1980–2024), grouped into energy, metals, agriculture, food, and other raw materials. A Commodity Price Network is constructed via Granger causality tests, using both pairwise and sparse VAR models, applied to price levels and first differences to account for potential non-stationarity. The results show that network topology is sensitive to the methodological approach. Nonetheless, consistent patterns emerge: metals and energy commodities maintain a central role, alongside certain agricultural products. In the sparser network configuration, energy commodities exhibit an average outgoing modularity 2.2 times higher than that of metals. Conversely, metals display an average ingoing modularity 5 % greater than energy commodities. Notably, iron shows 2.5 times more outgoing links than crude oil.

1. Introduction

This paper investigates the interdependence among commodity and natural resource market prices. The literature on this topic is extensive but often struggles with the issue of dimensionality when using price panels that span long time series and include a large number of commodities. To address this challenge, the paper proposes a novel methodological framework based on Granger Causality Networks (GCN), which combines Granger causality testing with Network Analysis (NA) to deal with large dimensionality and heterogeneous stochastic properties and to uncover and visualize the structure of price interdependence.

Commodity price interdependence has long attracted scholarly attention due to its implications for inflation transmission, supply chain vulnerability, and portfolio risk management (Boako et al., 2020; Ding et al., 2021; Esposti, 2021, 2024a,b; Fry-McKibbin et al., 2023; Kozian et al., 2025). Price shocks in one or more commodities can propagate across sectors and borders, affecting production costs, consumer prices, and macroeconomic stability. Moreover, the strategic importance of certain commodities, especially those deemed critical for national security (European Commission, 2024), further underscores the need for robust analytical tools to monitor and understand these interdependencies.

Despite its relevance, the empirical investigation of commodity price interdependence remains methodologically challenging. First, commodities may be connected either directly along the same supply chains or indirectly via other commodities or through economy-wide (or system) linkages. Therefore, interdependence may arise even among highly dissimilar commodities, involving a large number of items and complex relationships. Under such circumstances, traditional multivariate approaches, such as Vector Autoregressive (VAR) models, become computationally intensive and increasingly difficult to interpret as both the number of variables and the length of time lags grow. This dimensionality issue may be reduced ex-ante by limiting the analysis to a subgroup of commodities (selected by homogeneity or relevance) or by relying on common dimensionality reduction techniques (e.g., principal component analysis, factor analysis, cointegration, machine learning) (Shahzad et al., 2021; Esposti, 2021, 2024a, 2024b). However, such dimensionality reduction may be arbitrary and risk oversimplifying the underlying complexity, potentially leading to the loss of valuable information.

Second, commodity prices often exhibit heterogeneous stochastic properties, making it difficult to assume a common Data Generating Process (DGP) across the entire panel. This heterogeneity complicates the application of unified econometric models and calls for flexible approaches that can accommodate both stationary and non-stationary

E-mail address: r.esposti@staff.univpm.it.

https://doi.org/10.1016/j.resourpol.2025.105820
Received 26 July 2025; Received in revised form 4 November 2025; Accepted 10 December 2025
Available online 23 December 2025

series.

To address these challenges, this paper proposes an approach that integrates two empirical strategies for constructing Granger Causality Networks (GCNs) among commodity prices. Granger causality tests are first applied to a broad set of commodity price series. Based on these results, Commodity Price Networks (CPNs) are defined and analysed using tools from network theory. To tackle the issue of dimensionality, two alternative estimation methods are employed: Pairwise VAR (PW-VAR) and Sparse VAR (S-VAR). Both approaches allow for scalable inference while preserving the richness of the interdependence structure, although they differ significantly in their statistical foundations and the trade-offs they involve. Additionally, to account for stochastic heterogeneity, the analysis is conducted on both price levels and first differences, enabling robustness checks across different stationarity conditions.

This empirical strategy contributes to the literature by offering a novel and scalable framework for analysing commodity price interdependence. It is applied to a dataset of 49 commodity prices, grouped into five categories: energy, metals, agriculture, food, and other raw materials. Prices are observed monthly over a 45-year period, from January 1980 to December 2024, resulting in 540 observations per commodity. The resulting 49 × 49 matrix of connections provides a rich and interpretable representation of the commodity price system. To the best of our knowledge, this is the first application of network analysis grounded in Granger causality to such a comprehensive panel of commodity prices.

The remainder of the paper is structured as follows. Section 2 introduces the topic of commodity price interdependence, focusing on key stylized facts, recent literature, and policy relevance. Section 3 presents the methodological approach, highlighting its advantages and innovative aspects. Section 4 provides a detailed description of the dataset. Section 5 presents the results and assesses their robustness by comparing findings across different sets of series and network configurations. Section 6 discusses the main policy implications stemming from the proposed methodology and the obtained results. Finally, Section 7 concludes with some methodological reflections.

2. Commodity price interdependence and network analysis: a literature review

2.1. The key stylized facts

Over the years, a substantial body of empirical literature has emerged on the multivariate analysis of commodity prices (Byrne et al., 2020).[1] Especially after the 2007–2008 price turmoil, many empirical studies have investigated the common determinants of commodity price dynamics. These contributions point to several possible sources of co-movement in prices. However, not all of these sources necessarily imply interdependence among commodity prices, as they may simply reflect general common drivers related either to the real economy (e.g., population and economic growth on the demand side; increasing resource scarcity on the supply side) or to financial markets (e.g., growing speculative activity, exchange rate volatility, etc.) (OECD, 2010; Piot-Lepetit and M'Barek, 2011). More recent studies also highlight an additional general driver that may generate common dynamics across commodity prices, when expressed in nominal terms, namely, the rapid change in the inflation rate (Amaglobeli et al., 2022; Garzón and Hierro, 2022; Esposti, 2024a).

However, genuine cross-market interdependence, i.e., causal relationships across commodity prices, extends beyond simple price co-movement. Interdependence implies a deeper commonality driven by causation: it occurs when a shock to one price affects the dynamics of another (Listorti and Esposti, 2012). The economics of commodity price connectedness is ultimately what analysts, and the present study, care about (Diebold et al., 2017). To uncover these relationships, recent studies have developed various methodological approaches. Kozian et al. (2025, Appendix A) offer a comprehensive review and comparative evaluation of these methods.

Among the stylized facts emerging from this literature, three aspects are particularly worth emphasizing. The first concerns the complexity of the dynamics of these price series, and thus of their stochastic properties. Complex dynamics at the individual series level give rise to complex dependencies. As a result, identifying a common DGP across commodities that captures interdependence within these intricate dynamics often proves challenging.

The second stylized fact is that, despite the complexity of the underlying stochastic processes, both simple visual inspection and more sophisticated analyses confirm that commodity prices tend to move together, as repeatedly demonstrated during periods of rapid surges in commodity prices, such as in 2007 and 2021. This co-movement can involve closely related commodities (such as coal and crude oil) as well as seemingly unrelated ones (like wheat and gold) (Esposti, 2024a, 2024b).

As a consequence, the third stylized fact is that price interdependence may involve a wide range of commodities. Most commodities are part of many supply chains. These supply chains have become increasingly complex and global, and therefore increasingly interconnected. In addition, agent expectations and financial markets can link commodities that are only weakly connected in the real economy. As a result, beyond direct and simple linkages, the transmission of price shocks can also connect seemingly unrelated commodities. While direct connections are easily detectable and often widely studied (for instance, the corn–pork linkage; Quintino et al., 2021), indirect linkages are often barely visible, largely unknown, and difficult to define *ex ante*. Whether directly or indirectly, and with short or longer lags, shocks or fluctuations in the price of a single commodity can be transmitted to many others. This transmission can itself generate a form of amplification or reverberation of the original shock.

2.2. Methodological issues

The investigation of this complex connectedness requires an appropriate methodological framework. Although diverse in their specific techniques, the various approaches to this issue share a common underlying logic: interdependence is understood as the functional relationship linking the price of any i-th commodity within a group of N at time t to its own past values and to those of any other j-th commodity in the group, observed at previous times (t–s):

$$p_{i,t} = f\left(p_{i,t-s}, p_{j,t-s}\right) \ \forall i,j \in N. \tag{1}$$

A widely adopted explicit specification of equation (1) is the Vector Autoregressive (VAR) model, which describes the evolution of multiple time series by expressing each variable as a linear function of its own past values and those of other variables in the system. The linear structure makes VAR models functionally simple yet sufficiently powerful to capture complex dynamic interactions. They are particularly useful for studying causality, as they enable Granger causality tests to assess whether the past values of one variable help predict another (see below). However, investigating price interference through VAR modelling raises several challenges, which can be grouped into three thematic areas.

A first thematic area concerns the stochastic properties of the time series under consideration, with stationarity being the most prominent. VAR models require all variables to be stationary, or integrated of order 0 (I(0)), a condition that becomes increasingly difficult to meet as the number of price series (N) grows. Stationarity tests may yield inconclusive or conflicting results for individual series, depending on the

[1] The recommendations and suggestions provided by two anonymous reviewers on this literature review section are gratefully acknowledged.

specific testing procedures and specifications adopted (Enders, 2014). More critically, when N is large, establishing a common stochastic property across all series, namely, that they share the same order of integration, I(0), becomes exceedingly challenging. One potential solution for handling I(1), i.e. non-stationary, series is cointegration analysis, which leads to the specification of Vector Error Correction Models (VECMs). However, this approach requires that all series be unambiguously classified as I(1), and the number of cointegrating vectors may increase substantially with N, introducing further estimation complexities.

In general, as the number of price series increases, stationarity and integration testing become less reliable due to test size distortions, power loss, and cross-sectional dependencies. In such high-dimensional settings, it is often more pragmatic to abandon the goal of conclusively determining a common order of integration across all series. Instead, researchers may assume a common integration order, either I(0) or I(1), and proceed with modelling strategies accordingly (Stock and Watson, 2002; Lütkepohl, 2005).

The second thematic area emerging in recent literature concerns the presence of nonlinearities in commodity price dynamics and their interdependence. Conventional VAR models, being linear by design, tend to overlook persistent or transient nonlinear patterns in commodity price behaviour. Recent methodological advances highlight the importance of accounting for such nonlinearities and their potential sources, including speculative bubbles, volatility clustering, structural breaks, and regime shifts (Esposti and Listorti, 2013; Esposti, 2021, 2024a,b).

To address such limitations while preserving the VAR framework, several extensions to nonlinear or nonparametric forms have been proposed, including kernel-based VAR, Neural Network-based VAR (NN-VAR), and Gaussian process VAR (Signoretto and Suykens, 2015). Alternatively, more flexible approaches have emerged by moving beyond the VAR paradigm. In particular, Bayesian methods are increasingly employed to develop models capable of capturing complex dynamics more accurately (Boakye et al., 2024; Drachal and Pawłowski, 2024). Another promising avenue is the use of Machine Learning (ML) techniques to detect nonlinear connections among commodity prices (Kozian et al., 2025). Nonlinear correlations can also be explored through wavelet analysis (Boako et al., 2020; Nigatu and Adjemian, 2020; Kirikkaleli and Güngör, 2021; Mastroeni et al., 2022; Mutascu et al., 2022). In particular, Wavelet Cross-Correlation (WCC) enables the analysis of price interdependence across multiple time scales, allowing for the inclusion or exclusion of long-term trends and short-term cycles. However, the WCC's sensitivity to the filtering process may compromise the uniqueness of the resulting CPN matrix. Moreover, wavelet-based techniques may struggle to fully accommodate nonstationarity, potentially leading to spurious or misleading correlations.

The third, and more important here, thematic concerns high dimensionality. Ultimately, all solutions proposed in the recent literature and summarized above, tend to be computationally intensive and, in practice, become infeasible in high-dimensional settings, i.e., when the number of commodities is large and time dependence long. Recent literature in the field has emphasized that NA can offer a promising methodological approach. The logic of a network approach is to summarize price interdependence across a large set of N prices as a square (NxN) matrix also known as adjacency matrix:

$$\begin{bmatrix} f_{11}=0 & \cdots & f_{1N} \\ \vdots & \ddots & \vdots \\ f_{N1} & \cdots & f_{NN}=0 \end{bmatrix} \tag{2}$$

In equation (2), each element f_{ij} represents the presence and strength of a connection between the i-th and j-th prices (or network nodes), specifically the functional dependence of the i-th price on the past values of the j-th price. In unweighted networks, the matrix elements are binary: they take the value 1 if a connection (or edge) exists between two nodes, and 0 otherwise.

In weighted networks, the elements can assume any real value, reflecting the intensity or strength of the connection, depending on the nature of the relationship. Network connections can be either directional or bidirectional. We have an undirected network when $f_{ij}=f_{ji}$, and the adjacency matrix in equation (2) is symmetric. Conversely, we have a directed network when $f_{ij}\neq f_{ji}$, resulting in an asymmetric adjacency matrix. In this latter case, connections between nodes are also referred to as arcs.

By visualizing and quantifying the connections and interdependencies among commodity prices, NA enables the investigation of the system's structure or topology, the identification of "central" and "peripheral" commodities, as well as of clusters of commodities that tend to be more strongly interconnected (Schweitzer et al., 2009; Newman, 2010). The paper widely regarded as the origin of this field of study (Diebold and Yılmaz, 2014) employs a weighted dynamic network to examine time-varying connectedness. This is achieved through VAR modelling combined with variance decomposition. and subsequently investigate connectedness using NA.

Building on this framework, Granger Causality Networks (GCNs) have recently attracted attention (Sun et al., 2018) and can be considered a natural extension of the original approach. In this approach the relationship across commodity prices is typically captured by Granger Causality (GC) (Shi et al., 2018, 2020; Baum et al., 2023), an econometric method used to assess whether past values of one variable improve the prediction of another. GC thus helps uncover directional causal links, revealing which commodities exert influence on others over time.

GC is still based on VAR modelling[2] and estimation but does not require identifying assumptions on the long-term structural linkages and, consequently, does not distinguish between short and long-run causality (Dufour and Renault, 1998; Dufour and Taamouti, 2010). Just to mention a couple of recent applications of this GCN approach, Larrosa et al. (2024) investigate price leadership in the Argentinian retail tea market. Wang et al. (2021) propose a GCN in the time domain and frequency domain to investigate the interconnectedness of Chinese financial institutions. Carlos-Sandberg and Clack (2021) analyze the interdependence of oil prices across different regions and qualities, highlighting shifts in market influence, geopolitical events, or supply-demand dynamics.

The Granger causal relations can actually be described via two different network models (Eichler, 2012). The first consists of a network with N nodes whose arcs may change over time depending on the estimated coefficients of the corresponding lags of the VAR model (Carlos-Sandberg and Clack, 2021; Wang et al., 2021). Consequently, a VAR(N,K) model (with K indicating the number of lags) will generate K different NxN adjacency matrices representing the time-varying network, i.e., a *dynamic network*, similar to the dynamic Bayesian networks (Ghahramani, 1998). The second kind of network is a more compact representation, combining arcs from different lags of the VAR (N,K) model. This network is thus expressed by a NxN adjacency matrix that remains time-invariant. This *static network* indicates Granger causality whenever coefficients of some of the K lags point to a connection between two nodes. Here, we follow this latter idea as an unweighted static network model is more parsimonious and the consequent NA less computationally demanding, so it can better manage high dimensionality in the estimation stage (Ahelegbey et al., 2021).

It must be stressed that the construction of the price network (i.e., the adjacency matrix) through VAR model estimation does not, by itself, overcome the dimensionality issue arising from the large number of nodes. For example, the analysis by Diebold and Yılmaz (2014) is limited to only 16 nodes, raising serious concerns about the scalability of their

[2] The notion of Granger causality can be extended beyond the boundaries of the linear VAR framework, allowing for its application in nonlinear dynamical systems and more general statistical models (Shojaie and Fox, 2021).

approach to more complex networks. Moreover, VAR modelling tends to exacerbate this issue. In a VAR(N,K) model, the number of parameters to be estimated is Nx(KxN), since each of the N equations includes K lags for each of the N variables. Higher data frequency typically requires longer lags to capture indirect interdependence and seasonal or intra-year cycles. Consequently, when N is large and many lags are needed, the number of parameters becomes prohibitively large, making estimation unfeasible given the available observations T. According to Bernanke et al. (2005), standard VARs are typically limited to a small number of endogenous variables (usually less than ten) to preserve degrees of freedom. To address this high dimensionality, a GCN approach requires a more parsimonious VAR model. Two alternative empirical strategies are considered here. The first consists in performing pairwise VAR (PW-VAR) estimation followed by Granger causality (GC) tests. The second involves estimating sparse VAR (S-VAR) models, also followed by GC tests. Both approaches significantly reduce dimensionality, but differ in their statistical foundations and respective advantages and limitations.

Pairwise GC may be problematic for two major reasons. First of all, PW-VAR estimation ignores interactions among other variables while focusing on pairs, which can simplify the model fitting but may overlook multivariate dependencies. This limitation has been well documented (see, e.g. Lütkepohl, 1982; Shojaie and Fox, 2021) and may eventually imply inconsistent VAR estimates, and GC tests, due to omitted relevant variable bias. In spite of their limitations, however, bivariate tests of Granger causality continue to be widely used in many application areas, from economics (Chiou-Wei et al., 2008; Sun et al., 2018; Zhang and Broadstock, 2020; Larrosa et al., 2024) and finance (Hong et al., 2009) to neuroscience (Seth et al., 2015) and meteorology (Mosedale et al., 2006). A second issue with the PW-VAR GC approach is that it risks not being sufficiently parsimonious as it may generate many false positives (namely, spurious causal links are detected), especially when many lags are included. This occurs because, when many pairwise GC tests are performed in high-dimensional settings, each of these tests carries a certain probability of rejecting the null hypothesis of no causality by chance. As the number of tests increases,[3] the probability of at least one false positive approaches 1. The expected proportion of false positives is known as the False Discovery Rate (FDR) (Seth et al., 2015; Runge, 2018; Uematsu and Yamagata, 2025).

Given these limitations, Sparse VAR (S-VAR) modelling offers an effective alternative to PW-VAR when dealing with high-dimensional time series. By applying sparsity-inducing penalties, it reduces overfitting and simplifies the model structure. This approach jointly considers all variables, capturing complex interdependencies, while shrinking weaker connections to zero. The result is a sparse and interpretable Granger Causality Network (GCN), where only the most significant causal links are retained. Earlier contributions in this area, such as Litterman (1986) and Leeper et al. (1996), adopted a Bayesian framework with shrinkage priors to obtain stable estimates in moderately sized VAR models. Bayesian approaches continue to be considered a valid sparsity-inducing solution also in more recent works (George et al., 2008; Banbura et al., 2010; Ahelegbey et al., 2016, 2021).

Recent literature, however, has increasingly focused on directly selecting the nonzero entries by introducing sparsity-inducing penalties through Least Absolute Shrinkage and Selection Operator (LASSO) as the regularization technique (Lozano et al., 2009; Chudik and Pesaran, 2011; Basu and Michailidis, 2015).[4] More complex LASSO-based methods, such as adaptive LASSO, elastic net, and smoothly clipped absolute deviation (SCAD) (Breheny and Huang, 2011; Takada and Fujisawa, 2023), might enhance variable selection and sparsity, albeit at the cost of increased computational and methodological complexity. For this reason, these more sophisticated techniques have yet to be applied to the study of commodity price interdependence.

For the sake of comparison, this study employs both PW-VAR and S-VAR approaches to construct the Granger Causality Network (GCN) for a large set of commodity prices. While both methods address the issue of high dimensionality, the PW-VAR approach does not inherently guarantee sparsity in the resulting GCN. Greater sparsity can, however, be induced by adopting more stringent significance levels or more selective testing procedures for accepting Granger causality (see Section 3). PW-VAR models are generally more intuitive, interpretable, and straightforward to implement for detecting Granger causality, although they may suffer from estimation inconsistency due to the neglect of multivariate interactions.

In contrast, large VAR models with sparsity-inducing penalties (S-VAR) offer a more parsimonious, comprehensive, and theoretically grounded framework for high-dimensional settings, albeit at the cost of reduced interpretability. From an operational standpoint, the degree of sparsity achieved through S-VAR requires greater computational resources and careful tuning of penalty parameters. This introduces a degree of arbitrariness, as different regularization strategies may lead to different network structures, with the final choice often left to the analyst's discretion (see Section 3). Ultimately, the PW-VAR approach can still be regarded as a reliable approximation of a GCN derived from the more sophisticated S-VAR framework, with the sparsity and structure of the former ideally converging toward those of the latter.

Both VAR-based approaches to network identification assume linear relationships among prices. Nonetheless, to address the issue of nonlinearity discussed above, and following a widely adopted strategy in empirical research (Lütkepohl, 2005; Enders, 2014), a pragmatic approach can be adopted. Given the inherently linear nature of VAR, applying a logarithmic transformation to the original series may help capture more complex dynamics, such as multiplicative or power-law relationships.

A similarly pragmatic strategy is used to address another key issue: the potentially different stochastic properties of commodity price series. We consider both possible configurations, either most series are stationary (I(0)) or predominantly non-stationary (I(1)). In the first case, VAR models are estimated in levels; in the second, in first differences (Esposti, 2024b). Investigating price connectedness within GCNs using price levels captures both long-term and short-term interdependencies, but also carries the risk of identifying spurious causal relationships. In contrast, first differences emphasize short-term dynamics and help mitigate this risk.

By applying log-transformation to both price levels and their first differences the analysis can be extended further to quite different behavior of price dynamics. This allows to explore a broader range of price dynamics. For instance, a log-linear model in first differences implies that original price levels follow a multiplicative, path-dependent process. As a result, even if the differenced model is linear, the relationships among price levels remain nonlinear and cumulative (van Garderen, 2023). Repeating the GCN analysis across these different specifications and comparing the results helps identify patterns that are robust regardless of the underlying dynamics of the price series.

2.3. Policy relevance

At the heart of the extensive research effort devoted on commodity price interdependence lies a crucial policy concern: the volatility of resource and commodity prices represents a major source of economic instability, particularly in resource-dependent economies (Bredenkamp and Bersch, 2012). The United Nations Conference on Trade and Development (UNCTAD) 2023 report (UNCTAD, 2024) emphasizes that such volatility is a key source of socioeconomic stress in developing countries. Historically, high commodity price interdependence has triggered a range of policy responses focused on stabilizing economies,

[3] The number of variable pairs grows quadratically: for N variables, $N(N\text{-}1)$ tests are performed.

[4] For more details on LASSO estimation see Tibshirani (1996) and Yuan and Lin (2006).

managing inflation, and mitigating financial risks via a policy mix combining monetary, fiscal, and trade interventions, as well as exchange rate and macroprudential policies. The effectiveness and side effects of these responses vary depending on the context and the specific commodities involved (Gregorio, 2012).

A comprehensive review of policy tools to manage commodity price volatility lies beyond the scope of this study. However, one often overlooked insight deserves emphasis: the need for instruments that not only respond to volatility but also monitor and prevent it. International institutions consistently highlight the importance of data-informed policy frameworks. As reviewed by Zelingher (2024), several open-access platforms have been developed to this end. The FAO's Food Price Index tracks monthly changes in global food commodity prices and supports the Food Price Monitoring and Analysis (FPMA), which promotes market transparency, resilience-building, and early warning systems. The World Bank's Commodity Markets report recommends investing in infrastructure and data systems to enhance policy responsiveness (Baffes and Nagle, 2022). Likewise, the Global Food and Nutrition Security Dashboard offers real-time data on food price inflation and insecurity, enabling coordinated, evidence-based action. The *Commodity Markets Outlook* (World Bank, 2025) stresses how global shocks, such as the COVID-19 pandemic and the war in Ukraine, reshape commodity markets, urging governments to adopt flexible and forward-looking strategies. Other notable platforms include IFPRI's Food Security Portal (FSP) and the World Food Programme's Economic Explorer, part of the Vulnerability Analysis and Mapping (VAM) initiative. These systems incorporate nowcasting and early warning mechanisms to detect market volatility and support timely policy interventions.

Ultimately, policymakers now have access to a wide array of tools to monitor and respond to price shocks. Real-time surveillance systems, powered by advanced econometric models and AI-based forecasting tools, including neural networks, deep learning, and hybrid swarm-based systems, enable the detection of price bubbles and volatility clusters, thereby enhancing early warning capabilities (Ameur et al., 2024; Xavier et al., 2023). Alongside these quantitative approaches, qualitative tools such as participatory systems thinking engage stakeholders in mapping feedback loops and identifying leverage points for intervention. Periodically updated dashboards and data analytics visualize price surges, contagion across commodities, and links to inflation, guiding timely and targeted policy responses (Muflikh et al., 2021), whether sector-specific or economy-wide.

The network approach proposed in this study contributes to this evolving toolkit by offering a novel perspective. It maps connections across a wide range of commodities, not limited to agriculture, food, or energy, providing insights into underlying dynamics and identifying early movers. Continuously updating the network enhances real-time surveillance and supports timely, accurate policy responses. Designed to be interpretable even by non-technical users, this approach can translate results into a periodically updated dashboard that visualizes critical information. It enables the development of a real-time monitoring tool that can guide prompt policy action, distinguishing between sector-specific interventions and those requiring broader, economy-wide measures (Esposti, 2024a, 2024b).

3. The methodological approach

Our approach relies on Granger Causality Network (GCN) analysis to represent interdependencies among N commodity prices through a directional, unweighted adjacency matrix of size $N \times N$. The GCN framework offers two main advantages: (i) it provides a clear and intuitive mapping from Granger causality tests to binary network connections; and (ii) it remains computationally feasible even in high-dimensional settings, as testing is conducted within a suitable VAR framework that handles dimensionality and (non)stationarity more effectively than alternative methods.

The methodology unfolds in three stages. The first stage consists in the detection of causal links. We identify the elements of the adjacency matrix via Granger causality (GC) tests within VAR modeling. Two alternative specifications are considered: Pairwise VAR (PW-VAR) and Sparse VAR (S-VAR). The second stage concerns the construction of alternative networks. The two VAR specifications yield distinct versions of the CPN, each reflecting different assumptions and estimation strategies aimed at achieving sparsity. The third stage is the actual network analysis. We analyze the resulting networks using a set of indicators that describe the structure and intensity of interdependencies among individual prices and price groups.

3.1. Granger causality testing

Consider the N commodities whose price is observed over T time periods (months in the present case). As an explicit specification of (1) assume that the stochastic DGP representing the i-th price movement follows an autoregressive (AR) process:

$$p_{it} = \alpha_i + \delta_i t + \sum_{k=1}^{K} b_{ik} p_{it-k} \varepsilon_{it}, \forall i \in N; \forall t, k \in T; K < T \tag{3}$$

where p_{it} is the i-th commodity price at time t, α_i, δ_i, b_{ik} are commodity-specific unknown parameters to be estimated. α_i expresses the drift while δ_i the deterministic trend coefficient. Thus, α_i and δ_i indicate the long-term fundamental price level or the long-term deterministic trend, respectively, to which the actual price is expected to revert. b_{ik} are autocorrelation coefficients. ε_{it} is a disturbance term assumed to be normally, independently and identically distributed, $\varepsilon_{it} \sim NID(0,\ \sigma_i^2)$.

Under price interdependence, (3) is an incomplete representation of the underlying DGP. This latter has to include cross-price correlation terms as follows:

$$p_{i,t} = \alpha_i + \delta_i t + \sum_{k=1}^{K} \sum_{j=1}^{N} b_{ijk} p_{j,t-k} + \varepsilon_{it}, \forall i,j \in N; \forall t, k \in T;\ K < T \tag{4}$$

As (4) applies to all commodity prices, the actual stochastic process generating price series can be represented in a vector form, i.e., as a Vector AutoRegression process VAR(N,K):

$$\mathbf{p}_t = \mathbf{A} + \mathbf{D}t + \sum_{k=1}^{K} \mathbf{B}_k \mathbf{p}_{t-k} + \boldsymbol{\epsilon}_t \tag{5}$$

where $\mathbf{p}_t$ is the Nx1 vector of prices at time t and $\boldsymbol{\varepsilon}_t$ the Nx1 vector of the i.i.d. disturbance terms at time t. $\mathbf{A}$ is the Nx1 vector of drift coefficients, $\mathbf{D}$ is the Nx1 vector of deterministic trend coefficients, $\mathbf{B}_k$ is the NxN matrix of price correlation coefficients at generic lag k.

Provided that the series are stationary (i.e., I(0)), the VAR model in equation (5) can be consistently estimated in levels, as is standard practice (Enders, 2014). Alternatively, if the series are I(1), equation (5) can be specified in first differences. As previously noted, while preserving the linear structure of equation (5), it is also possible to capture potential nonlinear relationships among prices by estimating the model in a log-linear form. In practice, equation (5) can be estimated using different specifications of the price vector $\mathbf{p}_t$: levels of commodity prices, first differences, logarithms of price levels, and the first differences of the logarithms of prices. All these alternatives will be explored in the present study.

Once the model coefficients have been estimated, price p_j is said to Granger-cause price p_i if the past values of p_j have predictive power for the current value of p_i, conditional on the past values of p_i. Formally, the null hypotheses of no Granger causality from p_j to p_i involves testing whether all lag coefficients b_{ijk} are jointly equal to zero, $H_0 : b_{ij1} = b_{ij2} = \ldots = b_{ijK} = 0$. GC is assessed using a heteroskedasticity-consistent Wald test, whose test statistic asymptotically follows a chi-squared distribution with K degrees of freedom (Baum et al., 2023). It

is worth noticing that GC does not quantify the strength of the relationship between time series. Instead, it only indicates whether a causal link exists. Consequently, GC-based networks typically result in unweighted (binary) adjacency matrices.[5] To populate this $N \times N$ adjacency matrix (see Section 3.2), we assign $GC_{ij} = 1$ if the null hypothesis of no Granger causality from p_j to p_i is rejected at the chosen significance level. If accepted, $GC_{ij} = 0$ is assigned. Since the network is directional, whatever the value GC_{ij}, it can be either $GC_{ji} = 0$ or $GC_{ji} = 1$.

In practice, performing this battery of GC tests can be highly computationally demanding. First of all, in a VAR(N,K) model the number of parameters to be estimated (excluding drift and deterministic trend coefficients) is NxNxK (Morana, 2012). Second, **a** VAR(N,K) implies performing Nx(N-1) GC tests which entails estimating the model Nx(N-1)+1 times (1 unrestricted and Nx(N-1) restricted models). Third, for any equation of the model, NxK parameters must be estimated. Therefore, the number of available observations T must be sufficiently large to ensure identification and robust estimation, i.e., $T \gg N$xK. In the present case, where $N = 49$ and, at least, $K = 4$ (see Section 4) the total number of parameters to be estimated becomes 9604. The number of GC tests to be performed becomes 2352 which, in turn, implies estimating the model 2353 times. For each equation of the VAR model, 196 parameters must be identified and estimated with a number of observations $T = 540$.

To address this high-dimensionality problem in VAR models, one solution consists in estimating each pair of Granger causality measures, GC_{ij} and GC_{ji}, using pairwise VAR models, i.e. VAR(2, K) models:

$$\begin{aligned} p_{it} &= \alpha_i + \delta_i t + \sum_{k=1}^{K} b_{iik} p_{it-k} + \sum_{k=1}^{K} b_{ijk} p_{jt-k} + \varepsilon_{it} \\ p_{jt} &= \alpha_j + \delta_j t + \sum_{k=1}^{K} b_{jjk} p_{jt-k} + \sum_{k=1}^{K} b_{jik} p_{it-k} + \varepsilon_{jt} \end{aligned} \tag{6}$$

It means estimating Nx(N-1)/2 models each with 2K parameters and this seems much more feasible given the available T observations. By estimating bivariate models, this pair-wise VAR (PW-VAR) approach reduces complexity but may fail to capture the joint dynamics of the full system, potentially leading to omitted variable bias and spurious causal links arising from indirect effects. Furthermore, it is prone to generating false positives and, more generally, to overpopulating the adjacency matrix of the Granger Causality Network (GCN), resulting in a representation that lacks sufficient sparsity (see Section 3.2).

An alternative strategy to address high dimensionality is sparse VAR (S-VAR) modelling. S-VAR models are a variant of VAR models in which parsimonious restrictions are imposed on the model structure, meaning that many coefficients (and thus potential causal relationships) are set to zero. As a result, the number of parameters to be estimated is reduced, leading to a sparser coefficient matrix (Uematsu and Yamagata, 2025). S-VAR models aim to identify only the most relevant relationships among variables in a large multivariate dynamic system, enhancing interpretability and efficiency. Sparsity is achieved through variable selection techniques. In particular, during estimation, S-VAR models impose penalization or regularization constraints to shrink or eliminate insignificant coefficients, enabling full-system estimation while avoiding overfitting and preserving interpretability. One of the most widely used techniques, and the approach adopted in this study, is the LASSO regularization.

LASSO regularization can be described as follows. For each i-th price equation of the whole VAR(N,K) model, parameter estimation is performed according to a least square logic augmented by a penalization term. Considering generic equation (4), the following optimization problem is solved (Shojaie and Fox, 2021):

$$\min_{\{b_{ijk}\}} \left\{ \frac{1}{T} \sum_{t=K+1}^{T} \left(p_{i,t} - \alpha_i - \delta_i t - \sum_{k=1}^{K} \sum_{j=1}^{N} b_{ijk} p_{j,t-k} \right)^2 + \lambda \sum_{k=1}^{K} \sum_{j=1}^{N} |b_{ijk}| \right\} \tag{7}$$

where $\lambda > 0$ is the regularization (or penalization) scalar parameter that applies to the sum of the absolute values of all the NxK estimated coefficients (196 in the present case) of the individual i-th equation. In the context of LASSO, this sum is used as a penalty term in order to shrink many coefficients exactly to zero, and to automatically select the most relevant lagged variables for each equation. From this LASSO estimate it is thus possible to directly deduce whether price p_j GC price p_i (therefore, whether $GC_{ij} = 1$ or $GC_{ij} = 0$) if $b_{ijk} \neq 0$ for at least one k.

The lambda (λ) parameter is the key parameter in LASSO estimation, as it controls the amount of penalization applied to the model coefficients. A larger value of λ results in stronger penalization, reducing the number of variables selected in the model and inducing greater sparsity, whereas a smaller value leads to weaker penalization, thus limiting sparsity. λ is not estimated jointly with the model coefficients $\{b_{ijk}\}$ but is instead selected through an external procedure. Two alternative procedures are adopted in this study (see Section 3.2).

3.2. Building the Granger causality network

A GCN is an explicit form of (2) and is represented by an adjacency matrix like the following (**Sun et al., 2018**):

$$\mathbf{GCN} = \begin{bmatrix} GC_{11} = 0 & \cdots & GC_{1N} \\ \vdots & \ddots & \vdots \\ GC_{N1} & \cdots & GC_{NN} = 0 \end{bmatrix} \tag{8}$$

In this matrix rows represent the effects of causality while columns represent the sources of causality. In other words, the element in row i and column j indicates whether the price i is influenced (i.e., is Granger caused) by price j, namely that a fluctuation of node j can be transmitted to node i. If such a causation occurs than $GC_{ij} = 1$; otherwise, it is $GC_{ij} = 0$.[6] An additional key insight from this network structure is the role of indirect linkages, which, through direct GC connections, allow price shocks to propagate across all other commodities in the network. This characteristic is commonly referred to as *Network Topology*. Section 3.3 provides details how this topology can be investigated.

Given this common methodological framework based on GCN, the objective here is to find the most suitable way to define the network and have the best insight into commodity price interdependence. Five different adjacency matrices are considered and compared. Three are associated to PW-VAR models, two are obtained via S-VAR estimation. GCNs based on PW-VAR modelling differ for the underlying GC testing logic. The first (henceforth, GCN1) populates the adjacency matrix by rejecting the null hypothesis of GC at a 5 % confidence level. A second GCN (GCN2) adopts a more selective rejection criteria, namely 1 % confidence level as in Larrosa et al. (2024). Nonetheless, even this latter criterion might not guarantee enough sparsity of the matrix for the very large numbers of binary GC tests, Nx(N-1), to be performed to populate the matrix. When conducting such multiple hypothesis tests, the chance of obtaining false positives increases. To address this issue, the False Discovery Rate (FDR) offers a balance between discovery and reliability.

FDR is a statistical method used to correct for multiple comparisons when performing many hypothesis tests simultaneously. It controls the expected proportion of “false discoveries" (incorrectly rejected null hypotheses) among all the rejected hypotheses. The most common

[5] In **Zhou et al. (2022)** **a** weighted directed GCN is established using lag correlation between prices as weight.

[6] Dufour and Renault (1998) introduced an interpretation of GC distinguishing between short- and long-run causality. This concept is called h-step causality where one-step ($h = 1$) causality indicates a direct effects while indirect effects are expressed by h-step (with $h > 1$) causalities (see also **Uematsu and Yamagata, 2025**).

procedure for controlling FDR is the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995), which is followed here. First all p-values from the multiple tests are sorted in ascending order and ranked from 1 to Nx(N-1) (the total number of tests). For a desired FDR level α, representing the maximum acceptable proportion of false positives among the rejected hypotheses, a threshold value p_R is computed for any position in the rank, R: $p_R = (\alpha R)/[N(N-1)]$. Only null hypotheses with p-values less than or equal to p_R are rejected and enter the network with value 1. A third GCN (GCN3) is thus generated by applying this FDR procedure with $\alpha = 0.01$ in order to get a sparser network after PW-VAR estimation. It thus follows that moving from GCN1 to GCN3 via GCN2 we obtain a decreasing density, or increasing sparsity, of the network.

As anticipated, however, GCNs obtained via PW-VAR modelling may lead to inconsistent estimation of direct linkages and may miss relevant indirect linkages among prices. Inducing sparsity through a stricter rejection criterion does not necessarily resolve this issue. Therefore, GCN1-3 should be regarded as approximations of the true, but unknown underlying GCN. To better approximate this latent structure, the GCN is alternatively derived from S-VAR modelling using the LASSO estimation described in above. In this case, as well, two different GCNs are considered, depending on the procedure used to select the regularization parameter λ.

In one case (GCN4), λ is selected via cross-validation (CV), aiming to balance the trade-off between prediction accuracy and model sparsity. A grid of candidate λ values is considered. For each candidate, the model is trained on a subset of the data and validated on the remaining portion. The value of λ that minimizes the prediction error, measured by the mean squared error (MSE), is eventually selected. Alternatively, λ is selected using a conventional information criterion. Here, the Bayesian Information Criterion (BIC) is adopted. As BIC tends to select a larger λ, it imposes a stronger penalization on model complexity. Consequently, the GCN obtained using BIC (GNC5) is expected to be sparser than the one generated via CV (GCN4).

The comparison across this sequence of networks, from GCN1 to GCN5, is performed using the different time series. Commodity price levels are considered first. Then, to account for potential non-stationarity and nonlinearity, the same sequence of networks is generated using the first differences of prices, the logarithm of price levels and the first differences of the logarithm of prices.

3.3. Performing network analysis: the economics of connectedness

The five GCNs are investigated and compared via the same battery of network indicators (Diebold et al., 2017). In order to comprehensively describe the structure and functioning of the network, these indicators are here divided in four different dimensions: network topology; network correlation; leader nodes; clusters and communities.

3.3.1. Network topology

Network topology refers to the physical or logical arrangement of nodes (namely, commodities) within the network. It defines how nodes are interconnected and how shocks and fluctuations flow among them. Graphical visualization often represents the easiest way to express network topology: nodes and arcs are plotted as a graph with alternative layout algorithms determining the respective positioning. Nonetheless, besides visualization, topology can be more formally investigated via numerical indicators expressing the general structure of the network.

A first general indicator of the network topology concerns its *density* (D). Usually, density is expressed as the ratio of the number of connections (or arcs) in the network to the number of possible arcs. So, D indicates how close the network is to the maximum possible density.

As all GCNs here considered are directed networks, this indicator can be computed as $D = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} GC_{ij}}{N(N-1)}$, where GC_{ij} indicates the generic element of the adjacency matrix representing the network. The main implication of this indicator, from a more economic perspective, is that the lower the D the sparser the network which eventually implies that the network shows, in relative terms, less direct and more indirect linkages.

A second set of indicators that synthetically captures the incidence of direct versus indirect connections is based on distance analysis. Within a network, distance is usually expressed by the shortest path between two nodes in the network (d_{ij}). d_{ij} expresses the number or arcs of the path that connects the two nodes with the fewest number of arcs. Consequently, a first distance indicator expressing the overall network topology is the average of the shortest paths (or average path length): $APL = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} d_{ij}}{N(N-1)}$ with $i \neq j$. A second distance indicator is the longest shortest path: $LPL = \max d_{ij}$ with $i \neq j$. This latter indicator is also called the "diameter" of the network as it gives an indication of how "far apart" the more peripheral network's nodes are.

Another aspect of network topology, also known as *granularity*, concerns the presence of micro-level structures. An easy way to investigate this aspect is the census of dyads and triads. A dyad consists of two nodes and the connection between them. It is the simplest relationship in a network and represents a direct link between two individuals or entities. A triad consists of three nodes and the connections among them. [7] The frequency of dyads (i.e., the proportion of mutual connections) within the network measures how often pairs of nodes have reciprocal ties and, therefore, it also known as *reciprocity*. This feature can also be expressed in relative terms, that is, as a ratio between the observed reciprocity and the expected reciprocity in a random network of the same size and density. The frequency of closed triads is also often used as an indicator of the tendency of the network to form small groups or clusters. It is called *transitivity*. In the context of network analysis, transitivity expresses the likelihood that the adjacent nodes of a node are connected and it is measured as the ratio of the number of closed triplets (or triangles) to the number of connected triplets of nodes.

A conceptual shift from network-level to node-level analysis is represented by *peripherality*. Investigating peripherality concerns the emergence of core-periphery structures in the Granger Causality Networks (GCNs). This investigation within a network is usually performed by looking at the number and frequency of global and local bridges. A bridge is a link between two nodes that, if removed, prevents one from reaching the other. It is therefore essential for maintaining connectivity between them. A link is considered a local bridge if its endpoints share no common neighbours. In contrast, a global bridge is a link whose removal splits the network into two disconnected parts. Global bridges are crucial for preserving the overall connectivity of the network. [8]

3.3.2. Network correlation

The topology of a network can also be characterized relatively by quantifying its structural differences with respect to reference networks, namely networks with the same number of nodes but of known topology and structural properties (synthetic networks). In particular, five synthetic models are commonly considered as extremes along the structural spectrum (see also Section 5.3.1): geodesic, lattice, ring, random, and small-world networks (Watts, 2004).

The geodesic network exhibits high density and regularity, minimal local clustering and moderate global clustering due to its structured connectivity. The lattice network, typical of physical infrastructure, has a highly regular, grid-like structure, with nodes connected to their immediate neighbours in a fixed, repeating pattern. This results in high local clustering but low global clustering. In a ring network, each node is

[7] Triads can be open or closed. A closed triad is when all three nodes are connected to each other, forming a triangle. An open triad is when only two of the three nodes are connected to each other.

[8] Local bridges are always more numerous than global ones since only a few local bridges are also global.

connected to two neighbours, forming a closed loop. This leads to very low clustering, as only immediate neighbours are connected, and to a lack of core structures. In a random network, connections among nodes are created randomly, usually resulting in low clustering at both the local and global levels. A small-world network starts from a lattice or ring structure and then reconnects some links randomly. This produces mixed properties: most connections remain local, but some long-range links are introduced, enhancing overall connectivity.

Correlation analysis can also be useful to compare the five GCNs considered in this study, as well as to assess the similarity between networks derived from different data transformations, such as price levels, first differences, and logarithmic transformations. In general, GCN1–3 are expected to be highly correlated, as are GCN4–5, since they are generated using the same estimation strategies (PW-VAR and S-VAR, respectively), with the only difference being the degree of sparsity achieved. Conversely, the correlation between these two groups of GCNs can provide insight into the extent to which the estimation strategy influences the identification of connectedness among commodity prices.

Correlation between different price-based GCNs may also be informative. GCNs computed on price levels are expected to be highly correlated with those based on first differences, especially when long-term linkages, which are not captured in the differenced series, are either negligible or simply mirror short-term dynamics. In such cases, GCNs based on first differences may appear sparser, but not qualitatively different from those based on levels. Similarly, if GCNs constructed from log-transformed prices show high correlation with their linear counterparts, this would suggest that the assumption of linearity in price connectedness does not substantially alter the network structure.

In any case, to measure the correlation between two networks, the Quadratic Assignment Procedure (QAP) is adopted (Borgatti et al., 2018). Each adjacency matrix must first be flattened into a vector (row-wise or column-wise). A Pearson correlation coefficient is then computed between any pairs of resulting vectors. This approach is analogous to standard statistical correlation, but applied to the structural patterns of the networks. Similarly to a standard Pearson correlation test, it is also possible to assess the statistical significance of the observed correlation. After computing the correlation between two adjacency matrices, one of the matrices is randomly permuted (typically by shuffling rows, columns, or arcs) multiple times. For each permutation, the correlation is recalculated. The p-value is then estimated as the proportion of permutations in which the correlation is as extreme as, or more extreme than, the observed one.

3.3.3. Leader nodes

Identifying key nodes, or leader nodes, within a network is crucial for understanding the network's dynamics. To investigate this aspect, a sequence of centrality measures can be computed. They help to identify the most important nodes in a network based on their connections and positions according to different perspectives (Sun et al., 2018; Zhou et al., 2022; Marra et al., 2024): degree centrality emphasizes local connectivity; betweenness centrality and closeness centrality concerns control and efficiency of shocks' transmission within the network, respectively; eigenvector centrality focuses on influence based on connections.

Degree centrality (DC_i) is the simplest measure of node importance, indicating the number of direct connections (arcs) a node has. A node with high degree centrality has many connections and is often considered a key player in the network. In directed networks, DC_i is distinguished in out-degree and in-degree centrality (ODC_i and IDC_i respectively). Here, ODC_i measures how many prices are Granger caused by the i-th price, while IDC_i represents how many prices Granger causes the i-th price. The usual interpretation, therefore, is that ODC_i captures the relevance or influence of the i-th price within the network, that is, its transmission range. IDC_i, on the contrary, captures the dependence of the i-th price within the network. It is also designated as *price sensitivity* as it measures how much a given price is affected by price fluctuations of other prices.

Betweenness centrality (BC_i) is a measure that quantifies how often a node acts as a bridge along the shortest path between two other nodes. It indicates the influence a node has over the flow of information in the network: $BC_i = \sum_{j \neq h \neq i}^{N} \left(\sigma_{jh}^{i} / \sigma_{jh} \right)$, where σ_{jh} indicates the total number of shortest paths from generic node j to generic node h, and σ_{jh}^{i} the number of those paths that pass through the node of interest i. BC_i is used to measure the intermediation capacity and identifies transmission hubs. In our price transmission network, the stronger the betweenness centrality of price, the more it acts as the intermediary of causality among other prices. Nodes with high BC_i control information flow and act as bridges connecting different parts of the network. A node that has both high DC_i and high BC_i is likely to be a key node both in terms of local connections and network-wide influence.

Closeness centrality (CC_i) measures how close a node is to all other nodes in the network. It reflects the ability of a node to quickly interact with all others. It measures the average length of the shortest path from each node to the others: $CC_i = 1 / \sum_{j \neq i}^{N} d_{ij}$. In the present study, CC_i represents the transmission speed of price fluctuations. The greater the closeness centrality of a price, the shorter the transmission path between that price and other prices. This means that fluctuations in that price can be transmitted more quickly across the network. Nodes with high closeness centrality are typically well-positioned, often referred to as *influencers*, because they can efficiently disseminate shocks throughout the network

Finally, eigenvector centrality (EC_i) measures a node's influence based on the idea that connections to high-scoring nodes contribute more to a node's relevance than equal connections to low-scoring ones. If a node is connected to other high-centrality nodes, it has a higher eigenvector centrality. It is a sort of "second-level centrality" because it takes into account not just the number of connections (like degree centrality), but also the quality of those connections. For a given node i, its eigenvector centrality is given by the i-th element of the principal eigenvector, i.e., the eigenvector corresponding to the largest eigenvalue, λ_{max}, of the following eigenvalue equation: $\mathbf{GCN} \cdot \boldsymbol{x} = \lambda \boldsymbol{x}$, where **GCN** is the NxN adjacency matrix, $\boldsymbol{x}$ is an Nx1 eigenvector of **GCN** and λ is the respective eigenvalue.

3.3.4. Clusters and communities

In previous sections, we illustrated indicators suitable for investigating network properties from two opposite perspectives: the network as a whole and individual nodes. An intermediate level worth considering is that of clusters or groups of nodes. In this regard, the analysis can be approached from two different angles.

The first involves using indicators that detect the network's tendency to form local clusters, that is, the propensity to concentrate arcs in specific areas of the network, as opposed to a homogeneous distribution of arcs throughout the entire structure. The key concept and corresponding indicator in this context is the *clustering coefficient* (*CL*), which quantifies the extent to which nodes in a network tend to form tightly knit groups. It provides insight into the local cohesiveness of the network. In the present analysis, *CL* is calculated as the proportion of closed triads relative to all possible triads in the network. The main focus of the analysis might not be on the clustering tendency of the network per se, but rather on its implications for systemic risk. Highly segregated clusters tend to localize the transmission of risk within themselves, thereby limiting the spread of shocks across the entire network. In other words, the presence of well-defined clusters may act as a buffer, containing systemic disturbances within specific segments of the network. In terms of the systemic risk associated to a network, it is possible to use

the eigenvalue equation not just to measure individual node influence, but also to understand the system-wide capacity for shock transmission.[9] The maximum eigenvalue $\lambda_{\max}$ (or spectral radius) of the adjacency matrix reflects the network's capacity to propagate and amplify shocks. In dynamic systems, if $\lambda_{max}>1$, a small initial shock can grow over time thus representing an unstable system and indicating a higher systemic risk (Acemoglu et al., 2012).

The second analytical perspective on the role of groups of nodes within the network concerns the degree to which a network can be partitioned into distinct *communities*, namely groups larger than small local clusters, characterized by dense intra-community connections and sparse inter-community links. This network feature is often designated as *modularity*. Here, this indicator is computed as the ratio between intra-group and inter-group density, based on either incoming or outgoing connections. For instance, a value of 1.5 indicates that for every connection going outside the group, there are 1.5 connections remaining within the group. In the present study, communities of interest are represented by pre-determined commodity groups (see Section 4 for details). Accordingly, we do not use network analysis to identify emerging communities; rather, we aim to investigate whether, and how, the properties and performance of these ex-ante communities differ, as well as the nature of the connections within and between them.

4. Price series under scrutiny

As discussed above, the main novelty of the present contribution stems from the high-dimensional nature of the dataset. We consider the price of 49 commodities, which are grouped into five main categories (EMAFO): Energy (3 commodities), Metals (12), Agriculture (13), Food (12), and Other raw materials (9). All price series are monthly and span the period from January 1980 (1980M1) to December 2024 (2024M12), resulting in 540 time observations. Therefore, the dataset has dimensions N = 49 (commodities) and T = 540 (time periods), for a total of 26460 observations. All series are sourced from the International Monetary Fund (IMF) commodity price dataset. Table A1 (Annex 1) provides further details on the data sources, particularly regarding the specific product qualities represented and the markets from which the prices are collected.[10]

The International Monetary Fund (IMF) commodity price dataset appears particularly well-suited for the present analysis, both in terms of frequency and coverage. Its monthly frequency offers a valuable compromise between the number of time observations and the incidence of medium-to long-term linkages, which are the primary focus of this study. Alternative data sources may present limitations either in the time dimension or in commodity coverage by encompassing commodities of diverse nature. For instance, the FAO monthly Food Price Index, while useful, lacks the breadth of the IMF dataset as it focuses solely on agricultural and food products. On the other hand, high-frequency datasets, such as those provided by OpenDataBay,[11] Papers With Backtest[12] or Chicago Board of Trade (CBOT) Data Portal,[13] offer daily or weekly prices, which tend to overemphasize short-term dynamics. These may obscure the structural relationships of interest here.

In addition, such high-frequency datasets typically include futures prices as well, whereas this study focuses exclusively on spot prices, aiming to capture real economic linkages among commodities, independent of financial speculation. The focus on these real linkages also motivates why, following Esposti (2024a,b), here commodity prices are neither deflated nor adjusted for the possible presence of seasonality, particularly in the case of agricultural prices. The logic behind this choice is that we prefer to analyze the price series that economic agents really confront and on which they take decisions without possibly introducing artificial transformations.

Another key advantage of the IMF dataset is completeness. With few exceptions, monthly prices are consistently available over time, with minimal missing values and a standardized definition of commodity grades. Exceptions include Natural Gas, Cobalt, and Potash. For Natural Gas, the series begins in January 1985; earlier data were reconstructed by backward interpolation using U.S. annual prices from 1980 to 1985, as provided by the Energy Information Administration (EIA). For Cobalt and Potash, changes in grade definitions and regional benchmarks occurred in the early 1990s and early 2020s, respectively. Homogeneity was restored by calculating the ratio between overlapping price periods and applying it retrospectively.

The relevance of the selected commodities is underscored by the fact that six of them are included among the 34 Critical Raw Materials (CRMs), and four are listed among the 17 Strategic Raw Materials (SRMs), as defined by the European Commission (2024). It is important to note that all SRMs are also CRMs, but not all CRMs qualify as SRMs. For instance, aluminium is classified as both a CRM and an SRM. More broadly, many of the commodities considered here are crucial for key industries, defence, infrastructure, and overall economic stability.

Some limitations of the IMF dataset still merit attention. First, as the dataset aims to reflect global price dynamics, the selected markets are typically major international benchmarks. However, certain commodities, particularly agricultural ones, may exhibit more regional price behavior, potentially influenced by local macroeconomic shocks such as differential inflation rates. In these cases, although all prices are denominated in U.S. dollars, exchange rate fluctuations may still introduce short-term noise into relative price dynamics.

Another limitation concerns the absence of certain critical commodities, due either to delayed or incomplete data collection (e.g., chromium and manganese), or to their recent emergence as relevant market assets (e.g., lithium, silicon, and rare earth elements), particularly in the context of digital and renewable energy technologies. The exclusion of these commodities due to insufficient data may limit the generalizability of the findings, especially with respect to commodities traded in less liquid markets or originating from emerging economies, which typically exhibit lower trading volumes, higher price volatility, and reduced market depth.

Finally, for certain commodities, especially agricultural products and minerals, the adopted dataset may neglect relevant qualitative differences exist across grades or types (e.g., various purity levels for nickel and copper, or multiple quality categories for olive oil). This study present adopts the standard quality grade used in the IMF dataset, typically the most representative and relevant. Nonetheless, future research could explore how quality differentiation across commodities may influence price linkages.

The commodity grouping indicated above (EMAFO) may be somewhat arbitrary.[14] For instance, food commodities are distinguished from agricultural commodities because the former require a degree of industrial processing starting from agricultural raw materials. While this criterion may clearly separate wheat from olive oil, it is not so obvious in the case of beef and coffee. Moreover, several other sub-groups of commodities could also be proposed. The "agriculture" group could be subdivided into "colonial" (Banana, Coffee, Cocoa, Tea) and "non-

[9] Meng et al. (2014), for instance, propose using the sum of the eigenvalues of the GCN matrix, divided by the number of nodes (*N*), as a measure of systemic risk.

[10] These price series are proprietary and can not be made available within the paper's material. However, they can be freely downloaded at https://data.imf.org/?sk=471DDDF8-D8A7-499A-81BA-5B332C01F8B9 or requested at https://www.imf.org/en/Research/commodity-prices.

[11] https://www.opendatabay.com/data/financial.

[12] https://paperswithbacktest.com/datasets.

[13] https://www.cmegroup.com/company/cbot.html.

[14] Zhang and Broadstock (2020), for instance, consider a broad range of commodities aggregated into seven price indexes (Metal, Food, Precious, Oil, Raw, Beverage, Fertilizer).

colonial" (Barley, Beef, Corn, Lamb, Pork, Poultry, Rice, Sorghum, Soybean, Wheat) products. Moreover, colonial products can also be found in the 'Food' group, as is the case with coffee and tea. Besides the geographical provenience, "agriculture" and "food" groups themselves could be crossed and reassembled into "animal products" and "crops" or on the basis of prevalent use, like "cereals", "edible oils", "feed" just to give some examples. The group of "other raw materials" is characterised by having neither energy nor food use, but could be further disarticulated depending on the use itself like, for instance, "fertilizers", "fibres", "forest or wood products".

Nonetheless, while any ex-ante grouping could be questionable, the one here proposed can simply help to better interpret the results and assess the advantages of the proposed approach. At the same time, results emerging from NA can provide ex-post empirical support on whether the actual commodity prices clustering tends to be consistent with the adopted groups. Replications on other grouping logic can represent an interesting extension of the present study in future research.

5. Results

The results of the analysis are presented in accordance with the sequence of steps defined by the methodological framework outlined in Section 3. Particular emphasis is placed on key statistical findings (highlighted in bold within the result tables), while secondary results are discussed subsequently or included in the annexes. All tests, estimations, and calculations were performed using STATA version 19.5.

5.1. Stochastic properties of the price series

Granger causality is assessed within Vector Autoregression (VAR) models of price formation. As discussed in Section 2.2, this approach relies on the key assumption of stationarity, namely that all time series are I(0). Before constructing the Granger Causality Networks (GCNs), unit root testing is therefore necessary to verify this condition. The Augmented Dickey-Fuller (ADF) test is here employed to assess the presence of unit roots. This test, widely used in the literature (Sun et al., 2018; Zhou et al., 2022), evaluates the null hypothesis that a time series possesses a unit root, i.e., it is non-stationary (I(1)). Rejection of the null implies stationarity and thus suitability for inclusion in the VAR framework.

The ADF test is applied to the 49 price series, and its specification is selected for each series using a sequential general-to-specific approach (Enders, 2014). This consists of a stepwise testing strategy that begins with the most general specification (including both a trend and a drift, as in (3)) and simplifies it step by step by removing unnecessary deterministic components (trend and drift) based on their statistical significance. For any test (i.e., commodity), the selection of the lag order is made using the Bayesian Information Criterion (BIC). This criterion helps to balance the trade-off between model fit and parsimony, both avoiding overfitting and enhancing the power of the test.

Table A2 (Annex 2) summarizes the unit root test results. It reports only the p-values, while the full test outputs are available upon request. P-values greater than 0.1 are shown in bold, indicating cases in which we reasonably accept the null hypothesis of a unit root, that is, the series is I(1). In all other cases, we reject the null hypothesis and conclude that the series is I(0), possibly around a drift and/or a trend. In order to assess the robustness of these stationarity properties and, if necessary, identify the appropriate transformation to achieve stationarity, ADF tests are repeated on the four different sets of series: price levels, first differences of prices, logarithms of price levels, and first differences of the logarithm of prices.

The key finding is that, with few exceptions, results are largely correspondent across the 49 commodities: all price series can be considered I(0) as required by the VAR specification. They behave like mean-reverting processes, possibly around a drift and/or a deterministic trend eventually determined by the respective long-term market fundamentals (Esposti, 2024a). The few exceptions to this consistent I(0) evidence (namely, cases showing p-values above 0.1) are found in both price levels and price logarithms. These involve 7 commodities (about 14 % of the total set), concentrated among metals (Copper, Gold, Silver, Tin) and colonial agricultural products (Bananas and Cocoa beans). Beef is the only other case. For these commodities, the null hypothesis of a unit root should be accepted. These exceptions disappear when the first differences (or the logarithm of the first differences) are considered: all unquestionably behave as I(0) series.

Overall, conclusive evidence in favour of stationarity emerges from the ADF unit root tests: over 85 % of the price series are stationary (I(0)) at level, and 100 % become stationary after first differencing. Nonetheless, to assess whether, and to what extent, the presence of non-stationarity in a few series may affect the network analysis, the latter is henceforth conducted on both sets of data: price levels and first differences. As expected (Esposti, 2024a, 2024b), the logarithmic transformation does not alter the stationarity results. Therefore, log-transformed series are considered only in Section 5.4, where they are used to assess the robustness of the network analysis in light of potentially nonlinear relationships among commodity prices.

5.2. Granger causality

The CPN is constructed here as a GCN, that is, based on Granger Causality (GC) across prices. GC testing is conducted using five alternative approaches, each corresponding to a different VAR modelling strategy. From pairwise VAR (PW-VAR) estimation, with GC accepted at the 5 % confidence level, we obtain GCN1. Applying a stricter 1 % confidence level yields GCN2. When the 1 % level is further adjusted for false discovery rate (FDR), we obtain GCN3. GCN4 is derived from LASSO estimation of the sparse VAR (S-VAR) model, with the regularization parameter (λ) selected via cross-validation (CV). Finally, GCN5 is obtained by selecting λ based on the Bayesian Information Criterion (BIC).

Due to the high dimensionality of the dataset and space limitations, the VAR estimates are not reported here but are available upon request. In the following sections, we report and discuss only the GCNs obtained with these estimates.

5.2.1. Pairwise VAR models

To generate GCN1–GCN3, we separately estimate a battery of 1176 binary models, i.e. VAR(2, K). For each model, the inclusion of a drift and/or trend is determined based on the specification adopted in the ADF tests: if either of the two price series includes a drift and/or trend in its ADF specification, the same component is included in the corresponding VAR model.

Regarding the lag order K, it is worth recalling that indirect and complex linkages across prices may only become apparent after a sufficient number of periods. Moreover, for certain commodities, particularly agricultural products, seasonal effects may play a significant role. When using monthly data, considering a longer lag length (e.g., 12 months, $K = 12$) to account for these effects would, in the present case ($N = 49$), require estimating 564 parameters per equation in the VAR model, resulting in a total of 26508 parameters for the entire system. Therefore, from the perspective of model parsimony, the use of monthly data calls for an empirical strategy that appropriately balances the potential for longer-term effects with the need to limit model complexity. For this reason, the lag length is selected for each equation using the Bayesian Information Criterion (BIC), subject to the constraint that K ranges from a minimum of 4 to a maximum of 12.

Some key findings about the network topology are revealed by Figs. 1 and 2. The dashed arcs indicate inter-commodity connections, which are more frequent in denser networks and tend to vanish in sparser configurations. Figures display the network topology for all PW-VAR GCN variants. As expected, the transition from GCN1 to GCN3

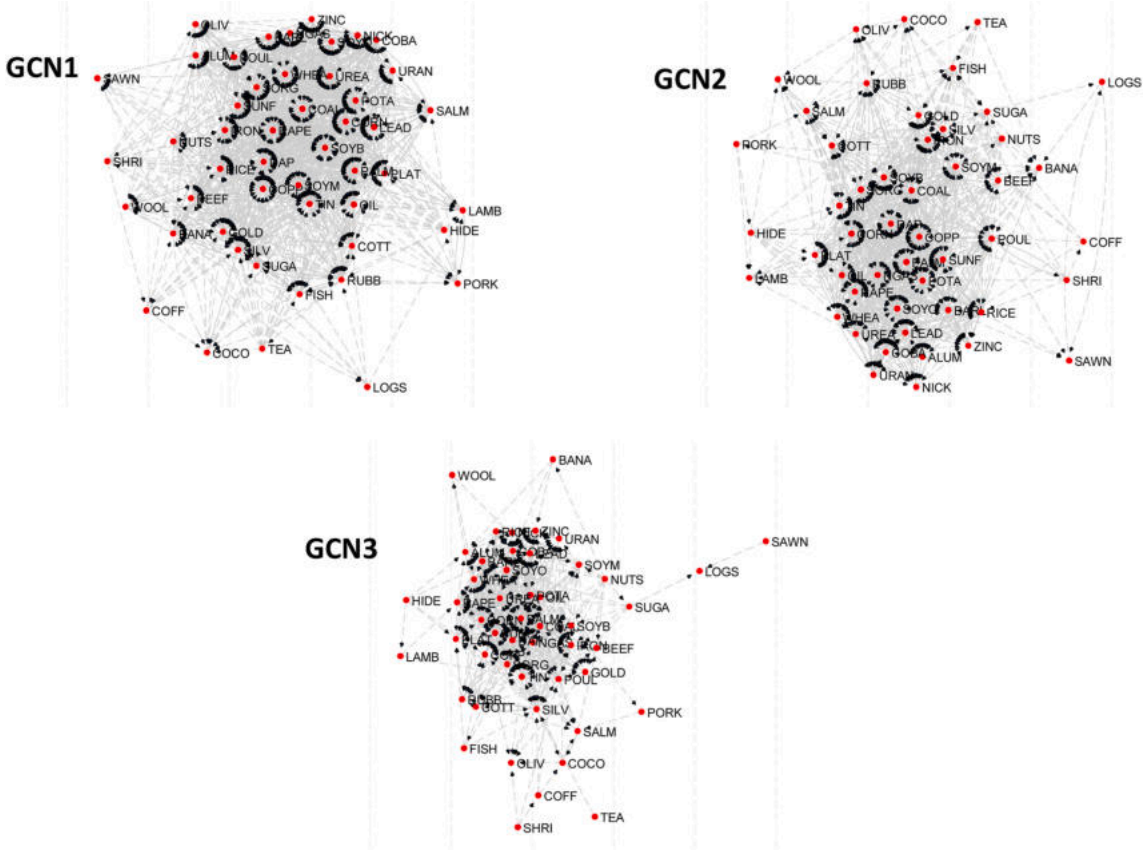

Fig. 1. Network structure GCN1-GCN3 – price levels (dashed arcs indicate links across commodities and arrows indicate the direction).

reveals a clear reduction in network density: the number of arcs decreases, and peripheral nodes become more isolated. This sparsification is especially evident in the networks based on first differences (Fig. 2).

Eventually, among the networks obtained through PW-VAR modelling, GCN1 based on price levels emerges as the densest case, while GCN3 based on the first differences of prices is the sparsest. This is made explicit in Annex 3 where the number of arcs is reported for each node (i.e., commodity) and for the whole network. This number reaches its maximum in GCN1 based on price levels, with 1049 arcs, corresponding to an average of about 21 linkages per node. Conversely, it is at its minimum in GCN3 based on the first differences of prices, with only 172 arcs and an average of 3.5 linkages per node, making it nearly seven times sparser than the former. Between these two extremes, we find, for instance, GCN3 based on price levels, with 496 arcs and an average of about 10 linkages per node, and GCN1 based on the first differences of prices, with 472 arcs and 9.6 linkages per node, respectively.

In all cases, the network topology reveals some core-periphery structure, with a group of central commodities strongly connected among themselves, and with all others and another body of commodities being relatively peripheral if not entirely isolated. This core-periphery structure becomes more evident moving from GCN1 to GCN3 and from networks based on price levels to cases based on the first differences of prices.

5.2.2. *Sparse VAR models*

To generate GCN4 and GCN5, we estimate a single VAR(49, K) model. Due to the high dimensionality of the system, the specification of deterministic components (drift and trend) and the selection of the lag order K cannot be based on an optimizing procedure. Therefore, both a drift and a trend are systematically included, allowing the penalized LASSO procedure to determine their relevance. The lag order is fixed at $K = 4$, as this is the most frequently selected specification in the PW-VAR estimations.

Figs. 3 and 4 display the network topology for all S-VAR GCN variants and highlight the structural differences between PW-VAR and S-VAR models. The key finding is that the extreme values in terms of sparsity are found with GCN4 based on price levels, with arcs 1034 and an average of about 21 arcs per node, and with GCN5 based on the first differences of prices, with 201 arcs and an average of 4.1 arcs per node (Annex 3). GCN4 thus appears more cohesive, with tighter clusters and more reciprocal links, while GCN5 shows a more hierarchical and directional topology, with fewer mutual arcs and more pronounced asymmetries. In GCN5 (Fig. 4), directional flows concentrate around a few central nodes, suggesting the emergence of systemic commodities that act as price transmitters.

Another key finding is that PW-VAR modelling allows to obtain at least the same range for sparsity of the S-VAR approach. Nonetheless, they may lead to a network with a different nature and structure. In order to better assess whether and how these GCN variants structurally differ, beyond sparsity, Table 1 reports the Pearson correlations (ρ) of the adjacency matrices corresponding to the different network versions (see Section 3.3.4 for details). The key evidence is that correlations are positive across all GCNs, but not all correlations are statistically significant at the 5 % or 10 % level. On the one hand, a strong correlation emerges among GCN1, GCN2, and GCN3, both in terms of price levels and first differences. The strongest correlations are observed within the PW-VAR models based on price levels, particularly between GCN1 and GCN2 ($\rho = 0.764$) and between GCN2 and GCN3 ($\rho = 0.754$), indicating that tightening the significance threshold reduces density but preserves network structure. GCN4 and GCN5 (S-VAR) show less pronounced yet statistically significant correlation ($\rho = 0.417$).

The interpretation is that switching from CV to BIC penalization sparsity is affected, but the overall structure tends to be preserved. This holds for both price levels and first differences, although the effect

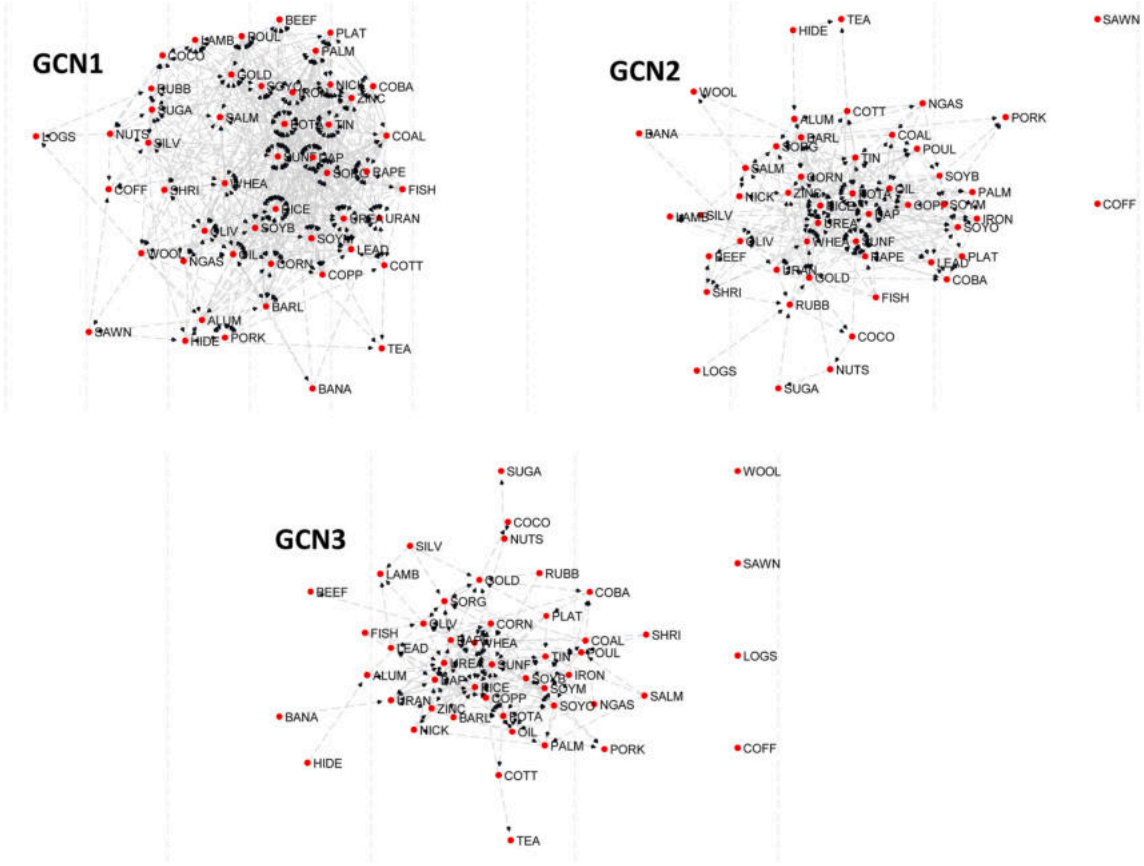

Fig. 2. Network structure GCN1-GCN3 – price first differences (dashed arcs indicate links across commodities and arrows indicate the direction).

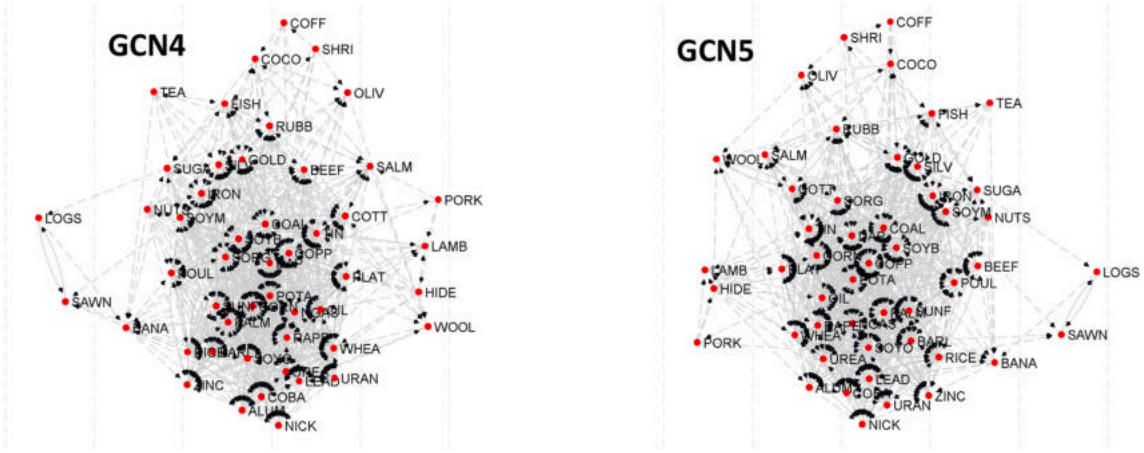

Fig. 3. Network structure GCN4-GCN5 – price levels (dashed arcs indicate links across commodities and arrows indicate the direction).

appears more pronounced in the latter case. On the other hand, what truly affects the network structure is the underlying VAR modelling approach, as PW-VAR and S-VAR GCNs rarely exhibit statistically significant correlations, except for a weak correlation between GCN3 and GCN5 in the case of price levels.

Cross-model correlations between networks based on price levels and first differences are generally low ($\rho < 0.3$), confirming that differencing alters network topology more substantially than model choice. However, particularly interesting is the significant correlation between GCN1–3 for price levels and GCN4–5 for the first differences of prices. Transitioning from price levels to first differences appears to substantially alter the network structure, but when this transformation is combined with a shift from PW-VAR to S-VAR modelling, the network structure seems to be substantially preserved, as suggested by the presence of significant, albeit mild, correlations in this case.

5.3. Network analysis

5.3.1. Network-wide metrics

The indicators illustrated in Section 3.3.1 provide a summary of network characteristics and highlight structural differences across GCNs. Three general aspects of the network topology, in particular, deserve attention and appropriate indicators: density, granularity, and peripherality. Table 2 reports these network-wide metrics. A geodesic

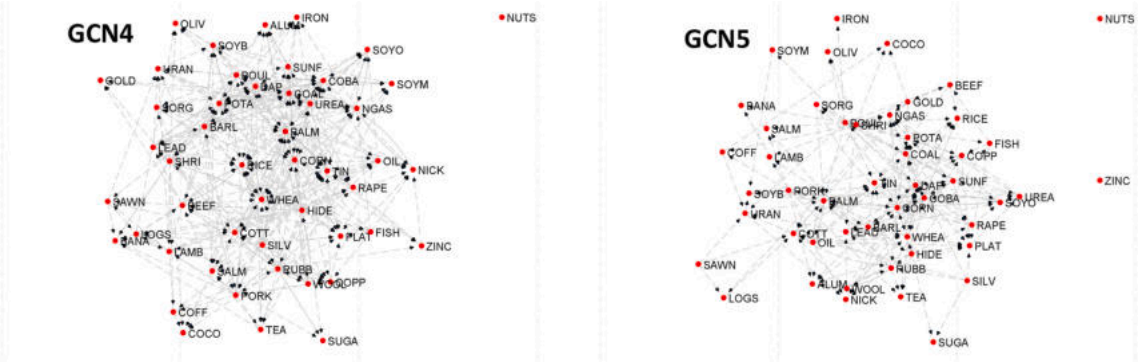

Fig. 4. Network structure GCN4-GCN5 – price first differences (dashed arcs indicate links across commodities and arrows indicate the direction).

Table 1
Pairwise Pearson correlations across GCN variants (QAP Method).

	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)
GCN1 (price levels)	1				
GCN2 (price levels)	0.764**	1			
GCN3 (price levels)	0.576**	0.754**	1		
GCN4 (price levels)	0.107	0.119	0.121	1	
GCN5 (price levels)	0.207	0.228	0.245*	0.417**	1

	GCN1 (price first differences)	GCN2 (price first differences)	GCN3 (price first differences)	GCN4 (price first differences)	GCN5 (price first differences)
GCN1 (price first differences)	1				
GCN2 (price first differences)	0.691**	1			
GCN3 (price first differences)	0.561**	0.813**	1		
GCN4 (price first differences)	0.086	0.094	0.098	1	
GCN5 (price first differences)	0.075	0.077	0.079	0.687**	1

	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)
GCN1 (price first differences)	0.191				
GCN2 (price first differences)	0.177	0.173			
GCN3 (price first differences)	0.191	0.196	0.203		
GCN4 (price first differences)	0.314**	0.327**	0.332**	0.182	
GCN5 (price first differences)	0.288**	0.298**	0.306**	0.157	0.171

*,** Statistically significant at 10 % and 5 % confidence level, respectively.

network is used for comparison. It is an undirected network in which points are connected by geodesics, i.e., the shortest paths between two points on a curved surface or within a geometric space.

5.3.1.1. Density. The key finding concerning density is that sparsity increases when moving from GCN1 to GCN3, from GCN4 to GCN5, and from price levels to first differences of prices. The densest network is GCN1 (price levels), with a density of 0.446, while the sparsest is GCN3 (first differences), with a density of 0.073, thus a reduction of over 80 %. As expected, average shortest path and diameter increase with sparsity, reaching 2.327 and 5 respectively in GCN3 (first difference), while they are 1.350 and 3 in GCN1 (price levels).

Table 2 also reports three conventional distance-based metrics that reflect density in terms of overall connectivity level of the network: the paths, the diameter and the average shortest path of the network. The "Paths (Largest Component)" refers to the total number of unique shortest paths that exist between all pairs of nodes. "Diameter (Largest Component)" expresses the length (in number of arcs) of the longest shortest path between any two nodes in the largest component. It indicates the maximum distance between the most distant nodes, providing a measure of the network's size in terms of connectivity. "Average Shortest Path (Largest Component)" is the average length of the shortest paths between all pairs of nodes in the largest component. It reflects how efficiently information or influence can travel across the network.

The key finding in this respect is that, only for GCN2-5 with first differences of prices, the largest component of the path is lower to 2352 since these GCNs are the only ones that present isolated nodes (see also Table A5 in Annex 4). The diameter and the average shortest path, as expected, increase with the increase of sparsity. Therefore, the indicator passes from 3 of GCN1 to 5 of GNC3 for both price levels and first differences, and from 2 of GCN4 to 3 of GNC5 and from 3 of GCN4 to 4 of GNC5 for levels and first differences, respectively. A similar gradient is observed in the average shortest path, which ranges from a minimum of 1.32 in GCN4 (price levels) to a maximum of 2.33 in GCN3 (first differences).

5.3.1.2. Granularity. Density alone does not offer sufficient insight into network granularity, that is, how connectivity is unevenly distributed depending on the presence of local structures or clusters. A network exhibits high granularity whenever it has several tightly connected subgroups and a heterogeneous distribution of link densities. Reciprocity is an helpful concept in this respect (Table 2). The key evidence is that this indicator varies between a maximum of 0.371 (GCN1, price

Table 2
Network-wide metrics across GCN variants compared to the corresponding geodesic network (maximum and minimum values by column in bold).

	Arcs	Density	Paths (largest component)	Diameter (largest component)	Average shortest path (largest component)	Reciprocity	Transitivity	Betweenness centralization	Indegree centralization	Outdegree centralization	λ_{max}
GCN1 (price levels)	**1049**	**0.446**	**2352**	3	1.350	**0.371**	0.789	0.033	**0.417**	0.310	0.524
GCN2 (price levels)	752	0.320	**2352**	3	1.528	0.290	0.764	0.051	0.397	0.354	0.616
GCN3 (price levels)	496	0.211	**2352**	**5**	1.837	0.195	0.924	0.055	0.359	0.380	0.707
GCN4 (price levels)	1033	0.439	**2352**	**2**	**1.325**	0.301	0.927	**0.029**	0.253	**0.508**	**0.425**
GCN5 (price levels)	355	0.151	**2352**	3	1.741	0.077	**0.955**	0.096	**0.164**	0.376	0.574
GCN1 (price first differences)	472	0.201	**2352**	3	1.719	0.192	0.504	0.059	0.391	0.220	0.513
GCN2 (price first differences)	251	0.107	2162	4	2.053	0.106	0.464	0.096	0.380	0.125	0.539
GCN3 (price first differences)	172	**0.073**	**1980**	**5**	**2.327**	0.096	0.612	0.088	0.223	**0.117**	0.745
GCN4 (price first differences)	349	0.148	2256	3	1.757	0.104	0.573	0.060	0.231	0.359	0.498
GCN5 (price first differences)	**201**	0.085	2162	4	2.077	**0.061**	**0.433**	**0.116**	0.168	0.274	**0.749**
Geodesic (undirected net)	1176	1	1176	1	1	1	1	0	0	0	3.992

levels) and a minimum of 0.061 (GCN5, first differences), values that are much lower than the hypothetical reciprocity within a geodesic network whose value is 1 (i.e., no form of granularity occurs). Eventually, reciprocity behaves very much like density and like the number of arcs, with the value of indicator declining from GCN1 to GCN3 and from GCN4 to GCN5, as well as passing from GCNs obtained with price levels to analogous networks obtained for first differences.

Reciprocity is just one of several dyadic properties and a deeper investigation of these latter may be informative on network granularity. Table 3 presents the dyadic configurations across GCN variants, highlighting the proportion of mutual, asymmetric, and null dyads. While in a geodesic network by design only mutual dyads are observed (100 %), the key evidence in terms of dyadic configurations is that in none of the GCNs this configuration is prevalent. Reciprocity is evidently associated with density and its presence declines as network sparsity increases moving from GCN1 (24 %) to GCN3 (7 %) and from GCN4 (20 %) to GCN5 (2 %) in the case of networks obtained with price levels. The same patters is observed for networks obtained with the first differences of prices where the incidence of mutual dyads ranges from a maximum of 7 % (GCN1) to a minimum of 1 % (GCN3 and GCN5). In all GCNs obtained with the first differences of prices, the null configuration is prevalent (up to 87 % in GCN3). This shift reflects a major topology change: the proportion of null dyads increases by over 50 % from GCN1 to GCN3. Asymmetric dyads remain relatively stable across variants, suggesting consistent unidirectional relationships.

Granularity may be investigated further via another indicator, transitivity (Table 2). It gives an indication of the degree to which nodes in a network tend to form local clusters: a high transitivity indicates a network showing tightly-knit clusters. The geodesic network has value 1 also for transitivity. The key finding about the GCNs is a gradient which is less regular and not fully correspondent with that observed for density and reciprocity. It seems that transitivity increases with sparsity since, in the case of price levels, it increases from GCN1 (0.789) to GCN3 (0.924) and it peaks in GCN5 (0.955). But it also declines when passing from networks obtained with levels to networks obtained with first differences. Eventually, the largest transitivity is found for GCN5 (price levels) while the lowest is for GCN5 (first differences) (0.433). The main implication of this evidence is that the adoption of different VAR modelling strategies, as well as of different data transformations, has an impact not only on the network sparsity but also, and more importantly, on its internal structure (granularity).

As for reciprocity, transitivity is just one of several triadic properties. The triad census classifies all possible configurations of triads in a directed network. Table 4 reports the distribution of 16 possible triadic configurations across GCN variants, providing insight into the local

Table 3
Dyadic configurations across GCN variants (maximum and minimum values by column in bold).

	Mutual	Asymmetric	Null	Total
GCN1 (price levels)	**284 (24 %)**	481 (41 %)	411 (35 %)	1176
GCN2 (price levels)	169 (14 %)	414 (35 %)	593 (51 %)	1176
GCN3 (price levels)	81 (7 %)	334 (28 %)	761 (65 %)	1176
GCN4 (price levels)	239 (20 %)	**555 (47 %)**	**382 (33 %)**	1176
GCN5 (price levels)	24 (2 %)	307 (26 %)	845 (72 %)	1176
GCN1 (price first differences)	76 (7 %)	320 (27 %)	780 (66 %)	1176
GCN2 (price first differences)	24 (2 %)	203 (17 %)	949 (81 %)	1176
GCN3 (price first differences)	15 (1 %)	**142 (12 %)**	**1019 (87 %)**	1176
GCN4 (price first differences)	33 (3 %)	283 (24 %)	860 (73 %)	1176
GCN5 (price first differences)	**11 (1 %)**	178 (15 %)	987 (84 %)	1176
Geodesic (undirected net)	1176	0	0	1176

Table 4
Triadic configurations across GCN variants (maximum and minimum values by column in bold).

	003	012	021D	021U	021C	030T	030C	102	120D	120U	120C	111D	111D	201	210	300	Total
GCN1 (price levels)	6.6 %	17.6 %	3.7 %	**5.4 %**	5.9 %	7.1 %	0.7 %	9.3 %	**4.6 %**	4.8 %	4.7 %	**7.9 %**	6.1 %	2.2 %	**9.5 %**	**3.9 %**	100 % (18424)
GCN2 (price levels)	16.4 %	28.1 %	3.9 %	5.0 %	6.0 %	5.5 %	0.5 %	**10.0 %**	2.3 %	3.0 %	2.6 %	5.1 %	4.3 %	1.3 %	4.5 %	1.4 %	100 % (18424)
GCN3 (price levels)	31.9 %	33.4 %	3.9 %	4.6 %	3.7 %	4.4 %	**0.1 %**	7.4 %	0.9 %	2.0 %	1.3 %	1.9 %	2.4 %	0.5 %	1.6 %	0.3 %	100 % (18424)
GCN4 (price levels)	**5.3 %**	**16.5 %**	**8.3 %**	4.9 %	**7.2 %**	**9.9 %**	**0.9 %**	5.1 %	3.2 %	**6.1 %**	**5.4 %**	5.2 %	**9.5 %**	**3.0 %**	7.6 %	1.7 %	100 % (18424)
GCN5 (price levels)	38.3 %	**38.8 %**	5.7 %	4.0 %	5.1 %	2.0 %	**0.1 %**	3.0 %	**0.1 %**	0.3 %	0.2 %	0.7 %	1.6 %	0.1 %	**0.0 %**	**0.0 %**	100 % (18424)
GCN1 (price first differences)	31.4 %	34.8 %	3.3 %	4.7 %	5.6 %	2.1 %	0.4 %	7.2 %	0.7 %	0.6 %	1.1 %	3.7 %	2.6 %	0.9 %	0.7 %	0.1 %	100 % (18424)
GCN2 (price first differences)	55.4 %	29.9 %	1.6 %	3.4 %	2.7 %	1.0 %	**0.1 %**	3.2 %	0.2 %	**0.1 %**	0.2 %	1.4 %	0.6 %	0.1 %	0.1 %	**0.0 %**	100 % (18424)
GCN3 (price first differences)	**66.8 %**	24.6 %	**1.0 %**	**1.6 %**	**1.9 %**	**0.3 %**	**0.1 %**	2.4 %	**0.1 %**	**0.1 %**	**0.1 %**	0.7 %	**0.3 %**	**0.0 %**	**0.0 %**	**0.0 %**	100 % (18424)
GCN4 (price first differences)	41.5 %	35.6 %	5.9 %	3.2 %	3.6 %	2.0 %	**0.1 %**	4.0 %	0.3 %	0.3 %	0.3 %	1.1 %	1.7 %	0.2 %	0.2 %	**0.0 %**	100 % (18424)
GCN5 (price first differences)	60.7 %	29.6 %	2.8 %	**1.6 %**	**1.9 %**	0.6 %	**0.1 %**	**1.7 %**	**0.1 %**	**0.1 %**	**0.1 %**	**0.3 %**	0.5 %	**0.0 %**	**0.0 %**	**0.0 %**	100 % (18424)
Geodesic (undirected net)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100 %	100 % (18424)

LEGEND - 003: No arcs between the three nodes; 012: One arc between two nodes, the third node is isolated; 021D: Two directed arcs originating from the same node towards two distinct nodes; 021U: Two directed arcs arriving at the same node from two distinct nodes; 021C: A chain of two consecutive directed arcs; 030T: A triangle with one directed arc missing; 030C: A triangle with one directed arc missing, but with an additional directed arc forming a "T"; 102: One directed arc and one bidirectional arc between three nodes; 120D: Two directed arcs originating from the same node and one bidirectional arc between two nodes; 120U: Two directed arcs arriving at the same node and one bidirectional arc between two nodes; 120C: A chain of two consecutive directed arcs and one bidirectional arc between two nodes; 111D: A triangle with one directed arc missing and one bidirectional arc between two nodes; 111U: A triangle with one directed arc missing and one bidirectional arc between two nodes, but with the opposite direction compared to 111D; 201: Two bidirectional arcs and one directed arc between three nodes; 210: A triangle with two bidirectional arcs and one directed arc; 300: A triangle with all three arcs bidirectional.

clustering and structural complexity of the networks. In the geodesic network only one triadic configuration occurs (that designated with 300, a triangle with all three arcs bidirectional) as an expression of homogenous density. The key finding here is that all GCNs show a more articulated set of triadic configurations. In denser networks (e.g., GCN1 and GCN2 with price levels), configurations such as 012 (one arc, one isolated node) and 030T (incomplete triangles) are relatively frequent, indicating partial clustering. As sparsity increases (e.g., GCN3 and GCN5), the prevalence of configuration 003 (no arcs among three nodes) rises sharply, reaching 66.8 % and 38.3 %, for the first difference and the level of prices, respectively. An increase of over 30 % compared to GCN1 in both cases. To further confirm the structural difference with respect to the geodesic network, fully connected triads (300) represent the less frequent configuration (<4 %) in all cases thus confirming the absence of dense, reciprocal clusters.

Positioned conceptually between granularity and peripherality, Table 2 also presents the network-level centralization indicators: betweenness centralization, indegree centralization, outdegree centralization. They measure how much a network is organized around its most central nodes, reflecting the overall inequality in node centrality across the network. High values of betweenness centralization indicate that a few nodes act as key bridges in the network. In a geodesic network, betweenness centralization is 0, because all nodes are equally central. Indegree centralization measures the extent to which incoming connections are concentrated on a few nodes. In a geodesic network it takes the value 0, since every node receives the same number of incoming links.

The key result here is that betweenness centralization appears to follow the opposite trend of density for levels, while for the first differences of prices GCN2 shows a higher value than GCN3. However, for both levels and first differences, GCN5 is the network version with the highest value (about 0.1), indicating a mild tendency to form a core-periphery structure compared to the geodesic network. On the contrary, the indegree centralization replicates the gradient of the density indicator with its value decreasing going from GCN1 to GCN3 and from GCN4 to GCN5 in both levels and first differences of prices. The highest value (0.417) is found for GNC1 for price levels. Overall, betweenness and degree centralization remain low overall, suggesting limited dominance by individual nodes.

A more mixed relationship with the density of the network is observed for outdegree centralization that measures the extent to which outgoing connections are concentrated on a few nodes. Also in this case, in a geodesic network it takes the value 0 (all nodes have the same number of outgoing links). Here, in the case of price levels, the indicator increases passing from GCN1 to GCN3 but also passing from GCN5 to GCN4, the latter being the largest observed value (0.508). In the case of the first differences of prices, the gradient returns to what was observed for the indegree centralization, with and higher value passing from GNC3 to GCN1 and from GCN5 to GCN4. Also in this case, the highest value is found for GCN4 (0.359).

5.3.1.3. Peripherality. As discused in Section 3.3.1, peripehreality is investigated by looking atthe number and frequency of global and local bridges. Table 5 presents the distribution of global and local bridges across GCN variants, along with the nodes involved (isolated nodes are reported in detail in Table A5, Annex 4), highlighting network-wide peripherality. It is worth noticing that the geodesic network does not present either local or global bridges as all nodes are connected to all other nodes.

The key finding is that the number of global bridges increases with network sparsity. Since this indicator expresses the tendency of a node to be isolated or peripheral, the number of global bridges tends to increase passing from 2 in GCN1 (price levels) to 16 in GCN3 (first differences), with corresponding frequencies substantially rising from 4 % to 33 %. Local bridges also proliferate in sparser networks, reaching 103 in GCN5

(price levels) and 78 in GCN5 (first differences). These bridges indicate critical links whose removal would fragment the network, underscoring the fragility of certain configurations.

It is not necessarily true, however, that networks obtained with PW-VAR modelling present less peripherality than those obtained with S-VAR modelling. For both price levels and first differences the GCN showing the largest number of global bridges, therefore the largest peripherality of groups of nodes, is GCN3. In this latter case, the frequency of global bridges, indicating how often the global bridging condition is met across all node pairs, is 20 % and 33 % for the network obtained with levels and first differences, respectively. Nodes more often affected by this peripherality tendency are also reported in Table 5 and mostly, though not exclusively, concern the group of other raw materials (Hide, for instance), and various soft commodities (e.g., Pork and Tea), while metals appear less peripheral.

5.3.2. Network correlation

Remaining at the network level, it is worthwhile to explore potential similarities between these GCNs and synthetic networks, i.e., networks with known properties and structures. Since each synthetic model has specific topological characteristics, it provides an interesting baseline to assess whether, and to what extent, the observed GCNs result from non-random mechanisms or specific system dynamics. In particular, real-world networks are characterized by emerging patterns resulting from self-organization within the network, patterns that are typically lacking in synthetic networks.

Table 6 compares the structural similarity of GCN variants with five synthetic networks (random, ring, small-world, lattice and geodesic networks) using pairwise Pearson correlations (see Section 3.3.2). The key finding is that, despite the previously discussed quantitative and qualitative (i.e., structural) differences among the ten variants, across all GCNs correlations with random, ring, small-world, and lattice networks are negligible (typically $|\rho| < 0.05$ and not statistically significant), indicating minimal structural resemblance. In contrast, correlations with the geodesic network are consistently negative and mostly statistically significant. In the case of price levels, this negative correlation always exceeds −0.6, ranging from −0.665 of GNC2 to −0.556 of GCN3 and highlighting a systematic deviation from idealized connectivity. In the case of first difference of prices, with the only exception of GCN1 (−0.644), GCN show weaker and a less consistent correlations, though not statistically significant only for GCN3.

Ultimately, no synthetic network provides even a mild approximation of our commodity price networks. The emergent properties of the latter cannot be synthetically reproduced through a mechanical generation of connections among nodes. The case of the geodesic model actually suggests that, while its synthetic structure offers some insight into the structure of the GCNs, it does so by highlighting what GCNs are not: real-world commodity price networks are not uniformly dense or fully connected and do not exhibit the highly dense and organized structure typical of the geodesic model. In fact, it is in the denser GCNs that the negative correlation with the geodesic model is strongest, indicating that this density does not result from a mechanical repetition of links among all nodes, but rather from the spontaneous emergence of both local and global clusters.

Table 6
Pairwise Pearson correlations across GCN variants and corresponding (49 nodes) synthetic networks (QAP Method).

	Random network	Ring network	Small network	Lattice network	Geodesic network
GCN1 (price levels)	−0.0062	−0.0477	−0.0451	−0.0138	−0.6564**
GCN2 (price levels)	0.0087	−0.0484	−0.0353	0.0010	−0.6645**
GCN3 (price levels)	0.0042	−0.0352	−0.0277	−0.0139	−0.5557**
GCN4 (price levels)	−0.0046	0.0276	0.0214	0.0074	−0.6138**
GCN5 (price levels)	−0.0148	0.0061	0.0147	−0.0016	−0.6414**
GCN1 (price first differences)	0.0093	0.0064	0.0102	0.0012	−0.6438**
GCN2 (price first differences)	0.0290	−0.0046	−0.0046	0.0004	−0.2583*
GCN3 (price first differences)	0.0093	−0.0020	−0.0020	−0.0018	−0.1551
GCN4 (price first differences)	−0.0272	−0.0349	−0.0393	−0.0043	−0.3643**
GCN5 (price first differences)	0.0037	−0.0478	−0.0478	−0.0194	−0.2404*

*,**Statistically significant at 10 % and 5 % confidence level, respectively.

5.3.3. Node-level metrics

The network-level metrics considered thus far contributes to reveal how model choice and data transformation jointly shape the topology and resilience of commodity price networks. Beside this aspect,

Table 5
In-degree and out-degree global and local bridges across GCN variants (maximum and minimum values by column in bold).

	Global bridges	Frequency	Local bridges	Nodes (Outgoing)	Nodes (Incoming)
GCN1 (price levels)	2	4 %	8	Pork, Rubb	Tea, Hide
GCN2 (price levels)	8	16 %	20	Pork, Poul, Soyb, Shri, Dap, Logs, Rubb, Sawn	Bana, Nuts, Pork, Coff, Shri, Tea, Hide, Logs
GCN3 (price levels)	11	20 %	33	Alum, Nuts(2), Pork, Soyb, Shri, Suga, Tea, Dap, Hide, Sawn	Silv, Bana, Lamb, Pork, Coff, Fish, Salm, Suga, Logs (2), Wool
GCN4 (price levels)	**0**	**0 %**	**6**	None	None
GCN5 (price levels)	1	2 %	**103**	Lamb	Pork
GCN1 (price first differences)	4	8 %	47	Barl, Sorg, Tea, Logs	Bana, Fish, Sunf, Rubb
GCN2 (price first differences)	11	22 %	61	Ngas, Bana, Barl, Coco, Lamb, Nuts, Sorg, Suga, Cot, Hide, Logs	Ngas, Alum(2), Uran, Bana, Nuts, Suga, Tea, Dap, Rubb(2)
GCN3 (price first differences)	**16**	**33 %**	67	Ngas, Coba, Gold, Lead, Bana, Coco, Nuts(2), Poul, Oliv, Cott(2), Hide, Pota(2), Urea	Coal, Alum(2), Plat, Uran, Barl, Beef, Coco, Nuts, Soyb, Whea(2), Suga, Sunf, Tea, Cott
GCN4 (price first differences)	7	14 %	71	Copp, Gold, Nick, Bana, Soym, Rubb, Wool	Coal, Gas, Alum, Tin, Soyb, Hide, Wool
GCN5 (price first differences)	13	27 %	78	Gold, Plat, Tin, Corn, Pork, Poul(3), Coff, Soym, Hide, Logs, Pota	Coal, Iron, Uran, Pork, Soyb, Shri, Soym, Suga, Cott, Hide, Logs, Urea(2)
Geodesic (undirected net)	0	0 %	0	None	None

however, one of the objectives of the analysis remains to investigate which are the leader commodities, namely the network nodes with systemic relevance. This can be achieved firstly by counting the connections of any single node and, then, by adopting more sophisticated indicators (see Section 3.3.3).

Figs. 5–8 display the number of arcs per node in various GCN variants. Due to space constraints, we only present the sparsest networks, namely, GCN3 and GCN5 for both levels and first differences of prices.[15] Figures present the same data as Tables A3 and A4 (Annex 3), but grouped by commodity type and arranged in descending order. In addition to incoming and outgoing arcs, the figures also report the node-by-node balance. This visualization complements the centrality metrics reported in Tables 7–10 and offers an additional perspective on how influence and interdependence are distributed within the network. A node, or group of nodes, can emerge either through its ability to generate shocks (outgoing causality) or its susceptibility to them (incoming causality). Core nodes typically score high on both dimensions, as effective shock transmission requires both reception and propagation. The balance between outgoing and incoming connections thus reveals whether a node acts primarily as a net generator of shocks or merely as a transmitter.

The key fact in Fig. 5 (GCN3 – price levels) is that metals dominate the upper ranks, with nodes such as Tin, Copper, and Zinc showing the highest number of connections. This is confirmed by Fig. 6 (GCN5 – price levels) but it reveals a more hierarchical structure: the distribution is steeper, with a few nodes concentrating most of the arcs, while many others remain marginal. This reflects the higher outdegree centralization and betweenness observed in GCN5, suggesting a shift toward directional influence and reduced redundancy. Nonetheless, minerals (e.g., Iron) confirm their prominence but now some soft commodities also become central (e.g., Palm oil, Hide).

In the case of first differences of prices, the key evidence is that the overall number of arcs per node drops significantly, and the distribution flattens. This sparsification confirms the effect of differencing on network topology. In Fig. 7 (GCN3) many nodes lose their central role, and inter-sectoral links diminish. Metals retain some prominence, but food and agricultural commodities become more relevant. Fig. 8 (GCN5) shows an even more polarized structure: a few nodes concentrate the majority of arcs, while the rest of the network is fragmented. It is confirmed that, beside metals, some agricultural and food commodities emerge as central (e.g., Corn, Palm oil).

In sum, Figs. 5–8 provide compelling evidence of how sparsity, directionality, and sectoral composition interact to shape the architecture of commodity price networks. They also show that node-level metrics vary significantly across GCNs, indicating that alternative VAR models and datasets affect not only network density but also the qualitative roles of individual and grouped nodes. This may be explained by the fact that increasing sparsity tends to highlight the most relevant linkages.

Beside the differences observed across the GCNs, some noteworthy regularities also emerge. Among the five commodity groups, a somewhat unexpected finding is that energy commodities do not play a particularly critical role within the network. They tend to act more as transmitters than as primary sources of shocks, as evidenced by their predominantly negative balance between outgoing and incoming connections, lower than that of most other groups. In contrast, metals consistently emerge as key sources of price shocks. Across most GCNs, their balance is not only largely positive but also typically the highest among all groups, indicating a structurally central role in driving network dynamics.

The behaviour of commodity groups, however, does not rule out the possibility that core and critical nodes may emerge within each of them. When core or leader nodes are identified not only based on the number of outgoing and incoming connections but also on their positive balance, some robust evidence emerges across the different GCNs. Among energy commodities, Natural Gas appears to act as a leader node more consistently than Oil and Coal. Copper, Gold, Iron, and Tin seem to be the key commodities among metals, while cereals, particularly Wheat, Rice, and Corn, emerge as critical nodes among agricultural commodities. Edible vegetable oils, especially Palm oil, also stand out as leader nodes among food products. Finally, within the group of other raw materials, fertilizers, notably Potash and Urea, appear to behave as more central nodes.

As discussed in Section 3, the centrality of a commodity within the network is not determined solely by its immediate neighbours, that is, the nodes directly connected to it, but also by its position within the entire structure. Specifically, a node's ability to connect, directly or indirectly, to all other nodes enhances its potential to propagate and even amplify shocks throughout the system. Global centrality measures are helpful to investigate this aspect.

Table 7 reports the node-level clustering coefficient that expresses the capacity of a node to act as an aggregator. A high clustering coefficient means the node's neighbours are well connected (forming tight-knit communities). Beside the already commented substantial variation emerging across GCN variants, the key finding is that slightly higher values tend to concentrate in the groups of metals and energy commodities. However, particularly when moving towards less dense GCNs, nodes with high values are found across all groups. Commodities such as Nickel, Lead, and Zinc exhibit high clustering in denser networks (e.g., Nickel reaches 1.000 in GCN3– price levels), indicating strong local interdependence. In contrast, several nodes (e.g., Aluminum, Tea, Hide) show zero clustering in sparser networks, reflecting isolation or weak local ties. Sectoral averages confirm this pattern: metals and energy commodities maintain higher clustering than agricultural and food items, especially in price level-based GCNs.

But the influence of a single node within the network can be better expressed by node-level degree centrality indicators (Tables 8–10): betweenness, closeness and eigenvector centrality. They provide complementary and largely concordant evidence. The key fact emerging from Table 8 is that higher betweenness centrality scores, pointing to commodities that act as strategic intermediaries in price transmission, are diffused across all commodity groups. GCN5 (first differences) shows the highest average centrality (101.11), with extreme values for commodities like Poultry (320.0), Corn (312.4), and Diammonium Phosphate (i.e., DAP) (289.2), indicating their pivotal role in connecting otherwise distant nodes. In contrast, GCN4 (price levels) has the lowest average (21.90), reflecting a more decentralized structure. Several nodes (e.g., Aluminum, Hide) exhibit zero centrality in specific variants, suggesting marginal influence. Sectoral averages reveal that, while metals and energy maintain consistent centrality across models, food commodities gain relevance in sparser networks.

The key finding about closeness centrality (Table 9) is the confirmation that passing from GCNs based on price levels to those based on the first differences of prices may substantially modify the evidence about centralization. GCN1 (price levels) has the highest average closeness (0.755), while GCN3 (price levels) drops to 0.567, a reduction of nearly 25 %. Commodities like Copper, DAP, and Rice consistently rank high, suggesting central roles in price transmission since they can reach other nodes more quickly. In contrast, nodes such as Logs, Tea, and Sawn show low closeness in sparse networks, indicating peripheral positions. Sectoral averages confirm that energy and metals maintain higher closeness than food and agricultural commodities.

Eigenvector centrality is only defined for connected networks, i.e., those without isolated nodes. Therefore, this indicator is available and reported for all GCNs based on price levels, and only for GCN1 based on first differences in prices. Table 10 presents node-level eigenvector centrality across these GCN variants. The key result confirms the central role of minerals and, to a lesser extent, of energy commodities. GCN3

[15] These are arguably the cases that exhibit the most robust connections. All other networks (GCN1, GCN2, and GCN4) are available upon request and provide largely consistent evidence.

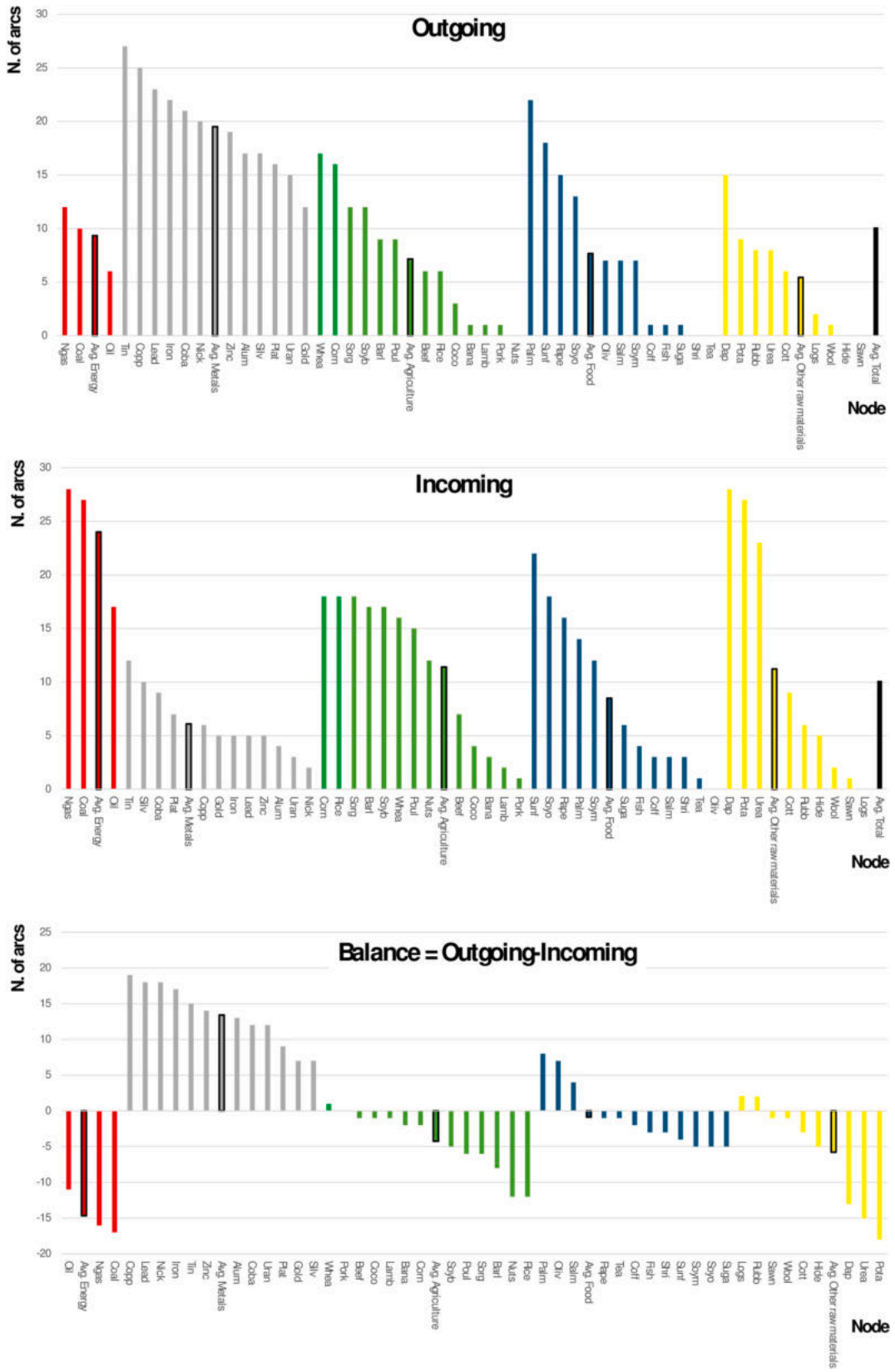

Fig. 5. Number of arcs per node in GCN3 by group and in descending order (price levels).

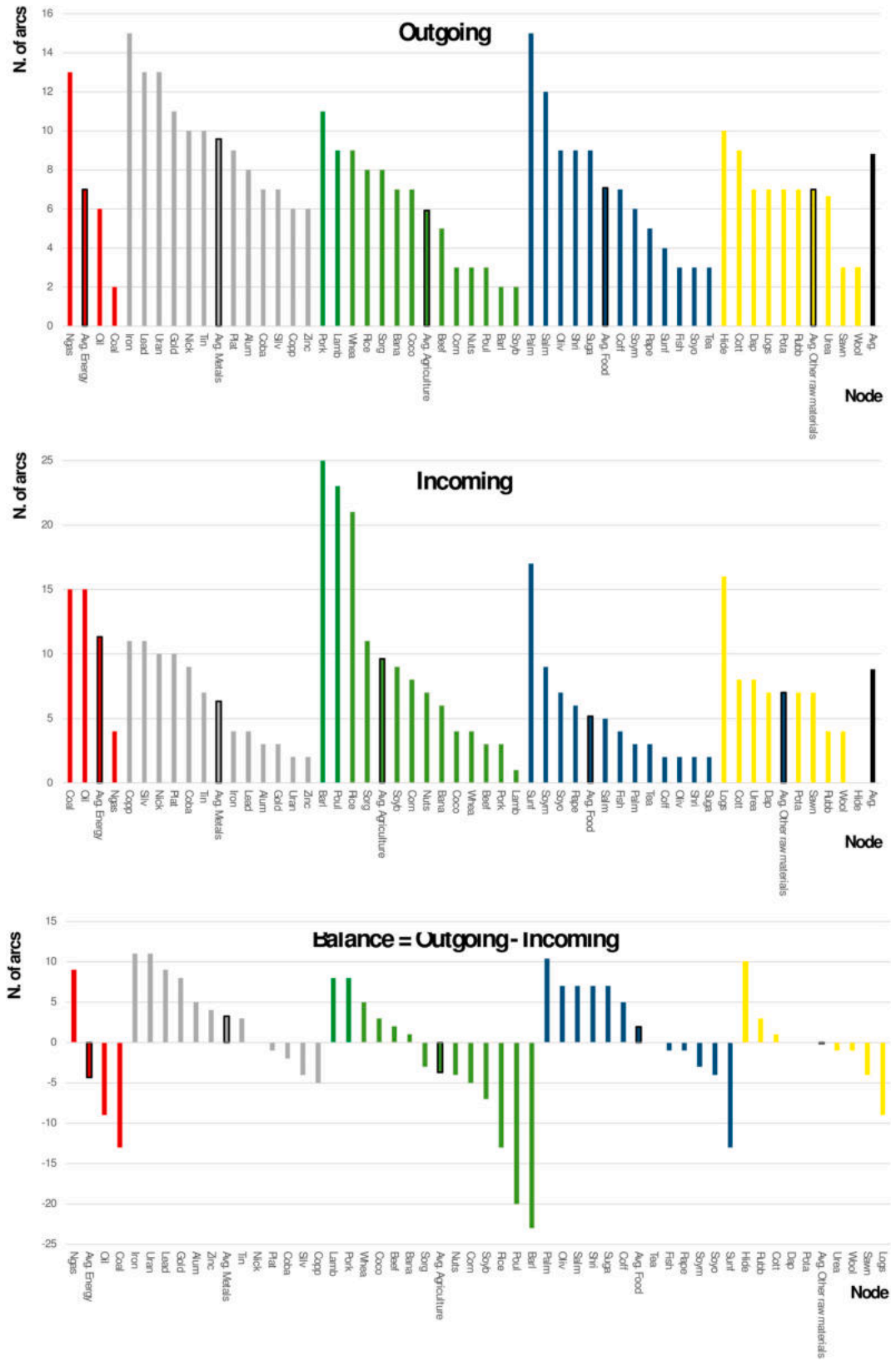

Fig. 6. Number of arcs per node in GCN5 by group and in descending order (price levels).

(price levels) shows the highest average centrality for energy commodities (e.g., Natural Gas: 0.216, Coal: 0.211), while sparse networks like GCN5 display lower values overall. DAP, Poultry, and Rice emerge as structurally central nodes in several variants, with Diammonium Phosphate reaching 0.246 in GCN1 (first differences), the highest observed value. In contrast, commodities like Logs, Tea, and Sawn show minimal centrality (e.g., Logs: 0.008 in GCN2), indicating peripheral roles. Sectoral averages confirm that metals and energy commodities

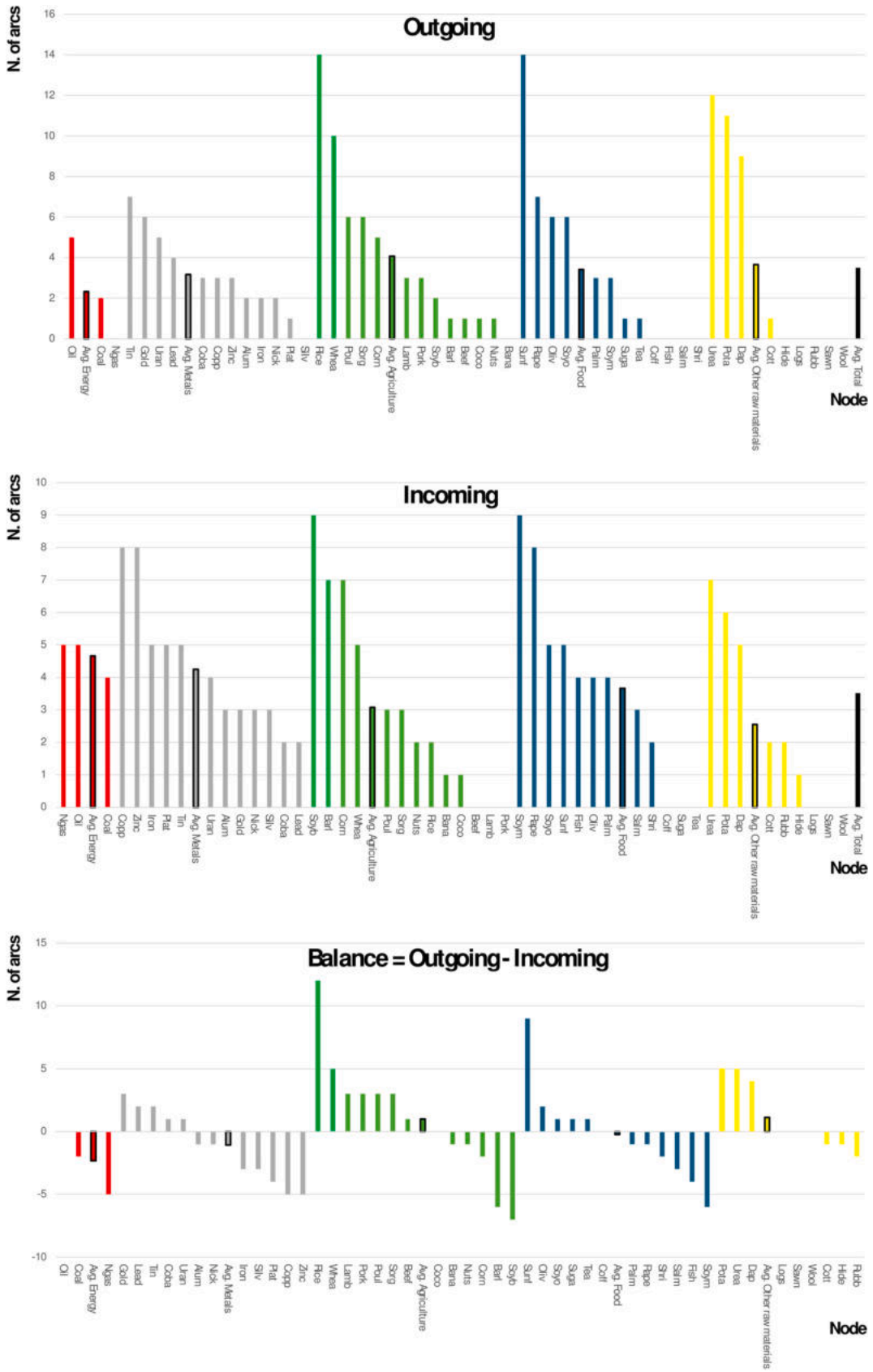

Fig. 7. Number of arcs per node in GCN3 by group and in descending order (price first differences).

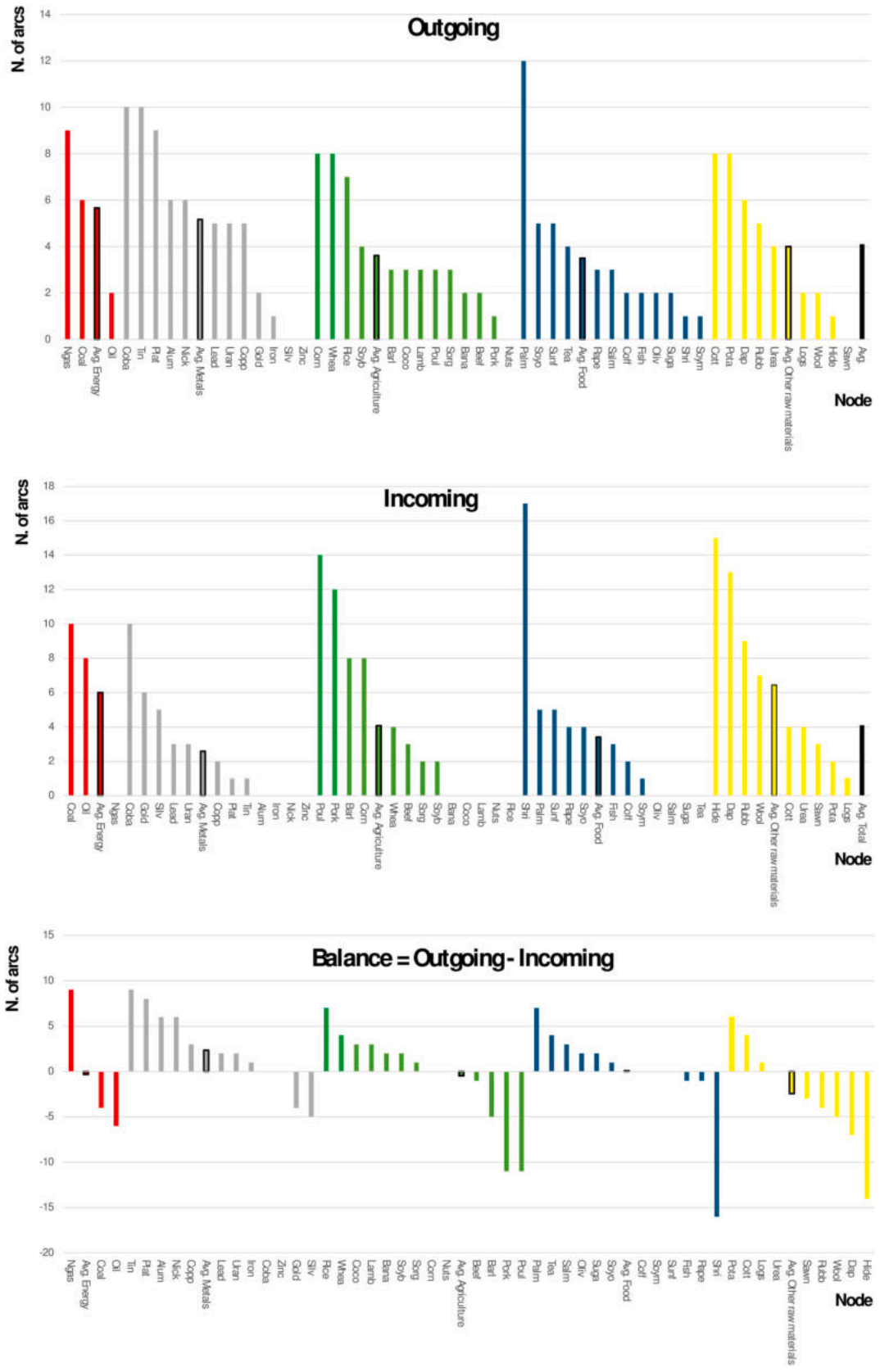

Fig. 8. Number of arcs per node in GCN5 by group and in descending order (price first differences).

Table 7
Node-level clustering coefficients across GCN variants (maximum and minimum values by column in bold).

	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)	GCN1 (price first differences)	GCN2 (price first differences)	GCN3 (price first differences)	GCN4 (price first differences)	GCN5 (price first differences)
Coal	0.603	0.475	0.403	0.521	0.190	0.333	0.214	0.083	0.289	0.264
Ngas	0.624	0.512	0.392	0.527	0.333	0.286	0.150	0.050	0.097	**0.000**
Oil	0.690	0.550	0.408	0.450	0.157	0.314	0.322	0.350	0.278	0.107
Alum	0.525	0.486	0.167	0.578	0.333	0.250	0.286	**0.000**	**0.000**	**0.000**
Coba	0.708	0.633	0.486	**0.722**	0.194	0.347	0.167	**0.000**	0.291	0.178
Copp	0.716	0.506	0.267	0.429	0.155	0.433	0.393	0.268	**0.000**	**0.000**
Gold	0.524	0.357	0.150	0.500	0.333	0.357	0.196	**0.000**	**0.000**	0.133
Iron	0.607	0.533	0.350	0.555	**0.000**	0.313	0.321	0.150	**0.000**	**0.000**
Lead	0.735	0.518	0.250	0.466	**0.000**	0.386	0.350	0.000	0.367	0.333
Nick	**0.750**	0.650	**1.000**	0.464	0.153	0.364	**0.400**	0.333	**0.000**	**0.000**
Plat	0.704	0.577	0.452	0.438	0.156	0.409	0.350	0.100	**0.000**	**0.000**
Silv	0.504	0.523	0.322	0.471	0.073	0.222	0.167	0.333	0.168	0.150
Tin	0.602	0.508	0.462	0.500	0.190	0.364	0.250	0.200	0.417	**0.000**
Uran	0.704	0.593	0.500	0.611	**0.000**	0.350	0.300	**0.417**	0.200	**0.000**
Zinc	0.709	**0.804**	0.800	**0.167**	**0.000**	0.358	0.347	0.250	**0.000**	**0.000**
Bana	0.431	0.476	0.333	0.437	0.200	0.333	**0.000**	**0.000**	**0.000**	**0.000**
Barl	0.608	0.517	0.463	0.469	0.162	0.231	0.133	0.095	0.182	0.196
Beef	0.507	0.482	0.238	0.350	**0.500**	0.286	**0.000**	**0.000**	0.200	**0.000**
Coco	0.472	0.321	0.083	0.533	0.083	0.300	**0.000**	**0.000**	**0.000**	**0.000**
Corn	0.595	0.554	0.467	0.515	0.214	0.244	0.268	0.167	0.218	0.196
Lamb	0.673	0.767	0.500	0.438	**0.000**	0.167	**0.000**	**0.000**	0.143	**0.000**
Nuts	0.548	0.415	0.197	0.458	0.143	0.097	**0.000**	**0.000**	**0.000**	**0.000**
Pork	0.321	0.083	**0.000**	0.515	0.333	**0.000**	**0.000**	**0.000**	0.167	0.061
Poul	0.655	0.521	0.381	0.439	0.140	0.319	0.150	0.333	0.154	0.082
Rice	0.561	0.459	0.399	0.413	0.093	0.286	**0.000**	**0.000**	0.167	**0.000**
Sorg	0.651	0.561	0.441	0.468	0.127	0.275	0.119	0.167	0.267	**0.500**
Soyb	0.567	0.487	0.346	0.500	0.194	0.357	0.227	0.153	0.304	**0.500**
Whea	0.654	0.562	0.508	0.476	0.167	0.373	0.300	0.350	0.167	**0.000**
Coff	0.500	0.400	0.333	0.450	**0.000**	0.333	**0.000**	**0.000**	0.167	**0.000**
Fish	0.562	0.482	0.333	0.454	0.083	0.273	0.333	0.167	0.143	0.167
Oliv	0.655	0.333	**0.000**	0.544	0.500	0.214	0.300	0.167	**0.000**	**0.000**
Palm	0.607	0.533	0.412	0.426	0.333	0.411	0.350	**0.000**	0.200	0.150
Rape	0.604	0.580	0.517	0.489	0.067	0.321	0.264	0.196	**0.500**	0.333
Salm	0.534	0.357	0.167	0.466	0.100	**0.491**	0.238	0.167	0.250	**0.000**
Shri	0.581	0.310	**0.000**	0.417	**0.000**	0.250	0.100	**0.000**	0.125	0.074
Soym	0.592	0.535	0.348	0.529	0.139	0.279	0.244	0.167	**0.000**	**0.000**
Soyo	0.677	0.589	0.497	0.515	0.214	0.348	0.400	0.400	0.250	0.167
Suga	0.575	0.471	0.233	0.460	**0.000**	0.095	**0.000**	**0.000**	**0.000**	**0.000**
Sunf	0.575	0.534	0.422	0.439	0.125	0.429	0.300	0.300	0.300	0.250
Tea	0.585	0.381	**0.000**	0.453	0.167	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Cott	0.621	0.516	0.458	0.431	0.143	0.250	0.250	**0.000**	0.300	0.333
Dap	0.566	0.461	0.361	0.498	0.119	0.346	0.400	0.300	0.180	0.122
Hide	0.591	0.470	0.250	0.451	**0.000**	0.100	**0.000**	**0.000**	0.145	0.114
Logs	**0.200**	**0.000**	**0.000**	0.482	0.125	**0.000**	**0.000**	**0.000**	0.048	**0.000**
Pota	0.639	0.567	0.439	0.690	0.262	0.304	0.196	0.133	0.300	**0.500**
Rubb	0.431	0.379	0.400	0.483	0.167	0.286	0.350	**0.000**	0.149	0.097
Sawn	0.333	**0.000**	**0.000**	0.486	0.143	**0.000**	**0.000**	**0.000**	0.250	0.333
Urea	0.613	0.594	0.514	0.471	0.286	0.319	0.268	0.167	0.187	0.250
Wool	0.645	0.500	0.500	0.507	**0.000**	0.167	0.000	**0.000**	0.200	0.119
Avg.	0.584	0.478	0.346	0.481	0.155	0.277	0.192	0.122	0.158	0.117
Avg. Energy	0.639	0.512	0.401	0.500	0.227	0.311	0.229	0.161	0.221	0.124
Avg. Metals	0.649	0.557	0.434	0.492	0.132	0.346	0.294	0.171	0.120	0.066
Avg. Agriculture	0.557	0.477	0.335	0.462	0.181	0.251	0.092	0.097	0.151	0.118
Avg. Food	0.587	0.459	0.272	0.470	0.144	0.287	0.211	0.130	0.161	0.095
Avg. Other raw materials	0.515	0.387	0.325	0.500	0.138	0.197	0.163	0.067	0.195	0.208

maintain higher eigenvector centrality than food and agricultural commodities.

As mentioned in Section 3.3.3, the eigenvalue equation associated with any GCN can be interpreted not only at the node level, but also at the network level. Specifically, the maximum eigenvalue, λ_{max}, of the adjacency matrix reflects the overall network's shock propagation potential. This value is reported in the last column of Table 2 for all GCNs under consideration. The key evidence is that the networks with the highest eigenvalues are not the densest ones. In the case of price levels λ_{max} increases from GCN1 (0.524) to GCN3 (0.707) and from GCN4 (0.425) to GCN5 (0.474). The same is observed for the networks based on first differences where the highest values are found for GCN3 and GCN5 at around 0.75. Notably, the networks with the highest λ_{max} tend to be the sparsest ones. In any case, λ_{max} is always well below 1, thus signalling substantial network stability and a tendency to attenuate rather than amplify signals throughout the network. The reference to the geodesic network is particularly interesting as in this network λ_{max} takes a value very close to the theoretical maximum of 4.

5.3.4. *Community structure*

Interacting nodes may give rise to an intermediate layer of network characteristics, situated between node-level and network-level

Table 8
Node-level betweenness centrality across GCN variants (maximum and minimum values by column in bold).

	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)	GCN1 (price first differences)	GCN2 (price first differences)	GCN3 (price first differences)	GCN4 (price first differences)	GCN5 (price first differences)
Coal	46.8	91.1	87.7	28.8	67.8	40.3	61.5	32.4	110.8	259.1
Ngas	37.5	**97.1**	**113.1**	17.9	101.7	31.4	51.5	**0.0**	151.0	**0.0**
Oil	27.8	29.8	5.9	19.0	92.2	94.7	54.2	27.7	26.2	44.3
Alum	16.7	16.0	48.2	16.4	19.5	47.1	46.9	39.3	**0.0**	**0.0**
Coba	19.4	35.5	79.9	5.2	50.3	37.1	36.0	50.8	165.3	207.7
Copp	27.7	45.9	33.0	11.9	74.5	7.7	13.2	14.9	2.4	7.7
Gold	31.1	47.9	18.5	7.2	63.2	91.5	173.9	164.1	1.3	82.9
Iron	48.2	32.9	19.9	29.5	141.4	77.6	66.3	57.0	**0.0**	**0.0**
Lead	15.1	37.2	20.5	**64.3**	56.2	35.0	47.1	28.6	36.9	33.1
Nick	18.0	9.5	15.1	33.7	117.2	34.0	16.0	7.1	15.3	0.0
Plat	18.0	43.3	50.7	46.5	127.5	18.4	14.4	30.8	**0.0**	21.7
Silv	31.1	43.6	102.2	16.1	94.6	36.2	**0.0**	**0.0**	**0.0**	**0.0**
Tin	38.3	75.1	106.4	16.2	71.8	53.8	79.5	58.9	118.8	156.0
Uran	27.7	28.0	8.1	43.4	50.2	38.9	115.7	96.4	42.5	80.7
Zinc	15.6	20.0	48.2	12.0	10.5	36.0	38.0	58.3	2.0	**0.0**
Bana	63.5	85.6	0.2	53.1	145.8	**1.0**	0.7	0.0	3.8	**0.0**
Barl	20.6	29.0	4.3	23.4	76.9	123.7	135.0	29.2	128.6	20.1
Beef	43.7	23.7	14.6	**2.6**	4.8	24.0	8.0	**0.0**	50.2	60.6
Coco	7.4	17.6	24.2	10.4	65.0	40.3	83.0	77.1	0.0	**0.0**
Corn	57.7	33.0	49.8	17.2	6.0	51.0	66.4	94.5	139.5	312.4
Lamb	12.1	6.9	2.5	80.0	8.6	45.1	2.9	0.0	50.7	**0.0**
Nuts	2.3	**1.2**	**0.0**	15.2	34.2	21.0	46.4	73.1	**0.0**	**0.0**
Pork	50.7	48.6	1.1	33.1	81.9	19.9	**0.0**	**0.0**	142.0	219.8
Poul	43.1	52.8	85.5	40.3	157.4	48.1	102.5	100.7	118.6	**320.0**
Rice	49.7	86.3	27.6	78.7	**176.7**	139.2	58.5	41.4	37.8	**0.0**
Sorg	22.8	22.8	19.1	17.3	78.7	106.8	151.2	90.5	21.1	6.8
Soyb	36.6	38.9	66.5	11.9	6.2	24.2	32.0	35.1	74.0	102.5
Whea	31.0	54.9	57.6	14.4	10.7	39.7	133.3	180.9	53.1	33.5
Coff	1.5	1.9	4.9	3.3	11.8	24.6	**0.0**	**0.0**	18.6	15.2
Fish	10.3	22.4	**0.0**	41.1	48.4	2.6	**0.0**	**0.0**	1.8	8.3
Oliv	14.4	6.9	**0.0**	51.5	130.9	73.0	98.1	118.8	**0.0**	**0.0**
Palm	35.5	36.8	63.4	33.9	17.4	31.1	7.0	32.3	**187.6**	149.0
Rape	47.8	21.0	44.6	9.7	23.9	54.1	96.7	142.7	12.4	8.5
Salm	19.9	37.3	55.2	77.7	71.7	29.6	50.8	**0.0**	76.2	**0.0**
Shri	2.5	48.1	**0.0**	8.1	21.1	33.8	33.0	**0.0**	79.5	107.3
Soym	38.9	51.6	22.7	23.5	92.4	46.4	20.3	42.6	16.9	1.1
Soyo	16.3	15.8	17.1	5.9	15.1	35.4	23.9	42.1	10.9	79.3
Suga	26.7	37.1	2.8	47.5	39.1	74.6	5.6	**0.0**	**0.0**	**0.0**
Sunf	51.4	44.2	71.8	18.6	29.0	95.8	217.3	177.6	25.2	112.2
Tea	**0.8**	2.0	**0.0**	59.9	18.4	2.2	**0.0**	**0.0**	11.1	**0.0**
Cott	53.6	60.2	8.7	21.5	90.8	63.9	43.1	39.0	172.6	107.1
Dap	**64.7**	77.3	107.3	28.1	63.9	**173.5**	112.4	54.7	105.1	289.2
Hide	3.0	11.6	**0.0**	12.8	**0.0**	54.1	**0.0**	**0.0**	18.2	89.5
Logs	5.3	47.6	**0.0**	25.3	139.5	1.7	**0.0**	**0.0**	24.8	**0.0**
Pota	44.8	62.7	66.2	6.2	77.0	163.0	**272.5**	**195.2**	76.5	5.5
Rubb	10.5	16.2	11.8	26.7	28.5	67.9	133.6	**0.0**	174.3	76.4
Sawn	11.2	20.2	**0.0**	17.6	85.9	4.3	**0.0**	**0.0**	18.8	**0.0**
Urea	29.2	25.9	23.8	22.3	103.2	42.2	98.9	145.4	84.2	100.5
Wool	5.5	3.0	2.4	26.2	2.4	88.4	**0.0**	**0.0**	50.1	20.4
Avg.	37.37	72.66	68.91	21.90	87.22	55.45	55.73	20.01	96.03	101.11
Avg. Energy	25.57	36.25	45.88	25.18	73.08	42.78	53.92	50.52	32.05	49.14
Avg. Metals	33.93	38.57	27.15	30.60	65.61	52.63	63.05	55.57	63.05	82.74
Avg. Agriculture	22.18	27.08	23.55	31.72	43.27	41.93	46.06	46.33	36.68	40.06
Avg. Food	25.31	36.06	24.46	20.75	65.69	73.22	73.38	48.25	80.50	76.50
Avg. Other raw materials	46.8	91.1	87.7	28.8	67.8	40.3	61.5	32.4	110.8	259.1

properties. This additional layer consists of groups larger than small local clusters, also known as communities. Here we consider as potential communities not those possibly surfacing spontaneously from node interaction but the predetermined five groups of commodities illustrated in Section 4 (EMAFO). Specifically, one key insight is whether, and to what extent, interdependence is stronger within the five groups than between them.

Table 11 investigates this community structure across the GCN variants using two complementary indicators that emphasize the centrality of commodity groups rather than of individual nodes. The first measures the proportion of within-group connections relative to the total number of connections for each group. However, this metric may be biased by group size disparities. For instance, the energy group includes only three commodities, while the agricultural group includes thirteen. To address this limitation, we also compute the modularity indicator (as defined in Section 3.3.4), which captures the relative density of intra-group versus inter-group connections.

The key fact emerging from Table 11 is that modularity values largely differ ranging from below 1 to over 5, depending on model and sector. Metals consistently exhibit the highest incoming connectivity (e.g., 52.6 % in GCN5 – price levels) and modularity (up to 2.149 in GCN5 – price levels). Energy commodities show the highest outgoing modularity

Table 9
Node-level closeness centrality across GCN variants (maximum and minimum values by column in bold).

	Geodesic	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)	GCN1 (price first differences)	GCN2 (price first differences)	GCN3 (price first differences)	GCN4 (price first differences)	GCN5 (price first differences)
Coal	1.000	0.842	0.787	0.676	0.800	0.593	0.578	0.522	0.485	0.632	0.593
Ngas	1.000	0.800	0.762	0.686	0.762	0.600	0.558	0.462	0.453	0.600	0.533
Oil	1.000	0.842	0.727	0.615	0.787	0.615	0.658	0.578	0.500	0.578	0.527
Alum	1.000	0.738	0.676	0.593	0.738	0.552	0.571	0.533	0.462	0.545	0.475
Coba	1.000	0.738	0.706	0.640	0.667	0.578	0.585	0.495	0.429	0.615	0.578
Copp	1.000	**0.906**	0.828	0.696	0.716	0.593	0.578	0.527	0.522	0.565	0.516
Gold	1.000	0.800	0.738	0.552	0.632	0.578	0.640	0.565	0.485	**0.480**	0.511
Iron	1.000	0.828	0.706	0.640	0.774	0.623	0.640	0.516	0.490	0.533	**0.381**
Lead	1.000	0.828	0.727	0.640	0.814	0.578	0.600	0.511	0.471	0.571	0.527
Nick	1.000	0.750	0.649	0.600	0.873	0.578	0.615	0.511	0.453	0.539	0.480
Plat	1.000	0.800	0.716	0.600	0.941	0.608	0.565	0.505	0.480	0.585	0.511
Silv	1.000	0.787	0.727	0.640	0.706	0.600	0.578	0.490	0.384	0.667	0.466
Tin	1.000	0.873	0.762	0.686	0.658	0.608	0.632	0.558	0.527	0.640	0.539
Uran	1.000	0.738	0.658	0.571	0.716	0.593	0.608	0.533	0.475	0.545	0.500
Zinc	1.000	0.762	0.658	0.600	**0.585**	0.539	0.608	0.539	0.511	0.490	**1.000**
Bana	1.000	0.738	0.593	0.453	0.706	0.558	0.462	0.393	0.333	0.527	0.432
Barl	1.000	0.800	0.727	0.600	0.774	**0.686**	0.615	0.558	0.495	0.640	0.522
Beef	1.000	0.750	0.658	0.545	0.608	0.522	0.558	0.485	0.358	0.585	0.466
Coco	1.000	0.585	0.565	0.495	0.696	0.552	0.552	0.453	0.366	0.500	0.414
Corn	1.000	0.842	0.716	0.640	0.738	0.511	0.623	0.565	0.539	0.667	0.571
Lamb	1.000	0.615	0.545	0.440	0.889	0.558	0.565	0.453	0.403	0.565	0.471
Nuts	1.000	0.716	0.632	0.558	0.706	0.558	0.539	0.393	0.397	**1.000**	**1.000**
Pork	1.000	0.565	0.511	0.425	0.800	0.578	0.558	0.407	0.390	0.585	0.539
Poul	1.000	0.800	0.706	0.615	0.923	0.667	0.578	0.516	0.500	0.623	0.593
Rice	1.000	0.828	0.716	0.593	0.906	0.667	0.716	0.686	0.571	0.640	0.495
Sorg	1.000	0.828	0.706	0.640	0.738	0.623	0.640	0.558	0.495	0.552	0.490
Soyb	1.000	0.842	0.762	0.667	0.593	0.558	0.615	0.545	0.522	0.565	0.485
Whea	1.000	0.814	0.716	0.623	0.750	0.571	0.608	0.571	0.565	0.585	0.516
Coff	1.000	0.578	0.505	0.457	0.676	0.545	0.527	**1.000**	**1.000**	0.522	0.453
Fish	1.000	0.706	0.623	0.485	**0.941**	0.500	0.565	0.485	0.449	0.565	0.490
Oliv	1.000	0.696	0.558	0.505	0.800	0.545	0.623	0.527	0.527	0.516	0.425
Palm	1.000	0.857	0.774	0.686	0.857	0.608	0.593	0.505	0.449	0.658	0.585
Rape	1.000	0.857	0.716	0.593	0.696	0.552	0.615	0.571	0.558	0.565	0.500
Salm	1.000	0.696	0.608	0.516	0.873	0.608	0.608	0.533	0.403	0.578	0.471
Shri	1.000	0.608	0.527	0.407	0.738	0.565	0.571	0.466	0.407	0.632	0.608
Soym	1.000	0.873	0.716	0.558	0.774	0.571	0.640	0.545	0.516	0.522	0.397
Soyo	1.000	0.787	0.738	0.649	0.676	0.558	0.600	0.527	0.500	0.522	0.475
Suga	1.000	0.750	0.658	0.516	0.762	0.552	0.571	0.361	**0.293**	0.495	0.390
Sunf	1.000	0.828	0.774	0.667	0.787	0.623	0.696	0.658	0.585	0.585	0.527
Tea	1.000	0.615	0.539	0.393	0.828	**0.490**	0.466	**0.348**	0.302	0.527	0.471
Cott	1.000	0.727	0.667	0.545	0.787	0.593	0.565	0.500	0.414	0.600	0.558
Dap	1.000	0.889	**0.842**	**0.706**	0.787	0.558	0.716	0.640	0.552	0.649	0.608
Hide	1.000	0.632	0.565	0.462	0.923	0.545	0.527	0.364	0.327	0.676	0.571
Logs	1.000	**0.527**	**0.440**	0.348	0.814	0.640	**0.414**	0.353	**1.000**	0.552	0.387
Pota	1.000	0.828	0.774	0.686	0.696	0.578	**0.716**	0.676	0.552	0.608	0.545
Rubb	1.000	0.762	0.649	0.545	0.750	0.565	0.565	0.527	0.417	0.658	0.539
Sawn	1.000	0.565	0.485	**0.259**	0.696	0.522	0.429	**1.000**	**1.000**	0.522	0.393
Urea	1.000	0.814	0.706	0.640	0.857	0.585	0.649	0.608	0.623	0.658	0.457
Wool	1.000	0.623	0.533	0.432	0.774	0.527	0.571	0.397	**1.000**	0.565	0.500
Avg.	1.000	0.755	0.669	0.567	0.765	0.577	0.589	0.531	0.509	0.588	0.551
Avg. Energy	1.000	0.828	0.759	0.659	0.783	0.603	0.598	0.521	0.479	0.603	0.540
Avg. Metals	1.000	0.796	0.713	0.621	0.735	0.586	0.602	0.524	0.474	0.565	0.538
Avg. Agriculture	1.000	0.748	0.658	0.561	0.756	0.585	0.587	0.506	0.457	0.618	0.483
Avg. Food	1.000	0.737	0.645	0.536	0.784	0.560	0.590	0.544	0.499	0.557	0.507
Avg. Other raw materials	1.000	0.707	0.629	0.514	0.787	0.568	0.572	0.563	0.654	0.610	0.593

Table 10
Node-level eigenvector centrality across GCN variants (maximum and minimum values by column in bold).

	Geodesic	GCN1 (price levels)	GCN2 (price levels)	GCN3 (price levels)	GCN4 (price levels)	GCN5 (price levels)	GCN1 (price first differences)
Coal	0.143	0.173	0.188	0.211	0.158	0.155	0.153
Ngas	0.143	0.161	0.188	0.216	0.145	0.172	0.100
Oil	0.143	0.169	0.181	0.173	0.152	0.183	0.174
Alum	0.143	0.137	0.142	0.148	0.138	0.102	0.088
Coba	0.143	0.145	0.162	0.182	0.110	0.136	0.155
Copp	0.143	**0.183**	0.202	0.203	0.123	0.150	0.147
Gold	0.143	0.156	0.163	0.112	0.082	0.136	0.172
Iron	0.143	0.167	0.156	0.169	0.144	0.193	0.191
Lead	0.143	0.164	0.166	0.187	0.154	0.135	0.164
Nick	0.143	0.145	0.139	0.162	0.172	0.145	0.187
Plat	0.143	0.160	0.167	0.136	0.185	0.169	0.132
Silv	0.143	0.151	0.161	0.140	0.121	0.142	0.094
Tin	0.143	0.178	0.185	0.199	0.106	0.177	0.198
Uran	0.143	0.136	0.137	0.123	0.128	0.156	0.142
Zinc	0.143	0.152	0.137	0.159	**0.064**	0.087	0.172
Bana	0.143	0.123	0.075	0.024	0.118	0.115	0.029
Barl	0.143	0.159	0.174	0.184	0.149	**0.252**	0.140
Beef	0.143	0.139	0.127	0.089	0.074	0.088	0.091
Coco	0.143	0.056	0.051	0.026	0.120	0.115	0.069
Corn	0.143	0.172	0.177	0.196	0.134	0.096	0.176
Lamb	0.143	0.078	0.062	0.018	0.176	0.109	0.103
Nuts	0.143	0.133	0.111	0.070	0.123	0.103	0.048
Pork	0.143	0.046	0.017	0.010	0.159	0.131	0.062
Poul	0.143	0.162	0.155	0.149	0.183	0.228	0.129
Rice	0.143	0.165	0.158	0.143	0.177	0.225	0.216
Sorg	0.143	0.169	0.167	0.186	0.133	0.184	0.186
Soyb	0.143	0.169	0.180	0.174	0.068	0.114	0.162
Whea	0.143	0.167	0.170	0.191	0.140	0.151	0.157
Coff	0.143	0.055	0.027	0.013	0.111	0.104	0.050
Fish	0.143	0.124	0.094	0.035	**0.188**	0.059	0.107
Oliv	0.143	0.118	0.055	0.036	0.153	0.096	0.142
Palm	0.143	0.174	0.191	0.218	0.170	0.180	0.151
Rape	0.143	0.175	0.169	0.171	0.119	0.111	0.179
Salm	0.143	0.116	0.097	0.050	0.174	0.175	0.148
Shri	0.143	0.081	0.031	0.009	0.136	0.124	0.097
Soym	0.143	0.174	0.154	0.107	0.150	0.126	0.184
Soyo	0.143	0.160	0.181	0.194	0.116	0.111	0.154
Suga	0.143	0.139	0.120	0.042	0.143	0.102	0.063
Sunf	0.143	0.169	0.196	0.209	0.152	0.186	0.213
Tea	0.143	0.081	0.039	0.006	0.161	**0.046**	0.022
Cott	0.143	0.132	0.143	0.108	0.152	0.161	0.099
Dap	0.143	0.181	**0.204**	**0.222**	0.153	0.113	**0.246**
Hide	0.143	0.086	0.064	0.030	0.185	0.110	0.041
Logs	0.143	**0.019**	**0.008**	0.002	0.157	0.180	**0.011**
Pota	0.143	0.168	0.189	0.209	0.122	0.127	0.239
Rubb	0.143	0.138	0.116	0.098	0.137	0.109	0.078
Sawn	0.143	0.045	0.018	**0.000**	0.116	0.071	0.014
Urea	0.143	0.166	0.171	0.198	0.170	0.151	0.200
Wool	0.143	0.082	0.041	0.023	0.148	0.081	0.081
Avg.	0.143	0.14	0.13	0.12	0.14	0.14	0.13
Avg. Energy	0.143	0.17	0.19	0.20	0.15	0.17	0.14
Avg. Metals	0.143	0.16	0.16	0.16	0.13	0.14	0.15
Avg. Agriculture	0.143	0.13	0.12	0.11	0.13	0.15	0.12
Avg. Food	0.143	0.13	0.11	0.09	0.15	0.12	0.13
Avg. Other raw materials	0.143	0.11	0.11	0.10	0.15	0.12	0.11

in sparse networks (up to 5.444 in GCN2 – first differences), indicating their role as cross-sector influencers. Agriculture and food sectors maintain moderate modularity and connectivity, while other raw materials display more variability, with outgoing modularity peaking at 2.420 in GCN5 (first differences).

The combination of these two indicators reveals substantial differences between GCNs based on price levels and those based on first differences of prices. In the former case, metals tend to form a strong community in terms of incoming connections, while it is significantly weaker for outgoing linkages. This suggests that metals play a key role in the network, as they tend to amplify incoming shocks within their group and subsequently transmit them to the rest of the network. In contrast, energy commodities appear to exhibit the opposite behaviour. In this case community linkages are relatively strong for outgoing connections and weaker for incoming ones. For food commodities, community linkages are weak in both directions, whereas the evidence is more mixed and, in any case, the community structure is weaker for agricultural products and other raw materials.

When GCNs based on the first differences of prices are considered, the results show notable deviations from the previous picture. The modularity of metals weakens, while it increases, both for incoming and outgoing linkages, in the case of energy commodities, which emerge as the strongest nucleolus within the network. For other raw materials, within-group connections become significantly more relevant, at least when the sparser networks (GCN4 and GCN5) are taken into account.

Table 11
Community structure across GCN variants: within-group connectivity and modularity (maximum and minimum values by row in bold).

	Energy	Metals	Agriculture	Food	Other raw materials
GCN1 (price levels)					
Incoming – Within (%)	**5.2**	**33.3**	24.0	17.2	15.4
Outgoing – Within (%)	**7.0**	18.3	**27.2**	23.3	19.0
Incoming – Modularity	0.833	**1.361**	0.904	**0.700**	0.838
Outgoing – Modularity	**1.150**	0.746	1.027	0.950	1.037
GCN2 (price levels)					
Incoming – Within (%)	**4.4**	**37.9**	20.3	17.2	14.2
Outgoing – Within (%)	**8.5**	18.0	**24.4**	22.6	18.6
Incoming – Modularity	0.705	**1.546**	0.765	**0.703**	0.772
Outgoing – Modularity	**1.390**	**0.736**	0.920	0.924	1.014
GCN3 (price levels)					
Incoming – Within (%)	**5.6**	**43.8**	16.2	16.7	9.9
Outgoing – Within (%)	14.3	**13.7**	**25.8**	18.5	20.4
Incoming – Modularity	0.902	**1.790**	0.611	0.681	**0.539**
Outgoing – Modularity	**2.333**	**0.558**	0.973	0.755	1.111
GCN4 (price levels)					
Incoming – Within (%)	**7.8**	26.5	27.4	**27.8**	18.1
Outgoing – Within (%)	**7.0**	23.4	**30.0**	23.5	23.6
Incoming – Modularity	**1.299**	1.084	1.034	1.133	0.988
Outgoing – Modularity	1.150	**0.957**	1.132	0.961	**1.286**
GCN5 (price levels)					
Incoming – Within (%)	**11.8**	**52.6**	27.2	24.2	19.7
Outgoing – Within (%)	19.0	34.8	**44.2**	**17.6**	20.0
Incoming – Modularity	2.044	**2.149**	1.025	**0.988**	1.071
Outgoing – Modularity	**3.111**	1.420	1.664	**0.721**	1.089
GCN1 (price first differences)					
Incoming – Within (%)	**11.4**	25.2	25.0	**25.9**	17.9
Outgoing – Within (%)	**13.8**	**27.3**	23.8	26.4	15.6
Incoming – Modularity	**1.978**	1.029	**0.942**	1.057	0.977
Outgoing – Modularity	**2.253**	1.114	0.896	1.077	**0.847**
GCN2 (price first differences)					
Incoming – Within (%)	18.2	21.8	**27.6**	23.7	**17.6**
Outgoing – Within (%)	33.3	**36.2**	22.9	23.0	**9.8**
Incoming – Modularity	**3.407**	**0.890**	1.040	0.969	0.961
Outgoing – Modularity	**5.444**	1.477	0.862	0.937	**0.536**
GCN3 (price first differences)					
Incoming – Within (%)	14.3	27.5	**35.0**	25.0	**13.0**
Outgoing – Within (%)	28.6	**36.8**	26.4	26.8	**9.1**
Incoming – Modularity	**2.556**	1.121	1.319	1.021	**0.710**
Outgoing – Modularity	**4.667**	1.504	0.996	1.096	**0.495**
GCN4 (price first differences)					
Incoming – Within (%)	**7.1**	25.5	29.0	18.3	19.5
Outgoing – Within (%)	**7.4**	14.4	29.0	14.3	40.0
Incoming – Modularity	**1.179**	1.039	1.094	**0.749**	1.060
Outgoing – Modularity	1.210	0.589	1.094	**0.583**	**2.178**
GCN5 (price first differences)					
Incoming – Within (%)	**11.1**	22.6	**30.2**	14.6	27.6
Outgoing – Within (%)	**11.8**	11.9	34.0	14.3	**44.4**
Incoming – Modularity	**1.917**	0.922	1.138	**0.598**	1.502
Outgoing – Modularity	1.922	**0.484**	1.283	0.583	**2.420**

5.4. Robustness check: comparison with transformed series

A final robustness check of the results can be performed by repeating the analysis with a logarithmic transformation of the series. As discussed in Section 2, a major challenge in studying commodity price interdependence is the presence, often temporary, of nonlinear dynamics and linkages. Failing to account for these nonlinearities may lead to an inappropriate or incomplete reconstruction of the interdependence network. Since logarithmic transformations can capture certain nonlinear patterns, it is worth assessing whether using these transformed series leads to substantially different network structures. Table 12 presents descriptive evidence on the structural properties of selected GCNs based on the logarithm of prices (hereafter referred to as transformed GCNs). Due to space limitations, only the sparser networks

Table 12
Main features of the GCN variants obtained with the logarithm of price levels.

	GCN3 (logarithms of price levels)	GCN5 (logarithms of price levels)
Pairwise Pearson correlation with other GCNs:		
GCN1 (price levels)	0.402**	0.149
GCN2 (price levels)	0.480**	0.151
GCN3 (price levels)	0.544**	0.158
GCN4 (price levels)	0.099	0.251*
GCN5 (price levels)	0.188	0.286**
GCN1 (price first differences)	0.112	0.033
GCN2 (price first differences)	0.087	0.015
GCN3 (price first differences)	0.118	0.019
GCN4 (price first differences)	0.218	0.063
GCN5 (price first differences)	0.209	0.091
Topology and centrality measures:		
Arcs	300	340
Density	0.127	0.145
Paths (largest component)	2256	2352
Diameter (largest component)	5	3
Average shortest path (largest component)	1.994	1.749
Reciprocity	0.047	0.083
Transitivity	1.133	0.573
Betweenness centralization	0.141	0.247
Indegree centralization	0.316	0.235
Outdegree centralization	0.401	0.661

*,** Statistically significant at 10 % and 5 % confidence level, respectively.

(GCN3 and GCN5) are considered, as these are more likely to reveal structural diversity. For the same reason, networks based on the first differences of the logarithm of prices are not reported. In any case, they lead to largely consistent conclusions.[16]

The upper part Table 12 reports the correlation coefficient between the untransformed GCNs and the transformed GCN3 and GCN5 networks. The key finding concerns the positive correlation between the transformed GCNs and all other GCNs considered thus far. In fact, the correlation is stronger and statistically significant only when the same underlying VAR model and the same data series (namely, price levels) are used. Specifically, a high positive correlation is observed between the transformed GCN3 and the untransformed GCN1 (0.402), GCN2 (0.480), and GCN3 (0.544), as well as between the transformed GCN5 and the untransformed GCN4 (0.251) and GCN5 (0.286). By contrast, no significant correlation is found between the two transformed GCNs and any of the untransformed GCNs based on the first differences of prices, suggesting that first differencing alters the network properties more substantially than the logarithmic transformation.

The lower part of Table 12 also presents some network-level metrics for the two transformed GCNs. The key evidence is that, when compared to their untransformed counterparts, GCN3 and GCN5 networks exhibit broadly similar properties. Nonetheless, the transformed networks appear slightly less dense and more granular, with a modest increase in the peripherality of some nodes. Transformed GCN3 shows higher transitivity (1.133), suggesting stronger local clustering. Transformed GCN5 exhibits higher betweenness (0.247) and outdegree centralization (0.661), indicating more hierarchical structure and directional influence. The transformed GCN5's diameter is also smaller (3 vs. 5), and its average path length is shorter (1.749 vs. 1.994), implying greater efficiency in price signal transmission.

The motivations and implications of these slight differences, potentially associated with the logarithmic transformation and, by extension, with the presence of nonlinear dynamics, may warrant further investigation in future research. Despite these differences, however, the findings confirm that logarithmic transformation preserves the core features of the network, thereby supporting the structural stability of the networks under potential nonlinear dynamics.

6. Policy implications

As discussed in Section 2.3, the volatility of resource and commodity prices poses a significant threat to economic and social systems. These fluctuations often reverberate through the broader economy, fueling inflation and jeopardizing macroeconomic stability. The strategic importance of certain commodities, coupled with growing market interdependencies, reinforces the urgency for policy-makers to adopt real-time monitoring and early-warning systems. Such instruments should be designed to detect price movements that could generate systemic risks. While forecasting commodity price turbulence remains a complex endeavour, continuous observation, supported by sound analytical frameworks, can yield critical insights for anticipatory governance and risk mitigation.

The analysis here proposed and the results obtained aim to contribute in this direction. They suggest that the commodity price network framework can be particularly helpful to assess how shocks in one or a few commodities may propagate across the broader market structure. By mapping the interdependencies among commodities, the network logic can also highlight those prices behaving like "canaries in a coal mine" by signalling the onset of generalized commodity price shocks and, consequently, major inflationary pressures. This perspective seems essential for shifting policy-making from fragmented responses toward a more integrated approach that not only mitigates shocks but also fosters systemic resilience.

In this context, the main policy implication of the present study consists in adopting a network-based logic to develop early warning and monitoring systems. These systems should be capable of tracking real-time price movements across interconnected commodities and generating timely alerts. The analytical strength of the commodity price network resides in its ability to uncover the transmission pathways of shocks, enabling policymakers to intervene before disturbances escalate into systemic crises.

The results produced in this study about such networks are particularly insightful, revealing two complementary findings. First, price interdependence is pervasive throughout the network, extending well beyond clusters of homogeneous commodities. Second, within these clusters, certain leader nodes emerge, acting as both generators and propagators of shocks, within their own group and across others. Notably, some nodes and groups occupy a more central position in the overall network structure, with metals standing out more prominently than energy commodities.

However, our empirical analysis also highlights that building such a tool requires several methodological steps. First, the network must be constructed by choosing among alternative topologies, sparse or dense, and considering the temporal dimension of the data, distinguishing between long-term price levels and short-term fluctuations. The current evidence suggests that a sparse network based on long-term price levels may yield the most meaningful insights. Second, once the network is established, it becomes crucial to identify key nodes and representative commodities within each group that can serve as sentinel indicators. Monitoring these nodes enables the quantification of shock transmission and the detection of early signs of systemic risk. Third, clear criteria must be defined to determine when a warning should be triggered, based on the behavior of these sentinel nodes. Finally, the network should undergo regular updates and refinements, integrating advanced analytical techniques and adjusting the commodity set to align with shifting market dynamics.

Although a detailed analysis of the fabric of policy action lies beyond the scope of this study, some suggestions regarding the operational use of the network-based tool can still be offered based on these conclusions. This use can move in multiple directions. It can support the deployment of targeted stabilization mechanisms, such as buffer stocks, subsidies, or strategic reserves, particularly when shocks are detected in critical nodes like energy or food commodities. These interventions should be directed toward the most exposed segments of the network, thereby minimizing the transmission of volatility. Moreover, the systemic nature of price propagation calls for cross-sectoral coordination, as shocks often traverse boundaries, from energy to transport, agriculture, and beyond, necessitating inter-sectoral collaboration and integrated policy responses. In addition to immediate interventions, the network can guide long-term resilience-oriented investments, either public or private. Insights derived from the network can help prioritize investments in infrastructure and systems that enhance the robustness of critical nodes, such as renewable energy or local food supply chains. Furthermore, the network can inform strategic communication efforts, equipping economic agents with the knowledge needed to anticipate and adapt to price changes.

7. Concluding remarks

This study explores the interdependence within and across different groups of commodity prices, addressing a central methodological question: Can a large-scale commodity price network be empirically constructed, and if so, how? What insights can such a network offer about price interdependence? These questions become particularly challenging when the number of commodities under consideration is large, raising both modeling and computational concerns. Empirically, our findings align with international studies that investigate commodity price co-movements and contagion effects. However, due to the dimensionality issue, most of these studies focus on specific contexts or

[16] These further transformed GCN variants are available upon request.

subsets of commodities (Larrosa et al., 2024), raising concerns about the generalizability of both the proposed methodologies and the resulting insights.

To tackle the dimensionality challenge, this study proposes a methodological framework based on Granger causality and subsequent network analysis. Specifically, a three-stage approach is introduced to generate multiple variants of the commodity price network. Applying this framework to a dataset of 49 commodities observed over 540 months (from 1980 to 2024) enables the identification and examination of the main structural features of this complex interdependence.

The study aims to advance both the theoretical and methodological understanding of commodity price interdependence. Theoretically, the results contribute to the notion that commodity markets function as interconnected systems, where localized shocks can propagate and amplify through structural channels. Methodologically, the proposed approach offers a flexible analytical framework with practical relevance for policymakers and market participants operating in an era of increasing market volatility and globally interconnected commodity markets.

However, the framework generates different network variants depending on the choice of VAR models and data processing techniques. On one hand, this diversity allows for an assessment of the robustness of key network features across alternative specifications. On the other hand, the results reveal that different modeling choices can significantly alter the network's structure and properties. The findings suggest a significant discrepancy between short-term and long-term connections, which are differently represented across the network derivations. This underscores the possibility that short-term volatility and long-term structural interdependence operate through distinct mechanisms.

Beyond the general interest and scope of the proposed analysis, the results raise concerns regarding the robustness and generalizability of the adopted methodology. The reliance on linear Granger causality may overlook nonlinear dynamics, while the static nature of the network may fail to capture evolving interdependencies over time. In this regard, a number of enhancements may be proposed, both to strengthen the methodological framework and to facilitate its comparison with alternative models.

Regarding the first aspect, the sensitivity of network topology to data transformation, model specification, and estimation strategy suggests further research to enhance the robustness and precision of VAR-based network derivations. This would help identify the best-performing configuration to be considered as a standard or benchmark. A deeper investigation into the most appropriate empirical construction of the commodity price network naturally leads to a second set of methodological enhancements. The study adopts the VAR framework due to its simplicity, interpretability, and ability to address dimensionality issues. Accordingly, the two model specifications employed, PW-VAR and S-VAR, are kept deliberately simple.

Future improvements to the PW-VAR model could involve explicitly accounting for the timing of Granger causality, i.e., identifying which lags contribute to causal relationships. This would enable a transition from a static network to a more sophisticated weighted and dynamic network (Diebold and Yılmaz, 2014). In the case of the S-VAR model, further refinement could be achieved by adopting more advanced sparsity-inducing techniques, such as adaptive LASSO (Takada and Fujisawa, 2023), to better define the sparsity of the adjacency matrix by also combining the PW-VAR and S-VAR approaches. Additionally, non-VAR-based approaches might be considered. If the dimensionality challenge can be addressed, methods such as wavelet analysis or machine learning could offer improved identification of connections under nonlinear multivariate dynamics.

Finally, additional refinements could be achieved through alternative data transformations. For example, converting prices to real terms by accurately accounting for inflation may offer deeper insights into the robustness of the findings. Moreover, expanding the dataset to include a broader range of commodities, especially those more recently gaining economic relevance, and extending the observation period could significantly improve the generalizability and analytical depth of the study.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author used Microsoft Copilot (GPT-4) in order to improve the readability and language of the manuscript. After using this tool, the author reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resourpol.2025.105820.

Data availability

The authors do not have permission to share data.

References

Acemoglu, D., Carvalho, V.M., Ozdaglar, A., Tahbaz-Salehi, A., 2012. The network origins of aggregate fluctuations. Econometrica 80 (5), 1977–2016.

Ahelegbey, D.F., Billio, M., Casarin, R., 2016. Sparse graphical vector autoregression: a bayesian approach. Annals of Economics and Statistics/Annales d' Economie et de Statistique 333–361.

Ahelegbey, D.F., Giudici, P., Hashem, S.Q., 2021. Network VAR models to measure financial contagion. N. Am. J. Econ. Finance 55 (C), 101318.

Amaglobeli, D., Hanedar, E., Hong, G.E., Thévenot, C., 2022. Fiscal Policy for Mitigating the Social Impact of High Energy and Food Prices. IMF Notes No 2022/001. International Monetary Fund, Washington.

Ameur, H., Boubaker, S., Ftiti, Z., Louhichi, W., Tissaoui, K., 2024. Forecasting commodity prices: empirical evidence using deep learning tools. Ann. Oper. Res. 339, 349–367.

Baffes, J., Nagle, P. (Eds.), 2022. Commodity Markets: Evolution, Challenges, and Policies. World Bank, Washington, DC.

Banbura, M., Giannone, D., Reichlin, L., 2010. Large Bayesian vector auto regressions. J. Appl. Econom. 25 (1), 71–92.

Basu, S., Michailidis, G., 2015. Regularized estimation in sparse high-dimensional time series models. Ann. Stat. 43 (4), 1535–1567.

Baum, C.F., Hurn, S., Otero, J., 2023. The dynamics of U.S. industrial production: a time-varying granger causality perspective. Econometrics and Statistics. https://doi.org/10.1016/j.ecosta.2021.10.012 (in press).

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B 57 (1), 289–300.

Bernanke, B.S., Boivin, J., Eliasz, P., 2005. Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach. Q. J. Econ. 120 (1), 387–422.

Boako, G., Alagidede, I.P., Sjo, B., Uddin, G.S., 2020. Commodities price cycles and their interdependence with equity markets. Energy Econ. 91, 104884.

Boakye, E.O., Heimonen, K., Junttila, J., 2024. Commodity markets and the global macroeconomy: evidence from machine learning and GVAR. Empir. Econ. 67, 1919–1965.

Borgatti, S.P., Everett, M.G., Johnson, J.C., 2018. Analyzing Social Networks, second ed. SAGE Publications, London.

Breheny, P., Huang, J., 2011. Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. Ann. Appl. Stat. 5 (1), 232–253.

Bredenkamp, H., Bersch, J., 2012. Commodity price volatility: impact and policy challenges for low-income countries. In: Arezki, R., Pattillo, C.A., Quintyn, M.G. (Eds.), Commodity Price Volatility and Inclusive Growth in Low-Income Countries. International Monetary Fund, Washington, pp. 55–67.

Byrne, J.P., Sakemoto, R., Xu, B., 2020. Commodity price co-movement: heterogeneity and the time-varying impact of fundamentals. Eur. Rev. Agric. Econ. 47 (2), 499–528.

Carlos-Sandberg, L., Clack, C.D., 2021. Incorporation of causality structures to complex network analysis of time-varying behaviour of multivariate time series. Sci. Rep. 11, 18880. https://doi.org/10.1038/s41598-021-97741-2.

Chiou-Wei, S.Z., Chen, C.-F., Zhu, Z., 2008. Economic growth and energy consumption revisited - evidence from linear and nonlinear granger causality. Energy Econ. 30 (6), 3063–3076.
Chudik, A., Pesaran, M.H., 2011. Infinite-dimensional vars and factor models. J. Econom. 163 (1), 4–22.
Diebold, F.X., Liu, L., Yilmaz, K., 2017. Commodity Connectedness. NBER Working Paper No. 23685, Washington. http://www.nber.org/papers/w23685.
Diebold, F.X., Yılmaz, K., 2014. On the network topology of variance decompositions: measuring the connectedness of financial firms. J. Econom. 182 (1), 119–134.
Ding, S., Cui, T., Zheng, D., Du, M., 2021. The effects of commodity financialization on commodity market volatility. Resour. Policy 73, 102220.
Drachal, K., Pawłowski, M., 2024. Forecasting selected commodities' prices with the Bayesian symbolic regression. Int. J. Financ. Stud. 12 (2), 34. https://doi.org/10.3390/ijfs12020034.
Dufour, J.-M., Renault, E., 1998. Short-run and long-run causality in time series: theory. Econometrica 66, 1099–1125.
Dufour, J.-M., Taamouti, A., 2010. Short and long run causality measures: theory and inference. J. Econom. 154, 42–58.
Eichler, M., 2012. Causal inference in time series analysis. In: Berzuini, C., Dawid, P., Bernardinelli, L. (Eds.), Causality: Statistical Perspectives and Applications. John Wiley & Sons, Hoboken (NJ), pp. 327–354.
Enders, W., 2014. Applied Econometric Time Series, fourth ed. John Wiley & Sons, Hoboken (NJ).
Esposti, R., 2021. On the long-term common movement of resource and commodity prices. A methodological proposal. Resour. Policy 72, 102010. https://doi.org/10.1016/j.resourpol.2021.102010.
Esposti, R., 2024a. Dating common commodity price and inflation shocks with alternative approaches. Bio base Appl. Econ. 13 (2), 171–201. https://doi.org/10.36253/bae-14060.
Esposti, R., 2024b. Who moves first? Resource price interdependence through time-varying granger causality. Nat. Resour. Model. 37 (3), e12396. https://doi.org/10.1111/nrm.12396.
Esposti, R., Listorti, G., 2013. Agricultural price transmission across space and commodities during price bubbles. Agric. Econ. 44 (1), 125–139.
European Commission, 2024. Regulation (EU) 2024/1252 of the European parliament and of the council of 11 April 2024 establishing a framework for ensuring a secure and sustainable supply of critical raw materials and amending regulations (EU) no 168/2013, (EU) 2018/2858, (EU) 2018/1724 and (EU) 2019/1020. Official Journal of the European Union, L, 3 May 2024. https://eur-lex.europa.eu/eli/reg/2024/1252/oj/eng.
Fry-McKibbin, R., Greenwood-Nimmo, M., Lin, Q., 2023. Commodity price cycles and the interdependence of commodity and equity markets. Available at: SSRN: https://ssrn.com/abstract=4557222 https://doi.org/10.2139/ssrn.4557222.
Garzón, A.J., Hierro, L.A., 2022. Inflation, oil prices and exchange rates. The Euro's dampening effect. J. Pol. Model. 44 (1), 130–146.
George, E.I., Sun, D., Ni, S., 2008. Bayesian stochastic search for Var model restrictions. J. Econom. 142 (1), 553–580.
Ghahramani, Z., 1998. Learning dynamic Bayesian networks. In: Giles, C.L., Gori, M. (Eds.), *Adaptive Processing of Sequences and Data Structures* (Lecture Notes in Computer Science, vol. 1387. Springer, Berlin, pp. 168–197.
Gregorio, J., 2012. Commodity prices, monetary policy, and inflation. IMF Econ. Rev. 60, 600–633.
Hong, Y., Liu, Y., Wang, S., 2009. Granger causality in risk and detection of extreme risk spillover between financial markets. J. Econom. 150 (2), 271–287.
Kirikkaleli, D., Güngör, H., 2021. Co-movement of commodity price indexes and energy price index: a wavelet coherence approach. Financ. Innov. 7, 15. https://doi.org/10.1186/s40854-021-00230-8.
Kozian, L.L., Machado, M.R., Osterrieder, J.R., 2025. Modeling commodity price co-movement: building on traditional time series models and exploring applications of machine learning algorithms. Decis. Econ. Finance. https://doi.org/10.1007/s10203-025-00512-1.
Larrosa, J.M.C., Gutiérrez, E.M., Uriarte, J.I., Ramírez Muñoz de Toro, G.R., 2024. Granger causality networks of price leadership in the retail tea market of Argentina. J. Revenue Pricing Manag. 1–10. https://doi.org/10.1057/s41272-024-00480-y.
Leeper, E.M., Sims, C.A., Zha, T., Hall, R.E., Bernanke, B.S., 1996. What does monetary policy do? Brookings Pap. Econ. Activ. 2, 1–78.
Listorti, G., Esposti, R., 2012. Horizontal price transmission in agricultural markets: fundamental concepts and open empirical issues. Bio base Appl. Econ. 1 (1), 81–108.
Litterman, R.B., 1986. Forecasting with Bayesian vector autoregressions-five years of experience. J. Bus. Econ. Stat. 4 (1), 25–38.
Lozano, A.C., Abe, N., Liu, Y., Rosset, S., 2009. Grouped graphical granger modeling for gene expression regulatory networks discovery. Bioinformatics 25 (12), i110–i118.
Lütkepohl, H., 1982. Non-causality due to omitted variables. J. Econom. 19 (2–3), 367–378.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer, Heidelberg.
Marra, A., Cucculelli, M., Cartone, A., 2024. So far, yet so close. Using networks of words to measure proximity and spillovers between firms. Eurasian Business Review 14 (4), 973–1000.
Mastroeni, L., Mazzoccoli, A., Quaresima, G., Vellucci, P., 2022. Wavelet analysis and energy-based measures for oil-food price relationship as a footprint of financialisation effect. Resour. Policy 77, 102692.
Meng, H., Xie, W.J., Jiang, Z.Q., Podobnik, B., Zhou, W.X., Stanley, H.E., 2014. Systemic risk and spatiotemporal dynamics of the US housing market. Sci. Rep. 4, 3655.
Morana, C., 2012. PC-VAR estimation of vector autoregressive models. Open J. Stat. 2, 251–259.
Mosedale, T.J., Stephenson, D.B., Collins, M., Mills, T.C., 2006. Granger causality of coupled climate processes: ocean feedback on the North Atlantic oscillation. J. Clim. 19 (7), 1182–1194.
Muflikh, Y., Smith, C., Brown, C., Aziz, A., 2021. Analysing price volatility in agricultural value chains using systems thinking: a case study of the Indonesian chilli value chain. Agric. Syst. 192, 103179.
Mutascu, M.I., Albulescu, C.T., Apergis, N., Magazzino, C., 2022. Do gasoline and diesel prices co-move? Evidence from the time-frequency domain. Environ. Sci. Pollut. Control Ser. 29, 68776–68795.
Newman, M.E.J., 2010. Networks: an Introduction. Oxford University Press, Oxford.
Nigatu, G., Adjemian, M., 2020. A wavelet analysis of price integration in major agricultural markets. J. Agric. Appl. Econ. 52 (1), 117–134.
OECD, 2010. Developments in commodity price volatility. Working Party on Agricultural Policies and Markets, Trade and Agriculture Directorate. OECD, Paris.
Piot-Lepetit, I., M'Barek, R. (Eds.), 2011. Methods to Analyse Agricultural Commodity Price Volatility. Springer, New York.
Quintino, D., Telo da Gama, J., Ferreira, P., 2021. Cross-correlations in meat prices in Brazil: a non-linear approach using different time scales. Economies 9 (4), 133. https://doi.org/10.3390/economies9040133.
Runge, J., 2018. Causal network reconstruction from time series: from theoretical assumptions to practical estimation. Chaos: An Interdisciplinary Journal of Nonlinear Science 28 (7), 075310.
Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A., White, D.R., 2009. Economic networks: the new challenges. Science 325, 422–425.
Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger causality analysis in neuroscience and neuroimaging. J. Neurosci. 35 (8), 3293–3297.
Shahzad, F., Bouri, E., Mokni, K., Ajmi, A.N., 2021. Energy, agriculture, and precious metals: evidence from time-varying Granger causal relationships for both return and volatility. Resour. Policy 74, 102298.
Shi, S., Hurn, S., Phillips, P.C.B., 2020. Causal change detection in possibly integrated systems: revisiting the money-income relationship. J. Financ. Econom. 18 (1), 158–180.
Shi, S., Phillips, P.C.B., Hurn, S., 2018. Change detection and the causal impact of the yield curve. J. Time Anal. 39 (6), 966–987.
Shojaie, A., Fox, E.B., 2021. Granger causality: a review and recent advances. arXiv: 2105.02675. Available at: https://arxiv.org/abs/2105.02675.
Signoretto, M., Suykens, J.A.K., 2015. Kernel methods. In: Kacprzyk, J., Pedrycz, W. (Eds.), Springer Handbook of Computational Intelligence. Springer, Heidelberg, pp. 577–605.
Stock, J.H., Watson, M.W., 2002. Forecasting using principal components from a large number of predictors. J. Am. Stat. Assoc. 97 (460), 1167–1179.
Sun, Q., Gao, X., Wen, S., Chen, Z., Hao, X., 2018. The transmission of fluctuation among price indices based on Granger causality network. Phys. Stat. Mech. Appl. 506, 36–49.
Takada, M., Fujisawa, H., 2023. Adaptive lasso, transfer lasso, and beyond: an asymptotic perspective. arXiv preprint arXiv:2308.15838. https://doi.org/10.48550/arXiv.2308.15838.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. Roy. Stat. Soc. B 58 (1), 267–288.
Uematsu, Y., Yamagata, T., 2025. Discovering the network granger causality in large vector autoregressive models. J. Am. Stat. Assoc. https://doi.org/10.1080/01621459.2025.2450836.
UNCTAD, 2024. Trade and Development Report 2023: Growth, Debt and Climate – Realigning the Global Financial Architecture. United Nations Conference on Trade and Development, Geneva: United Nations.
van Garderen, K.J., 2023. Forecasting levels in Loglinear unit root models. Econom. Rev. 42 (9–10), 780–805.
Wang, G.J., Si, H.B., Chen, Y.Y., Xie, C., Chevallier, J., 2021. Time domain and frequency domain Granger causality networks: application to China's financial institutions. Finance Res. Lett. 39, 101662.
Watts, D.J., 2004. The 'New' science of networks. Annu. Rev. Sociol. 30, 243–270.
World Bank, 2025. Commodity Markets Outlook, April 2025. World Bank, Washington, DC.
Xavier, A., Fernandes, B., De Oliveira, J., 2023. A hybrid swarm-based system for commodity price forecasting during the Covid-19 pandemic. IEEE Access 11, 74379–74387. https://doi.org/10.1109/access.2023.3293738.
Yuan, M., Lin, Y., 2006. Model selection and estimation in regression with grouped variables. J. Roy. Stat. Soc. B 68 (1), 49–67.
Zelingher, R., 2024. Democratising agricultural commodity price forecasting: the AGRICAF approach. arXiv:2410.20363v5. (Accessed 2 March 2025).
Zhang, D., Broadstock, D.C., 2020. Global financial crisis and rising connectedness in the international commodity markets. Int. Rev. Financ. Anal. 68, 101239.
Zhou, X., Zheng, S., Zhang, H., Liu, Q., Xing, W., Li, X., Han, Y., Zhao, P., 2022. Risk transmission of trade price fluctuations from a nickel chain perspective: based on systematic risk entropy and granger causality networks. Entropy 24, 1221. https://doi.org/10.3390/e24091221.