UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

INVESTIGATING COMMODITY PRICE INTERDEPENDENCE WITH GRANGER CAUSALITY NETWORKS.

ROBERTO ESPOSTI

QUADERNO DI RICERCA n. 498

ISSN: 2279-9575

July 2025

The views expressed in the articles are those of the authors and do not involve the responsibility of the Department.

Abstract

This paper investigates the interdependence among prices in the commodity and natural resource market segment. The analysis is performed using a large dataset made of about 50 commodity prices observed with monthly frequency over a period of almost half a century (1980-2024). These different commodities are clustered in five groups (energy, metals, agriculture, food, other raw materials) in order to discriminate the interdependence within and between groups. The adopted method consists in building a Commodity Price Network (CPN) defined via Granger causality tests. These tests are performed with two alternative empirical strategies: pairwise VAR models estimation (pairwise Granger Causality) and sparse VAR models estimation (sparse VAR Granger Causality). Both price levels and price first differences are considered in order to take the possible non-stationarity or price series into account. Network analysis is performed on the different networks obtained using these alternative series and modelling approaches. Results suggest relevant differences across series and methods but some solid results also emerges, particularly pointing to a generalized interdependence that still assigns a central role to some metals and agricultural products.

Keywords: Commodity Prices, Price Interdependence, Granger Causality, Network Analysis, Sparse VAR Models.

JEL Classification: C32, Q02, O13

Address:

Roberto ESPOSTI
Department of Economics and Social Sciences - Università Politecnica delle Marche, 60121, Ancona (Italy). Email: r.esposti@staff.univpm.it

Investigating Commodity Price Interdependence with Granger Causality Networks.

Roberto Esposti*
*Department of Economics and Social Sciences - Università Politecnica delle Marche (Ancona, Italy); email: r.esposti@staff.univpm.it

ORCID ID: 0000-0002-1656-0331

Abstract

This paper investigates the interdependence among prices in the commodity and natural resource market segments. The analysis is performed using a large dataset comprising approximately 50 commodity prices observed at a monthly frequency over a period of almost half a century (1980-2024). These different commodities are clustered in five groups (energy, metals, agriculture, food, and other raw materials) in order to discriminate the interdependence within and between groups. The adopted method consists in building a Commodity Price Network (CPN) defined via Granger causality tests. These tests are performed using two alternative empirical strategies: pairwise VAR model estimation (pairwise Granger Causality) and sparse VAR model estimation (sparse VAR Granger Causality). Both price levels and price first differences are considered in order to take the possible non-stationarity or price series into account. Network analysis is performed on the different networks obtained using these alternative series and modelling approaches. Results suggest relevant differences across series and methods but some solid results also emerge, particularly pointing to a generalized interdependence that still assigns a central role to some metals and agricultural products.

Keywords: Commodity Prices, Price Interdependence, Granger Causality, Network Analysis, Sparse VAR Models.

JEL Classification: C32, Q02, O13

1. Introduction

The magnitude and direction of interdependence (or connectedness) among commodity prices is a research field that received considerable attention over the past few decades (Boako et al., 2020; Esposti, 2021, 2024a,b; Fry-McKibbin et al., 2023; Kozian et al., 2025, just to mention a few recent studies). The relevance of this subject is twofold. From the real economy perspective, since most of these commodities enter many production processes and at different stages of the supply chains, such interdependence affects the linkage across sectors and this makes price shocks transmit along supply chains finally spreading through the whole economy. Understanding the interdependence across commodity prices thus allows to assess how individual price shocks may transmit over the whole economy making it more vulnerable and more exposed to inflation risks and business cycle downturns**.** In addition, several countries regard some of these commodities as critical or strategic for their own national security (European Union, 2024), which highlights their vulnerability to shocks and fluctuations in the respective prices. From the point of view of the financial markets, understanding the integration and connectedness among commodity prices is essential for risk management and portfolio allocation and this may contribute, in turn, to global financial stability (Boako et al., 2020; Ding et al., 2021).

Besides its relevance, however, this topic has attracted the attention of researchers also because it is methodologically challenging. In general terms, commodities may be connected either directly along the same supply chains (or networks) or indirectly via other commodities or through economy-wide (or system) linkages. Therefore, interdependence may also occur among very different commodities. Some can behave as pivotal/hubs/central nodes since they affect many others but are not affected by any or very few. On the contrary, some commodities are peripheral or marginal: affected by few, not affecting any. So, the number of commodities involved in this interdependence may be large and the relationship may be complex.
The quantification and qualification of this interdependence within such complexity is the main focus of the present paper. Two methodological issues, in particular, seem relevant here. One problem concerns the dimensionality, that is, the number of commodities involved within the analysis. The typical investigation of interdependence is based on estimation VAR models and consequent Granger causality testing. But large VAR models may raise problems both from the computational point of view and in terms of result interpretability. Therefore, dimensionality is often reduced ex-ante by limiting the analysis to a subgroup of commodities (selected by homogeneity or relevance) or by relying on some rank reduction technique (jut to mention a few and depending on the data under analysis: Principal Component Analysis, PCA; latent Factor Analysis, FA; Cointegration; Machine Learning techniques), namely collapsing the original set of commodities to a much smaller number of factors, or indexes, statistically capturing the common part of the variability of the original series (Esposti, 2021, 2024a,b). This dimensionality reduction may be arbitrary and, in any case, may result in the loss of information by oversimplifying the complexity of the abovementioned relationships.

The second methodological problem concerns the stochastic properties of this set of commodities. Even when interdependent, these prices may still show heterogenous stochastic features. Consequently, finding a common data generation process Data Generating Process (DGP) over a large set of prices can be challenging if not unfeasible and this may prevent the use of an unifying methodology to investigate the commonality across commodities.

To deal with these two methodological challenges, the present paper proposes an approach also known as Granger Causality Networks (GCN) and consisting in a Network Analysis (NA) grounded on Granger Causality (GC) tests. On the basis of the latter, a Commodity Price Network (CPN) is constructed. On the basis of the former, the complex interdependence occurring within this CPN is investigated. At the time of writing this article, to the best of our knowledge, this is the first attempt to exploit the potentials of NA to investigate the complex interdependence across commodity prices. Therefore, this methodological proposal represents the main original contribution this paper aims to provide to the literature in the field.

NA by itself, however, does not solve the problem of dimensionality. Performing GC tests on a wide set of time series would imply the estimation of a very large VAR. In practice, this estimation is unfeasible whenever high-frequency data (monthly in the present case) requires a large number of lags. To make GC test feasible two empirical strategies can be adopted. The first consists in performing pairwise VAR (PW-VAR) estimation and consequent GC tests. The second consists in estimating sparse VAR (S-VAR) models then followed GC tests. Both solutions allow a significant reduction of the dimensionality but substantially differ in terms of statistical justification and in terms of pros and cons. For the sake of comparison, both approaches are adopted here.

In order to evaluate the potentially distinct stochastic properties of the price series, appropriate testing procedures can be implemented, particularly to investigate their stationarity. However, when multiple series are considered, it is highly unlikely that a common DGP can be identified. Therefore, the aforementioned dual approach to constructing the NA is replicated here for the price series in both levels and first differences, aiming to highlight evidence that remains robust regardless of stationarity properties.

This empirical strategy is here applied to a set of 49 commodity prices grouped in five categories: energy, metals, agriculture, food, and other raw materials. Prices are observed monthly over a period

of 45 years, from January 1980 to December 2024. The dataset thus consists of 540 observations for any commodity and generates a 49x49 matrix representing the network connection across prices.

The rest of the paper is structured as follows. Section 2 introduces the topic of commodity price interdependence focusing on the relevant recent literature in the field and the main stylised facts it highlights. Section 3 illustrates the adopted methodological approach and discusses its advantages and novelty. Sections 4 details the adopted dataset and provides some descriptive evidence. Results are illustrated in Section 5, where their robustness is also examined by comparing evidence emerging from different sets of series and different constructions of the network. Section 5 concludes by drawing some methodological and policy implications.

2. Commodity price interdependence and network analysis: methodological issues and literature review

Over the years, a substantial body of empirical literature has emerged on the multivariate analysis of commodity prices (Byrne et al., 2020). 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.) (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, as in the present study—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). This study aims to quantify and characterize these causal relationships, identifying which commodity prices are causally connected and how such links generate indirect and complex interdependencies. 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 their stochastic properties. Complex dynamics at the individual series level give rise to complex dependencies. As a result, identifying a common data-generating process (DGP) across commodities that captures interdependence within these intricate dynamics often proves challenging. In this regard, recent methodological advances emphasize the need to account for non-linearities and their potential causes—such as bubbles, volatility clustering, structural breaks, or regime shifts (Esposti and Listorti, 2013; Esposti, 2021, 2024a,b; Lin et al., 2024). Machine learning and Bayesian approaches are also increasingly employed to develop more flexible and accurate models capable of capturing these complexities (Boakye et al., 2024; Drachal et al., 2024).

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,b).

As a consequence, the third stylized fact is that price interdependence may involve a wide range of commodities. 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. The dimensionality of commodity price interdependence thus makes it resemble a network, whose structure and functioning reflect the aforementioned complexity and commonality.

The economics of commodity price connectedness is ultimately what analysts care about (Diebold et al., 2017). First, most commodities are part of many—if not most—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*. The network perspective not only highlights the complexity and multifaceted nature of this economic system, but—more importantly—offers valuable insights into its underlying structure.

This investigation requires an appropriate methodological framework. Recent literature in the field has emphasized that NA can offer a promising methodological approach. The paper widely regarded as the origin of this field of study (Diebold and Yilmaz, 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. Forecast Error Variance Decomposition (FEVD) concentrates on long-term connectedness and, therefore, requires some arbitrary structural identification restriction (Zhang and Broadstock, 2020). Moreover, the Diebold and Yilmaz (2014) analysis is limited to just 16 nodes, and it raises serious doubts about the scalability of their approach to more complex networks, as dimensionality increases rapidly with the number of nodes.

Granger Causality Networks (GCN) have recently attracted attention (Sun et al., 2018) and can be considered a variant of that original approach. GC is still based on VAR modelling[1] 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 by considering a GCN sorted by the χ^2 statistic of each pairwise estimation considering the lag structure of 8 periods. 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) analyse the interdependence of oil prices across different regions and qualities adopted within a dynamic GCN, highlighting shifts in market influence, geopolitical events, or supply-demand dynamics.

GCN seems a natural approach to investigating price connectedness, as this ultimately involves identifying first-moving prices—those at the core of supply chains whose shocks trigger responses in other prices (Esposti, 2024a). In time-series econometrics, this concept is typically captured by Granger Causality (GC) (Shi et al., 2018; Shi et al., 2020; Baum et al., 2023), an econometric tool used to assess whether one time series can help forecast another. A variable is said to Granger-cause another if past values of the former improve the prediction of the latter. Granger Causality thus helps uncover directional causal relationships, revealing which commodities directly influence others over time.

At the same time, Network Analysis (NA) models relationships among multiple commodities by analysing causality and dependencies across assets (Schweitzer et al., 2009; Newman, 2010). These methods help to identify "central" commodities (those exerting the greatest influence on others) and "peripheral" commodities in price dynamics. NA allows commodity prices to be represented as nodes

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

within a graph, where the connections (edges) between nodes represent interdependencies among commodity prices. These edges can be either directed or undirected, depending on the nature of the relationship—for example, directional or bidirectional correlations.

A network of N nodes can be represented by an $N \times N$ adjacency matrix, where each element indicates the presence and strength of a connection between nodes. In unweighted networks, the matrix elements are binary (binary adjacency matrix): 1 if a connection exists between two nodes, and 0 otherwise. In weighted networks, the elements can assume any positive value, reflecting the intensity or strength of the connection between the corresponding nodes (weighted or valued adjacency matrix). NA enables the visualization and quantification of the connections and interdependencies among commodity prices, allowing for: the identification of central commodities (hubs) that exert a greater influence on others; the detection of driver commodities—those that act as key market influencers (e.g., oil)—which affect a wide range of other prices, with significant implications for investment strategies and planning; the study of the system's topology (Diebold and Yilmaz, 2014) and the identification of clusters of commodities that tend to move together.

GCNs allow to map the direction and strength of price influence among commodities, leading to insights into how price shocks in one market propagate through others (e.g., crude oil influencing metal prices). In GCNs these causal relationships among commodities are modelled as directional links (directed edges) between nodes (commodities).

The Granger causal relations can actually be described via two different network (and graphical) models (Eichler, 2012). The first consists of a network with N nodes but whose arcs may change over time depending on the estimated coefficients of the corresponding lags of the VAR model. 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*. This representation is similar to that in dynamic Bayesian networks (Ghahramani, 1998). By applying Granger causality networks to both time domain and frequency domain, Wang et al. (2021) investigate connectedness at different frequencies, thus providing evidence also on the dynamics of these linkages. By adopting five distinct methods, Carlos-Sandberg and Clack (2021) investigate the dynamics of connectedness by assessing how causality structures evolve over time.

The second kind of network is a 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 the network model is more parsimonious and the consequent NA less computationally demanding (Ahelegbey et al., 2021), though it does not necessarily imply lower dimensionality in the estimation stage.

A static GCN approach is adopted here to investigate price connectedness among a large set of commodities. Although GCN may seem a rational and informative strategy to investigate price interdependence across a large set of commodities, it is worth stressing that, whether static or dynamic, the main empirical challenge of the GCN approach lies in handling high dimensionality—namely, a large number of nodes. The aim of this study, therefore, is to begin with a static network to compare different empirical strategies for GCN under high dimensionality, in order to identify the most feasible, reliable, and robust solution—if any. For this reason, we do not consider weighted or dynamic networks at this stage. However, once the most effective empirical strategy is established, extending the proposed approach to these more complex cases represents a promising direction for future research in the field.

The dimensionality issue arises not only for the large number of commodities involved but also because the GCN approach encounters some major issues about VAR modelling and estimation. First of all, VAR modelling by itself tends to exacerbate the dimensionality problem. In a VAR(N,K) model the number of parameters to be identified and estimated is NxKxN. In this respect, the higher the frequency the longer is the lag to be admitted in VAR modelling since indirect interdependence may

take time and seasonal or in-year cycles can occur. The consequence is that, if N is already large and many lags must be admitted, the number of parameters explodes and estimation may become unfeasible given the available observations T. According to Bernanke et al. (2005), for instance, in order to save degrees of freedom, standard VARs are typically limited to a small number of endogenous variables (namely, N), usually less than ten, due to the curse of dimensionality.

Due to this high dimensionality a GCN requires a more parsimonious VAR model specification. For the sake of comparison, two empirical strategies are considered here. The first consists in performing pairwise VAR (PW-VAR) estimation and consequent GC tests. The second consists in estimating sparse VAR (S-VAR) models, followed by GC tests. Both solutions allow a significant reduction of dimensionality but substantially differ in terms of statistical justification and of pros and cons.

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. Granger causality based on only two variables severely limits the interpretation of the findings: without adjusting for all relevant covariates, a key assumption of Granger causality is violated. This limitation has been well documented (see, e.g. Lütkepohl, 1982; Shojaie and Fox, 2021) and eventually implies inconsistent VAR estimates due to omitted relevant variable bias. In spite of their limitations, bivariate tests of Granger causality continue to be widely used in many application areas, from economics (Chiou-Wei et al., 2008) 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, or connection, by chance. As the number of tests increases,[2] the probability of at least one false positive approaches 1. The expected proportion of false positives is known as the False Discovery Rate (FDR) (Runge, 2018; Seth et al., 2025; Uematsu and Yamagata, 2025).

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. Granger causality was then typically assessed using classical hypothesis testing methods, such as the F-test. More recent works (Lozano et al., 2009; Chudik and Pesaran, 2011; Basu and Michailidis, 2015) have increasingly focused on directly selecting the nonzero entries via sparsity-inducing penalties, often by augmenting the VAR loss function. These works introduce sparsity-inducing penalties by adopting LASSO (Least Absolute Shrinkage and Selection Operator) as the regularization technique (Tibshirani, 1996; Yuan and Lin, 2006). LASSO estimation of S-VAR models is adopted here. But Bayesian approaches continue to be considered as an alternative to regularization methods for analysing large VAR processes (George et al., 2008; Banbura et al., 2010; Ahelegbey et al., 2021).

While promoting sparsity, LASSO estimation can still capture complex interactions among multiple variables. However, it typically requires greater computational resources and careful tuning of penalty parameters. In addition to its computational cost, LASSO introduces a degree of arbitrariness, as

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

sparsity can be achieved through various regularization strategies, with the final choice often left to the analyst's discretion (see Section 3).

For the sake of comparison, we employ both PW-VAR and S-VAR approaches to construct the Granger Causality Network (GCN) of a large set of commodity prices. While both methods address high dimensionality, the PW-VAR approach does not inherently ensure sparsity in the resulting GCN. At the same time, P-VAR models can be considered a more intuitive, interpretable, and simpler method for detecting Granger causality, albeit at the cost of potential estimation inconsistency. In contrast, large VAR models with sparsity-inducing penalties offer a more parsimonious, comprehensive and theoretically grounded framework for high-dimensional settings, though they may be less immediately interpretable. In practice, however, the degree of sparsity obtained via S-VAR depends on the penalization which is eventually established by the analysist. Also in the PW-VAR approach more sparsity can be induced by adopting a more selective significance level or testing procedure to accept Granger causality (see Section 3). In this respect, PW-VAR can still be regarded as a reliable approximation of a GCN based on the more sophisticated S-VAR approach, with the sparsity and structure of the former ideally converging to that of the latter. This may explain why, even just for comparative purposes, several recent contributions continue to adopt this approach (Sun et al., 2018; Zhang and Broadstock, 2020; Larrosa et al, 2024).

A further complication in applying VAR modelling to GCN analysis, particularly in high-dimensional settings, concerns the stochastic properties of the time series under consideration—most notably, their stationarity. VAR models require that all variables be stationary, a condition that is increasingly difficult to satisfy as the number of series (N) grows. Stationarity tests may yield inconclusive or conflicting results for individual series, depending on the specific testing procedures and specification used.

When N is large, it becomes exceedingly difficult to establish a common stochastic property across all series—namely, that they share the same order of integration. This lack of conclusiveness poses a problem not only for the S-VAR approach, where a VAR(N, K) must be estimated consistently, but also for the PW-VAR framework. In the latter case, it is highly unlikely that none of the $N\times(N-1)$ estimated VAR(2, K) models would involve pairs of commodities with differing integration orders. One potential solution for dealing with I(1) series is cointegration analysis, leading to the specification of Vector Error Correction Models (VECMs). However, this approach requires that all involved series can be unambiguously classified as I(1), and moreover, the number of cointegrating vectors may increase substantially as N grows, posing additional estimation challenges.[3]

As the number of time series increases, stationarity and integration testing become increasingly unreliable or inconclusive due to test size distortions, power loss, and cross-sectional dependencies. In such high-dimensional environments, it is often more pragmatic to assume a common integration order—either I(0) or I(1)—and proceed with modelling strategies accordingly (Stock and Watson, 2002; Lütkepohl, 2005). It follows that when N is large, it is reasonable to abandon the idea of conclusively determining a common order of integration across all series. A more pragmatic approach is to consider both possible configurations: either all (or most) of the series are I(0), or they are predominantly I(1). I(0) series imply VAR models in levels, whereas I(1) series require VAR models in first differences of prices.

Investigating price connectedness within GCNs using price levels versus first differences is clearly not equivalent. While price levels may capture both long-term and short-term interdependencies, they carry the risk of identifying spurious causal relationships. In contrast, first differences emphasize

[3] Granger causality testing among non-stationary series can also be performed within a Lag-Augmented VAR (LA-VAR) modelling framework (Dolado and Lütkepohl, 1996; Yamada and Toda, 1998). Some applications of this extended Granger causality testing have already been presented (Baum et al., 2023), also with specific reference to commodity prices (Shahzad et al., 2021; Adeosun et al., 2023; Aharon et al., 2023; Esposti, 2024b). It remains questionable, however, whether LA-VAR modelling may be appropriate under high-dimensionality so this solution is not considered here but could represent an interesting extension of the approach proposed in this study.

short-term dynamics and help mitigate this risk. However, the goal here is to identify the most effective empirical strategy for conducting GCN analysis on a large set of commodity prices. Therefore, we do not seek to determine whether using price levels is preferable or more appropriate than using first differences—or vice versa. Rather, we aim to highlight robust findings and any differences that emerge from comparing the two cases.

A further issue concerning the VAR framework underlying GCN analysis, which recent literature has begun to explore, is the presence of possible nonlinearities in price dynamics and, consequently, in their interdependence. VAR models—whether PW-VAR or S-VAR—are linear by design: each variable (i.e., a commodity price) is modelled as a linear function of its own past values and those of other commodities in the network. As a result, this framework may overlook persistent or transient nonlinear patterns in commodity price behaviour.

A possible way to capture connections across prices in a more sophisticated—and specifically nonparametric—manner that also admits nonlinearities is the adoption of Machine Learning (ML) techniques (Kozian et al., 2025). Nonlinear correlation among prices can also be investigated through wavelet analysis (Boako et al., 2020; Nigatu and Adjemian, 2020; Kirikkaleli and Güngör, 2021; Mastroeni et al., 2022; Mutascu et al., 2022). As with other nonparametric techniques, wavelets are flexible tools for investigating nonlinearities such as cycles or structural breaks in price series. They can thus offer a more realistic representation of interactions among commodity prices. In particular, Wavelet Cross-Correlation (WCC) allows for the analysis of price interdependence at different time scales, which makes it possible to include or exclude long-term trends and short-term cycles. However, the WCC's sensitivity to the filtering process—returning different correlation values depending on how the signal is decomposed—implies that the corresponding elements in the matrix defining the Commodity Price Network (CPN) may lack uniqueness. Moreover, WCC is still a pairwise method, and may therefore overlook complex indirect linkages among multiple commodities. Additionally, wavelet techniques may struggle to fully account for nonstationarity, potentially leading to spurious or misleading correlations. To address such limitations, various extensions of the VAR model to nonlinear or nonparametric forms have been proposed, such as kernel-based VAR, neural network-based VAR (NN-VAR), and Gaussian process VAR (Signoretto and Suykens, 2015).

Ultimately, all solutions proposed in the recent literature to address the linearity assumption tend to be computationally intensive and, in practice, become infeasible in high-dimensional settings (i.e., when N is large). Therefore, while not suitable for the present analysis, these approaches nonetheless represent a promising direction for future research. Also in this respect, we prefer here to adopt a more pragmatic, namely feasible, approach. Though maintaining the linear structure of VAR models, nonlinearities may still be admitted in the behaviour of the original series via logarithmic transformation. Given the inherently linear structure of VAR models, applying a logarithmic transformation to the variables may help in capturing more complex dynamics and nonlinear relationships in levels, such as multiplicative or power-law dynamics.

This strategy is widely adopted in empirical studies (Lütkepohl, 2005; Enders, 2014) and can be applied here to both price levels and their first differences. For instance, a log-linear model in first differences implies that the original price levels follow a multiplicative and path-dependent process. Consequently, even though the differenced model is linear, the relationship among price levels is nonlinear and cumulative (van Garderen, 2023). Repeating the GCN analysis on these different series and comparing the respective results can therefore help identify evidence that appears robust regardless of the underlying dynamics of the price series and their interdependence.

3. The methodological approach

Our approach is based on the Granger Causality Network (GCN) analysis. The core idea lies in representing the interdependencies among N commodity prices as a directional, unweighted

adjacency matrix of size $N \times N$, which we refer to as the Commodity Price Network (CPN). Each element of this matrix takes a value of 0 or 1, depending on whether the commodity in the column Granger-causes the price response of the commodity in the corresponding row. In this context, connectedness means Granger causality. Thus, the matrix is constructed based on Granger causality tests performed within a VAR model. The matrix is unweighted, as connections are binary and do not capture the intensity of the relationship. It is also directional, since a causal effect from commodity *i* to commodity *j* does not imply a reverse effect. As such, the matrix is non-symmetric by construction.

The proposed methodology thus consists of a three-stage approach. The first stage detects the elements of the adjacency matrix via Granger causality testing. Depending on how the VAR model is specified (PW-VAR or S-VAR) the first stage has two alternative versions, each with different GC testing procedures. In the second stage, these variants generate different and alternative possible CPNs. The third stage performs the Network Analysis on these alternative CPN versions by making use of various indicators expressing the network structure as well as the degree and nature of interdependence across prices and groups of prices.

3.1. Granger causality testing

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

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

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

$$(2) \quad p_{i,t} = \alpha_i + \delta_i t + \sum_{k=1}^{K} \sum_{j=1}^{N} b_{ijk} p_{j,t-k} + \varepsilon_{it}, \qquad \forall i, j \in N; \forall t, k \in T;\ K < T$$

As (2) 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*):

$$(3) \qquad \mathbf{p}_t = \mathbf{A} + \mathbf{D}t + \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \mathbf{B}_k \mathbf{p}_{t-k} + \boldsymbol{\epsilon}_t$$

where $\mathbf{p}_t$ is the *N*x1 vector of prices and $\boldsymbol{\varepsilon}_t$ the *N*x1 vector of the i.i.d. disturbance terms at time *t*. $\mathbf{A}$ is the *N*x1 vector of drift coefficients, $\mathbf{D}$ is the *N*x1 vector of deterministic trend coefficients, $\mathbf{B}_k$ is the *N*x*N* matrix of price correlation coefficients at generic lag *k*.

Provided that the series are stationary (i.e., I(0)), the VAR model in equation (3) can be consistently estimated in levels, as is standard practice (Enders, 2014). Alternatively, if the series are I(1), equation (3) can be specified in first differences. As previously noted, while preserving the linear structure of equation (3), it is also possible to capture potential non-linear relationships among prices by estimating the model in a log-linear form. In practice, equation (3) 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_j, 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, $H_0: b_{ij1} = b_{ij2} = ... = b_{ijK} = 0$. Granger causality 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, unlike the Pearson correlation coefficient, 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.[4] To populate this *N*×*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, we assign $GC_{ij} = 0$. Since the network is directional, whatever the value GC_{ij}, it can be either $GC_{ji} = 0$ or $GC_{ji} = 1$.

In practice, however, 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 *N*x*N*x*K* (Morana, 2012). Second, a VAR(*N*,*K*) implies performing *N*x(*N*-1) GC tests which entails estimating the model *N*x(*N*-1)+1 times (1 unrestricted and *N*x(*N*-1) restricted models). Third, for any equation of the model, *N*x*K* parameters must be estimated. Therefore, the number of available observations *T* must be sufficiently large to ensure identification and robust estimation, i.e., T>>*N*x*K*. In the present case, where *N* = 49 and *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:

$$
\begin{aligned}
p_{it} &= \alpha_i + \delta_i t + \sum_{k=1}^{K} b_{iik} p_{it-k} + \sum_{k=1}^{K} b_{ijk} p_{jt-k} + \varepsilon_{it} \\
p_{jt} &= \alpha_j + \delta_j t + \sum_{k=1}^{K} b_{jjk} p_{jt-k} + \sum_{k=1}^{K} b_{jik} p_{it-k} + \varepsilon_{jt}
\end{aligned}
\tag{4}
$$

It means estimating *N*x(*N*-1)/2 models each with 2*K* parameters and this seems much more feasible given the available *T* observations. This PW-VAR approach estimates bivariate models for each pair of variables, reducing complexity. However, this approach 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 multivariate dynamic system, enhancing interpretability and efficiency,

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

particularly when dealing with a large number of variables or long time series. 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 Least Absolute Shrinkage and Selection Operator (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 (2), the following optimization problem is solved (Shojaie and Fox, 2021):

$$(5)\quad \min_{\{b_{ijk}\}} \left\{ \frac{1}{T} \sum_{t=K+1}^{T} \left(p_{i,t} - \alpha_i - \delta_i t - \sum_{k=1}^{K} \sum_{j=1}^{N} b_{ijk} p_{j,t-k} \right)^2 + \lambda \sum_{k=1}^{K} \sum_{j=1}^{N} |b_{ijk}| \right\}$$

where $\lambda>0$ is the regularization or penalization scalar parameter that applies to the sum of the absolute values of all the *N*x*K* 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, inducing sparsity, 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, allowing more variables to remain in the model and 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).

An interesting combination of the PW-VAR and S-VAR approaches could consist of first performing pairwise Granger causality tests and then using the resulting p-values to inform the LASSO estimation of the S-VAR model. This approach would require an extension of the standard LASSO that employs variable-specific penalty weights, a method commonly referred to as adaptive LASSO or weighted LASSO (Takada and Fujisawa, 2023). Such a strategy would allow the regularization process to incorporate statistical evidence of predictive relevance, potentially enhancing model interpretability and forecasting performance. However, since the objective here is to compare alternative PW-VAR and S-VAR constructions of the GCN, this option is not pursued in the present analysis, although it may represent a promising direction for future research in this field.

3.2. Building the Granger Causality Network

A GCN is represented by an adjacency matrix like the following (Sun et al., 2018):

$$(6)\quad \mathbf{GCN} = \begin{bmatrix} GC_{11} = 0 & \cdots & GC_{1N} \\ \vdots & \ddots & \vdots \\ GC_{N1} & \cdots & GC_{NN} = 0 \end{bmatrix}$$

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 predicted) 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 $GC_{ij} = 0$.[5] An additional key insight from the network structure is the role of indirect linkages, which, through direct Granger causality (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. The main reason consists in the very large numbers of binary GC tests, *N*x(*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 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 *N*x(*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 these issues. 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 CGNs 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. BIC tends to select a larger λ, as it imposes a stronger penalization on model complexity. Consequently, the CGN obtained using BIC (GNC5) is expected to be sparser than the one generated via CV (GCN4).

The comparison across this sequence of networks, from CGN1 to CGN5, is performed using the different time series. Commodity price levels are considered first. Then, to account for potential non-stationarity and non-linearity, the same sequence of networks is generated,–and the comparison

[5] 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).

repeated,–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; leader nodes; clusters and communities; network correlation.

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 them 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. [6] 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

[6] 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.

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.

The topology of a network can also be characterized relatively by quantifying its structural differences with respect to reference networks of known topology (synthetic networks). This aspect will be explored in Section 3.3.4.

3.3.2. Leader nodes

Identifying key nodes, or leader nodes, within a network is crucial for understanding the network's structure and 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 the 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 will have 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 $\mathbf{GCN}$ is the NxN adjacency matrix, $\boldsymbol{x}$ is an Nx1 eigenvector of $\mathbf{GCN}$ and λ is the respective eigenvalue.

3.3.3. 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.[7] The maximum eigenvalue λ_{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.

3.3.4. Network correlation

The topology of a network can also be described in relative terms, by assessing how its structure differs from that of other networks. This comparison is often made with synthetic networks, namely networks with the same number of nodes but well-understood structural properties. In particular, five synthetic models are commonly considered as extremes along the structural spectrum (see also Section 5.3.1 and Annex 5): 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 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

[7] 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.

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. However, 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.

4. Price series under scrutiny

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.[8]

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. The IMF aims to reflect global price dynamics through these market prices; therefore, the selected markets are typically among the most important globally and are commonly used as reference prices, even by agents operating in more localized markets. As such, these prices can be legitimately considered as proxies for otherwise unavailable global market prices. Nonetheless, it is acknowledged that some commodity prices may exhibit a more regional character than others (e.g., agricultural commodities compared to energy commodities). While not all price series are based on U.S. markets, all prices are denominated

[8] The availability of monthly price series of natural gas starts in 1985M1. The previous period has been thus recovered by interpolating backward the 1980-1985 US yearly natural price series as provided Energy Information Administration (EIA).

in U.S. dollars. This ensures that the dataset is not affected by commodity price fluctuations driven by exchange rate movements (Garzón and Hierro, 2022).

The relevance of the selected commodities is further 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 Union (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. Ensuring a prompt response to potential supply crunches, often signalled by sudden price spikes, is increasingly urgent for contemporary economies.

The novelty of the present contribution lies partly in the methodology, which, while not entirely new, builds on recent approaches and includes a comparative assessment of alternative solutions. However, the main innovation stems from the high-dimensional nature of the dataset, which includes nearly 50 individual commodities spanning diverse groups and a long period of time with relatively high-frequency observations. Unlike Zhang and Broadstock (2020), who also consider a broad range of commodities but aggregate them into seven price indexes (Metal, Food, Precious, Oil, Raw, Beverage, Fertilizer), this study focuses on disaggregated commodity-level data, allowing for a more granular analysis of price dynamics across heterogeneous markets.

The commodity grouping indicated above (EMAFO) may be somewhat arbitrary. 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 assumption 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-colonial” (Barley, Beef, Corn, Lamb, Pork, Poultry, Rice, Sorghum, Soybean, Wheat) products. The latter can also be further divided into “animal products” and “crops”. Moreover, colonial products can also be found in the 'Food' group, as is the case with coffee and tea.

Colonial products are exemplary of commodities that are produced by a limited group of countries. They are mostly export-oriented commodities and, for this reason, they are expected to be less and differently connected with other (i.e., non-colonial) commodities. They may also involve another possible source of price interdependence that has nothing to do with some supply-chain connection or financial market connection. Since they mostly come from a limited number of countries, macroeconomic shocks in one of these economies may transmit to a whole range of apparently unrelated commodities, not necessarily colonial.

Besides the geographical provenience, “agriculture” and “food” groups themselves could be crossed and reassembled 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”. However, the key argument here is that 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.

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 (Crain and Lee, 1996). The logic behind this choice is that we prefer to analyse the price series that economic agents really confront and on which they take decisions without possibly introducing artificial transformations. It remains true, however, that indirect effects and, more generally, complex structural dynamics only become

apparent after a sufficient number of periods. When adopting a monthly frequency, considering a lag length of 12 months (K=12) in the present case (N=49) would require estimating 564 parameters for each 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 appropriate empirical strategy that balances potential seasonal effects with the need to limit model complexity (see Section 3.2).

5. Results

The results of the analysis are presented following the sequence of steps outlined in Section 3 (see also Sun et al., 2018; Ahelegbey et al., 2021; Zhou et al., 2022). 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 a multivariate representation of the commodity price formation process, namely a VAR model. This requires that all series share the same order of integration and, in particular, that they are all I(0) (stationary processes). Prior to constructing the GCNs, unit root testing must therefore be performed to verify this condition. Here, the standard Augmented Dickey-Fuller (ADF) test is adopted (Sun et al., 2018; Zhou et al., 2022) whose null hypothesis is that the series have a unit root, i.e. I(1) (non-stationary) processes. The 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 (1)) and simplifies it step by step by removing unnecessary deterministic components (trend and drift) based on their statistical significance. Moreover, 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 four different sets of series: price levels, first differences of prices, logarithms of price levels, and first differences of the logarithm of prices.

With few exceptions, results are largely correspondent across the 49 commodities: all price series can be considered I(0). 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 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. To assess whether, and to what extent, the presence of non-stationarity in a few price series may affect the network analysis, the latter will henceforth be conducted on both sets of series: price levels and first differences of prices. Since, as expected (Esposti, 2024a,b), the logarithmic transformation does not alter the evidence emerging from the tests, these transformed series will only be considered in Section 5.4 to assess the robustness of the network analysis compared to potentially non-linear relationships among commodity prices.

5.2. Granger causality

As detailed in Section 3, 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 derived from these estimates.

5.2.1. Pairwise VAR models

To generate GCN1–GCN3, we separately estimate a battery of 1176 VAR(2, K) models. 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 selected for each equation using the BIC, subject to the constraint that K ranges from a minimum of 4 to a maximum of 12.

Figures 1 and 2 display the network topology for all PW-VAR GCN variants. As expected, the GCN becomes progressively sparser (i.e., less dense) both when moving from GCN1 to GCN3 and when transitioning from networks based on price levels to those constructed from the first differences of prices. 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. 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.

This is made explicit in Annex 3 (see also Table 2), 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 practice, switching from networks based on price levels to those based on first differences reduces the number of connections, therefore the network density, to a similar extent as adopting a stricter significance threshold for Granger causality. However, while the latter mainly affects the number of connections, the former may also alter the nature and structure of the network itself and this aspect deserves further investigation.

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.

Figures 3 and 4 display the network topology for all S-VAR GCN variants. Also in this case, 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). In practice, PW-VAR modelling allows to obtain at least the same range for sparsity of the S-VAR approach. Nonetheless, as previously noted, they may lead to a network with a different nature and structure.

To better assess whether and how the GCN variants structurally differ, beyond sparsity, Table 1 reports the Pearson correlations of the adjacency matrices corresponding to the different network versions. As expected, correlations are positive across all GCNs, but not all correlations are statistically significant at the 5% or 10% level, indicating that some networks share structural similarities, while others differ substantially. On the one hand, a strong correlation emerges among GCN1, GCN2, and GCN3, both in terms of price levels and first differences. This confirms that adopting a stricter significance level increases sparsity but does not substantially alter the network structure. A similar pattern is observed in the GCNs obtained when 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 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. Particularly surprising and 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, as indicated by the lack of significant correlations between analogous GCNs across the two sets of series. Conversely, 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 provide a summary of network characteristics and highlight structural differences across GCNs. Table 2 reports these network-wide metrics. A geodesic 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.

Three general aspects of the network topology deserve attention and appropriate indicators: density, granularity, and peripherality. Concrrning the former aspect, the density indicator confirms that sparsity increases when moving from GCN1 to GCN3, from GCN4 to GCN5, and from price levels to first differences of prices. In any case, the densest variant (GCN1 with price levels) is still less than half as dense as the geodesic network. Table 2 also reports three conventional distance-based metrics that reflect the overall connectivity level 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. In the present study, only for GCN2-5 with first differences of prices, the largest component of the path is lower to 2352 since these cases 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, it 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 (based on price levels) to a maximum of 2.33 in GCN3 (based on first-differenced prices).

However, as discussed in Section 3, 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 and transitivity are helpful concepts in this respect and can be expressed with appropriate indicators. In practice, reciprocity is measured as the proportion of mutual connections (dyads) relative to all directed ties. In the present case, reciprocity varies between a maximum of 0.371 (GCN1 obtained with price levels) and a minimum of 0.061 (GCN5 obtained with the first differences of prices), values that are much lower than the hypothetical reciprocity within a geodesic network whose value is 1. 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). 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. While in a geodesic network only mutual dyads are observed, so no form of granularity occurs, in none of the present GCNs this configuration is prevalent even though it is evidently associated with density and its presence declines moving from GCN1 to GCN3 and from GCN4 to GCN5, as well as moving from networks obtained with price level to those obtained with the first differences of prices.. In all GCNs obtained with the first differences of prices, the null configuration is prevalent. Only in denser networks obtained with price levels (GCN1 and GCN2) the prevailing configuration is the asymmetric dyad, that is, a pair of nodes connected by a one-way relationship, where the tie goes from one node to the other but not vice versa.

Granularity may be investigated further via another indicator, transitivity. Formally, transitivity is measured as the ratio of closed triads (or triangles) to all possible triads in the network. 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 GCNs here considered show 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 it increases from GCN1 to GCN5 but also declines when passing from networks obtained with levels to networks obtained with first differences. Eventually, the largest transitivity is found for GCN5 with price levels while the lowest is for GCN5 with the first differences of prices. 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. There are 16 possible configurations and they are detailed in Table 4. It emerges that, while 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, all GCNs show a more articulated set of triadic configurations. In general terms, for GCN obtained with price levels, the prevalent configuration is that designated with 012 indicating one arc between two nodes, while the third node is isolated. The rarest is that indicated either by 030C (a triangle with one directed arc missing, but with an additional directed arc forming a "T") or the abovementioned 300. In the case of GCN based on the first differences of prices, the prevalent configuration is that of either no triads (003: no arcs between the three nodes), if the exclude the case

of GCN1 behaving like networks obtained with price levels. To further confirm the structural difference compared to a geodesic network, in all cases the less frequent configuration is 300.

A final synthetic indicator of network-level granularity is the overall clustering coefficient of a network (see Section 3.3.3). At the node level, this coefficient measures the tendency of nodes to form triads (Section 5.3.2). At the network level, the clustering coefficient expresses the average node-level clustering value reported at the bottom of Table 7. It thus indicates how much the nodes in a network tend to cluster together, forming local clusters. Results obtained indicate that the average clustering coefficient follows the same gradient observed for density: it decreases from GCN1 to GCN3 and from GCN4 to GCN5 for both levels and first differences of prices, but it also decreases passing from a GCN with levels to the analogous network with first differences of prices. This correspondence between density and the clustering coefficient of the network is confirmed by the fact that the geodesic network takes a value of 1, which is the theoretical maximum and it is about double the maximum value observed for a GCN (0.584), more or less the same ratio observed for density (Table 2).

Positioned conceptually between granularity and peripherality, Table 2 presents the network-level centralization indicators (see Sections 3.3.2 and 5.3.2). 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 centrality 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. In the present case, 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.

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. Here, indegree centralization replicates the gradient of the density indicator with its value increasing going from GCN3 to GCN1 and from GCN5 to GCN4 in both levels and first differences of prices. The highest value (0.417) is found for GNC1 for price levels. A more mixed relationship with the density, or sparsity, 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 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 (0.359) is found for GCN4.

The investigation of peripherality within a network is usually performed by looking at the number and frequency of global and local bridges. The latter indicator also serves as a conceptual shift from network-level to node-level analysis (see next section). 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.[9]

Table 5 presents the global and local bridges across GCN variants, along with the nodes involved (isolated nodes are reported in detail in Table A5, Annex 4). Geodesic network does not present either local or global bridges as all nodes are connected to all other nodes. Since this indicator expresses the tendency of a node to be isolated or peripheral, the number of global bridges tends to increase passing from GCN1 to GCN3 for both levels and first differences of prices. The same occurs passing from

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

GCN4 to GCN5 and, for all GCNs, from networks obtained with price levels to networks obtained with the first differences. However, it is not necessarily true 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 tendency to 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, while the group that seems less involved is clearly that of metals.

It is finally worthwhile to explore potential similarities between these GCNs and synthetic networks, i.e., networks with known properties and structures. These networks are usually very regular and well-organized, resulting in a structure with highly homogeneous density (or sparsity).[10] This comparison thus allows for the analysis and comparison of the fundamental structural properties of a real-world network. 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.

In order to assess the similarity between our GCNs and the synthetic cases, Table 6 reports the respective pairwise Pearson correlations. The results provide clear and concordant evidence. Despite the previously mentioned quantitative and qualitative (i.e., structural) differences among the ten GCN variants resulting from different VAR models and data series, all GCNs show very low correlation (always below 0.05 in absolute value) with all synthetic networks, with the sole exception of the geodesic model. In this latter case, by contrast, all GCNs exhibit a similar, statistically significant (except for GCN3 based on the first differences of prices), and negative correlation. This correlation is particularly strong for all GCNs derived from price levels.

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: they 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.

5.3.2. Node-level metrics

The objective of this section is investigating which are the more relevant, critical or central nodes, namely commodities. This can be achieved firstly and easily by counting the connections of a single node. Then, by adopting more sophisticated indicators (see Section 3.3.2). Figures 5-8 report the number of arcs per node in the various GCN variants. They present the same data as Tables A3 and A4 (Annex 3), but grouped by commodity type and arranged in descending order. Moreover, in addition to incoming and outgoing arcs, the figures also report the node-by-node balance. Due to space constraints, we only present the sparsest networks, namely, GCN3 and GCN5 for both levels and first differences. These are arguably the cases that exhibit the most robust connections. All other

[10] Figure A1 (Annex 5) displays the topology of these synthetic networks generated with 49 nodes (i.e., the commodity prices under consideration).

networks (GCN1, GCN2, and GCN4) are available upon request and provide largely consistent evidence.

It emerges that the node-level metrics significantly differ across the GCNs and this further demonstrates that adopting alternative VAR modelling and datasets does not only imply a quantitative difference within the network, namely density, but also a qualitative difference, namely the role of single nodes and groups of nodes. This may explained by the fact that looking for more sparsity also brings about a selection of the most relevant linkages within the network. It is also worth noting that a node, or a group of nodes, can emerge either through its capacity to generate a shock (outgoing causality) or through its propensity to be affected by one (incoming causality). Core nodes typically exhibit high values in both dimensions, as transmitting a shock within a network requires the ability to both receive and propagate the signal. However, the balance between outgoing and incoming connections reveals the extent to which a node can be considered a net generator of shocks within the network, rather than merely a transmitter.

Besides the abovementioned differences emerging across the GCNs, some regularities are worth noting. Regarding the five groups of commodities, one somewhat unexpected result is that energy commodities are not particularly critical within the network. They appear to act more as transmitters than as sources of shocks, as indicated by the balance between outgoing and incoming connections, which is predominantly negative and, in any case, lower than that of most other groups. Conversely, the group of metals emerges more as a source rather than a transmitter of price shocks within the network. In most GCNs, their balance is not only largely positive but also typically the highest among all groups. For the other three groups, the picture is more mixed. They appear to be, at least for several of their respective commodities, more peripheral to the core of the network, with lower average values for both outgoing and incoming connections, and a balance that tends to be closer to zero compared to energy commodities and metals.

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 and sunflower oils, also stand out as leader nodes among food products. Finally, within the group of other raw materials, fertilizers, notably potassium and diammonium phosphate, appear to behave as more central nodes.

As discussed in Section 3, however, the centrality of commodities within the network cannot be fully captured by simply counting and categorizing local connections. In a network, a node's importance is not determined solely by its immediate neighbors, 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. To investigate this more comprehensively, global centrality measures can be employed. In this context as well, the main objective remains to identify results that appear robust across different network configurations.

Table 7 reports the node-level clustering coefficient that expresses the capacity of a node to a act an aggregator. A high clustering coefficient means the node's neighbours are well connected (forming tight-knit communities). A low clustering coefficient suggests the node's neighbours are not well connected. 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 al groups. The highest value usually concerns metals, when price levels are concerned, while in the case of the first differences of prices the highest values are found also for commodities belonging to agricultural and food products or to other raw materials. The lowest values are most frequently found in this latter category.

But the influence of a single node within the network can be better expressed by degree centrality indicators (Tables 8-10). As discussed in section 3.3.2, different indicators can be computed at the node level: the betweenness, the closeness, the farness and the eigenvector centrality. They provide more or less complementary and largely concordant evidence. A higher betweenness centrality score (Table 8) indicates that the node plays a more crucial role in connecting different parts of the network. It is confirmed that the group of metals tends to show higher betweenness centrality but, also in this case, commodities with the highest values can be found in all groups. In particular, when passing from GCNs based on price levels to those based on the first differences, these highest values tend to shift from metals and energy commodities to the other groups, including other raw materials. In this latter group, though not exclusively, we also find the commodities with the lowest value of this centrality indicator.

Higher closeness centrality (Table 9) indicates that a node is more central, meaning it can reach other nodes more quickly. It clearly emerges here how passing from GCNs based on price levels to those based on the first differences of prices may substantially modify the evidence about centrality. In the former case, it is confirmed that metals show the highest values. In the latter case, however, this leadership is contested by the group of other raw materials in which, nodes with high centrality are found usually among the group of fertilizers or wood products. Farness centrality (Table A6, Annex 4) is a similar indicator to closeness centrality (it is essentially the inverse): nodes with higher farness centrality are less central, meaning they are farther away from other nodes in the network. This indicator provides evidence that largely corresponds to what has been illustrated for the closeness centrality.

Eigenvector centrality is a more sophisticated concept and measure, as it evaluates the importance of a node within the network not only based on its direct connections, but also by considering the importance of the nodes it is connected to. A node with a high eigenvector score is therefore linked to influential and well-connected nodes, whereas a lower score indicates connections to less important commodities. Since eigenvector centrality is only defined for connected networks, i.e., those without isolated nodes, this indicator is available and reported for all GCNs based on price levels, and only for GCN1 when based on first differences in prices. Table 10 clearly shows that, on average, minerals, and to a lesser extent, energy commodities, are more central than commodities in other groups, which tend to exhibit similar average performance. However, highly central individual commodities can be found across all groups, notably copper among metals, rice among agricultural products, sunflower and palm oils among food commodities, and fertilizers, particularly potassium and diammonium phosphate, among other raw materials.

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 capacity to propagate and amplify shocks. This value is reported in the last column of Table 2 for all GCNs under consideration. Interestingly, if we exclude the cases in which the presence of isolated nodes causes λ_{max} to equal 1 (namely, GCN2–4 based on the first difference of prices), the networks with the highest eigenvalues are not necessarily the densest ones. On the contrary, λ_{max} increases from GCN1 to GCN3 and from GCN4 to GCN5, both in the case of price levels and for networks based on first differences. 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 in this context, as λ_{max} takes a value very close to the theoretical maximum of 4.

Eventually, a diffuse form of leadership emerges within the network, with critical nodes found across different commodity groups. Commodities that are central within their respective groups become critical in spreading shocks throughout the entire network via indirect connections. The relevance of both within-group and between-group connections, however, deserves further investigation.

5.3.3. Community structure

Interacting nodes may give rise to a third layer of network characteristics beyond network-level and node-level properties. In this section, we aim to explore this additional layer represented by groups larger than small local clusters, also known as communities. In addition to density, granularity, and peripherality, the presence of communities of relatively homogeneous commodities contributes to a more complete understanding of the structure of the GCNs under analysis. 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. Specifically, one key insight we are interested in is whether, and to what extent, interdependence is stronger within the five groups than between them.

Table 11 examines this community structure across the GCN variants, using two indicators. The first is the proportion of within-group connections relative to the total number of connections for each group. However, this indicator may be biased by the fact that some groups contain fewer commodities (only three in the case of the energy group, thirteen in the case of agricultural products), which affects the proportion. To mitigate this effect, we also use modularity, as defined in Section 3. In practice, this metric captures the ratio between intra-group and inter-group density of connections, based on either incoming or outgoing links.

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, certain regularities emerge: metals tend to form a strong community in terms of incoming connections, while it is significantly weaker for outgoing linkages. The opposite pattern is observed for energy commodities, where 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. 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.

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.

5.4. Robustness check: comparison with transformed series

A final assessment that may help to evaluate the robustness of results obtained concerns repeating the analysis using the logarithmic transformation of the series (i.e., the logarithm of price levels and their first differences). As discussed in Section 2, a major challenge in studying commodity price interdependence is the presence, often temporary, of non-linear 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 for selected GCNs based on the logarithm of prices (hereafter referred to as transformed GCNs). Due to space limitations, only the sparse networks (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.[11]

A positive correlation emerges between the transformed GCNs and all other GCNs considered thus far. As expected, however, 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, GCN2, and GCN3, as well as between the transformed GCN5 and the untransformed GCN4 and GCN5. 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, confirming 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. When compared to their untransformed counterparts (GCN3 and GCN5), these networks exhibit broadly similar properties. Nonetheless, particularly in the case of GCN3, the transformed network appears slightly less dense and more granular, with a modest increase in the peripherality of some nodes. The motivations and implications of these slight differences, potentially associated with the logarithmic transformation and, by extension, with the presence of nonlinear dynamics, may therefore warrant further investigation in future research.

6. Concluding remarks and policy implications

This study investigates the interdependence within and among different groups of commodity prices. A substantial body of literature exists on this topic, with many recent contributions emphasizing the high complexity of such interdependence, partly due to the large number of diverse commodities involved. Properly investigating this interdependence, while accounting for this diversity, presents a major dimensionality challenge. To address this issue, the present study proposes a methodological solution based on the concept of Granger causality and subsequent network analysis. We therefore introduce a three-stage approach capable of generating several variants of the commodity price network.

Applying this approach to a set of 49 commodities observed over 540 months (from 1980 to 2024) allows us to assess its ability to reveal the main structural features of complex price interdependence. At the same time, it helps to evaluate whether, and to what extent, these features are robust across the different network variants that the approach can generate. In this respect, the results indicate that alternative VAR modelling and, above all, the use of first-differenced series instead of price levels can significantly alter the network's structure and properties. This suggests some role for real long-term linkages, arguably driven by broad and shared macroeconomic conditions, which are typically disregarded when first differencing is applied. It also points to a significant discrepancy between short-term and long-term connections, which may be difficult to detect when using lower-frequency data.

Nonetheless, some robust evidence appears to emerge. Price interdependence is widespread across the entire network, extending well beyond groups of homogeneous commodities (the abovementioned EMAFO groups). Within each of these groups, certain leader nodes seem to emerge, acting as both generators and propagators of shocks, both within their own group and across different groups. However, some nodes and groups appear to be more central to the overall network structure, with metals, more than energy commodities, standing out in this regard.

Beyond the interpretation of the results, on which further investigation is clearly needed, the proposed approach and its empirical application point to some potential methodological improvements. In particular, this study highlights two possible enhancements across both stages of the analysis. In the first stage, alternative modelling approaches to the VAR framework adopted here, such as wavelet

[11] These further GCN variants are available upon request.

analysis or machine learning techniques, may help refine the identification of connections across nodes, especially under nonlinear multivariate dynamics. The second stage could benefit from a more refined definition and analysis of the network, particularly by leveraging the different timing of price interdependence as captured by Granger causality. This could lead to the construction of a weighted and/or dynamic network. Both solutions may significantly enhance the resulting network analysis, although they would also introduce substantial computational complexity.

Besides the main methodological focus of the study, some policy considerations can also be drawn from the approach and results presented. The strategic relevance of certain commodities, combined with their increasing interdependence and the resulting rise in market instability, makes it urgent for policy decision-makers to be equipped with real-time surveillance or early-warning tools. These tools should be capable of identifying those prices whose movements may have a major systemic impact, potentially triggering widespread systemic risks. The approach proposed in this study, along with its possible future enhancements, may help to identify and continuously update the set of prices that behave like 'canaries in a coal mine,' signalling the onset of generalized commodity price shocks and, consequently, major inflationary pressures. This type of critical information could support early, targeted policy interventions in the most vulnerable markets, aimed at stabilizing them promptly and thereby preventing or mitigating the amplification of shocks across the entire network.

Table 1 – Pairwise Pearson correlations across GCN variants.

	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.

Table 2 – Network-wide metrics across GCN variants compared to the corresponding geodesic network.

	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

Table 3 - Dyadic configurations across GCN variants.

	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.

	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.

Table 5 - In-degree and out-degree global and local bridges across GCN variants.

	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),	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

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

	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**
	Random network	Ring network	Small network	Lattice network	Geodesic network
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.

Table 7 – Node-level clustering coefficients across GCN variants (maximum and minimum values 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

Table 7 (continued)

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

Table 8 – Node-level betweenness centrality across GCN variants (maximum and minimum values 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

Table 8 (continued)

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

Table 9 – Node-level closeness centrality across GCN variants (maximum and minimum values 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

Table 9 (continued)

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 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

Table 10 (continued)

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

Table 11 – Community structure across GCN variants: within-group connectivity and modularity

	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

Table 11 (continued)

	Energy	Metals	Agriculture	Food	Other raw materials
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

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)
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.

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

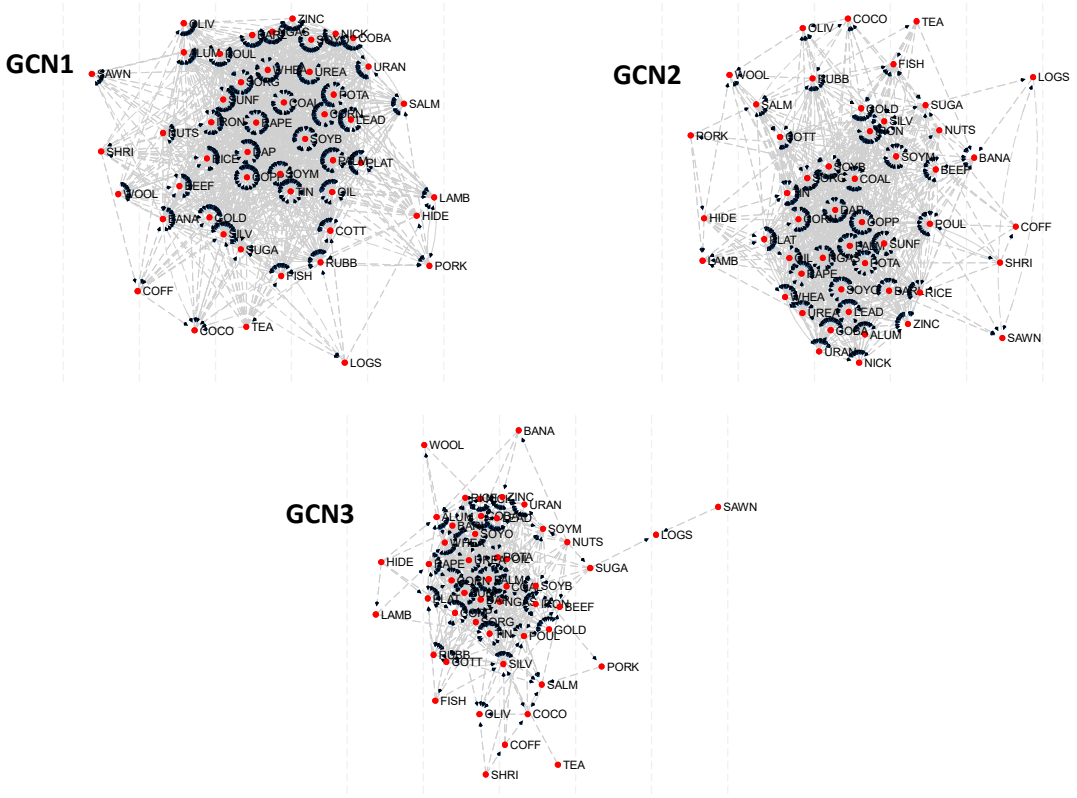


Figure 2 – Network structure GCN1-GCN3 – price first differences (dashed arcs indicate links across commodities and arrows indicate the direction).

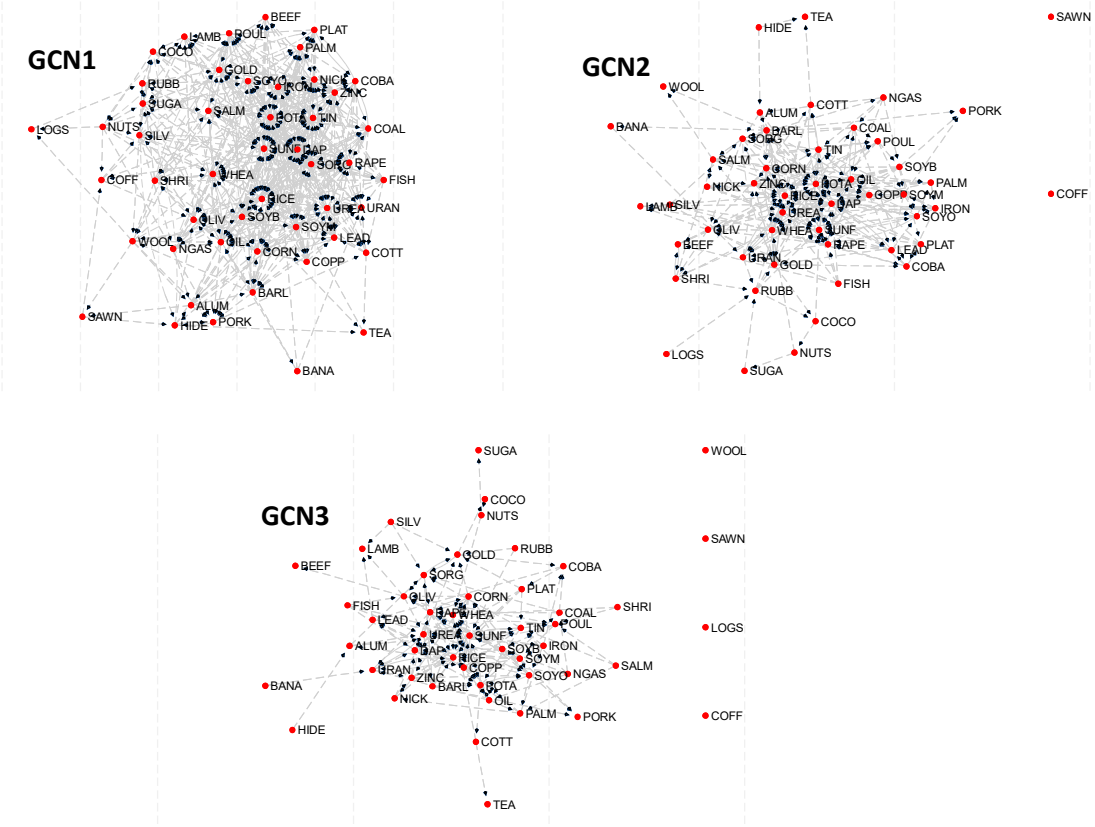


Figure 3 – Network structure GCN4-GCN5 – price levels (dashed arcs indicate links across commodities and arrows indicate the direction).

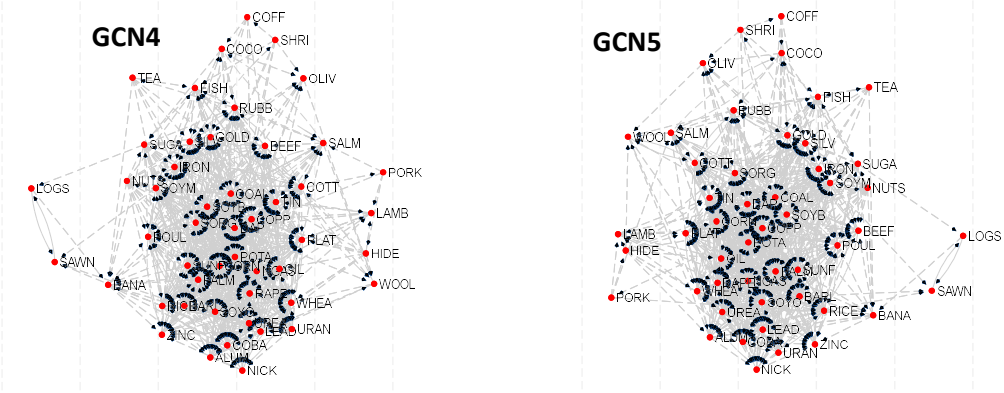


Figure 4 – Network structure GCN4-GCN5 – price first differences (dashed arcs indicate links across commodities and arrows indicate the direction).

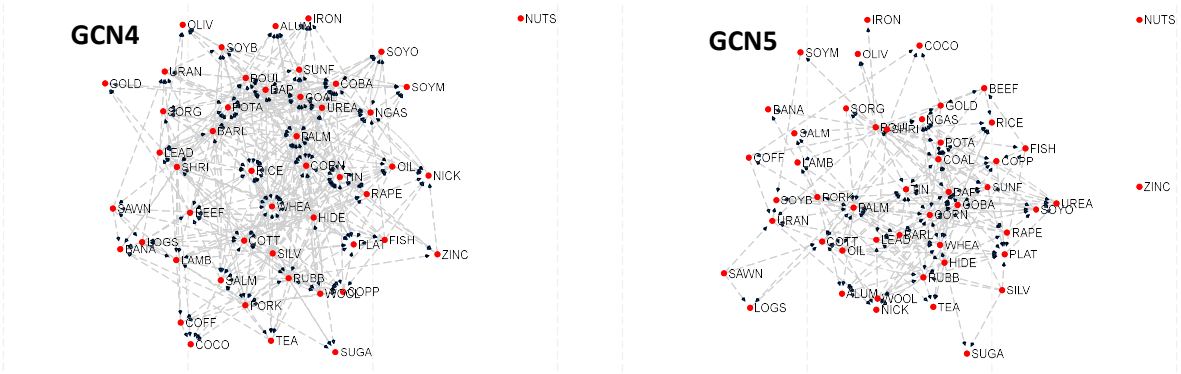


Figure 5 – Number of arcs per node in GCN3 by group and in descending order (price levels).

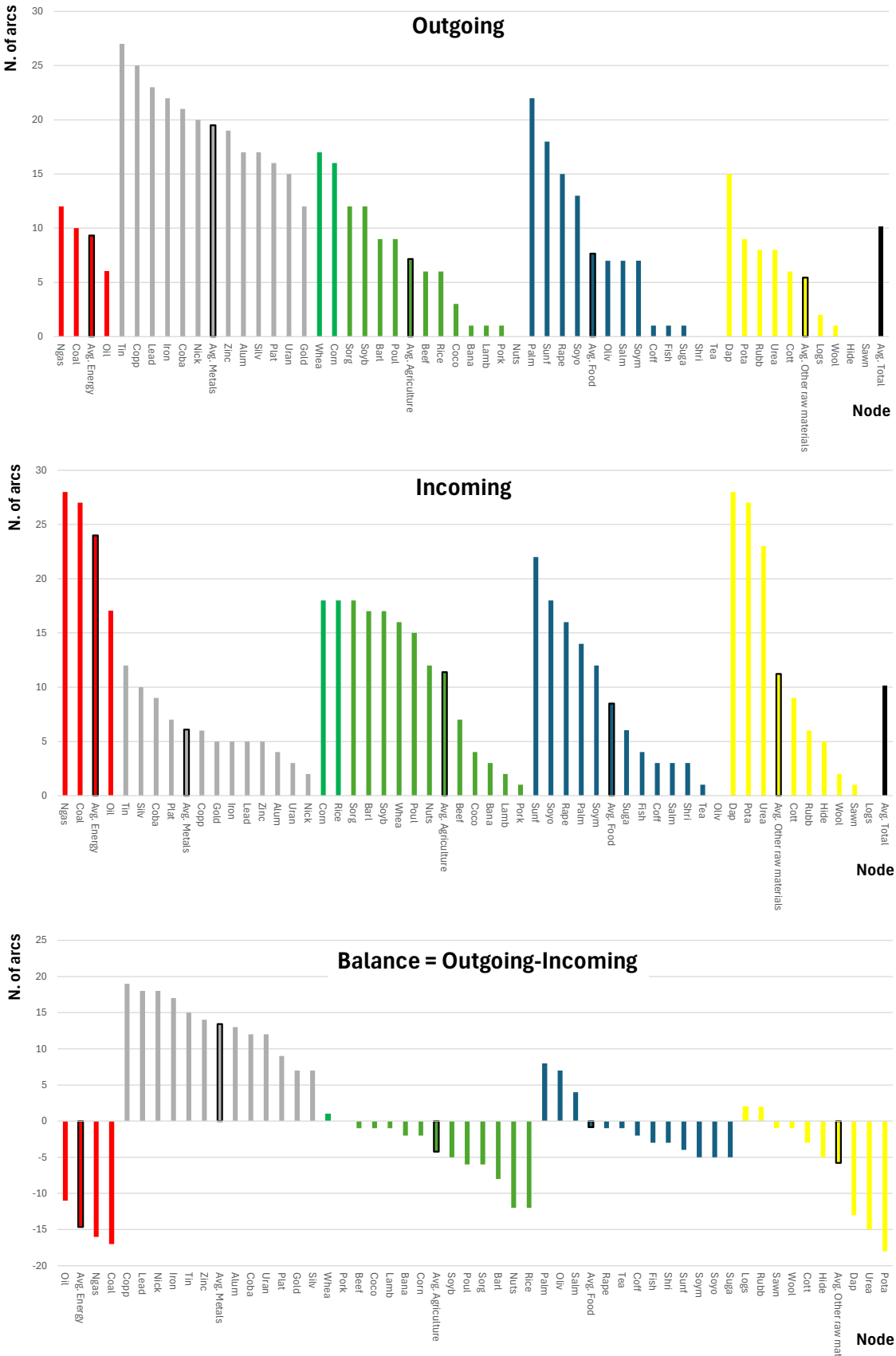


Figure 6 – Number of arcs per node in GCN5 by group and in descending order (price levels).

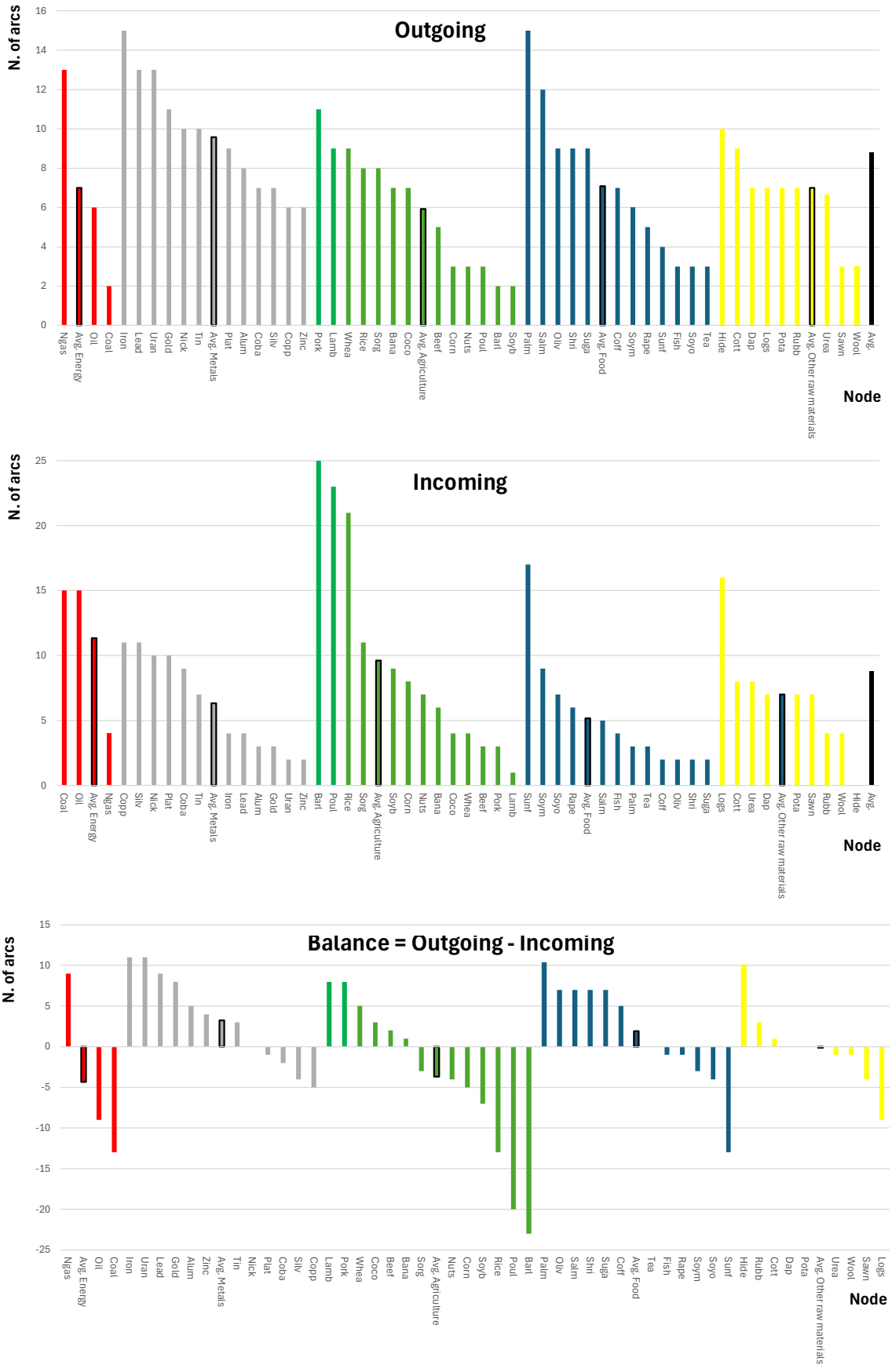


Figure 7 – Number of arcs per node in GCN3 by group and in descending order (price first differences).

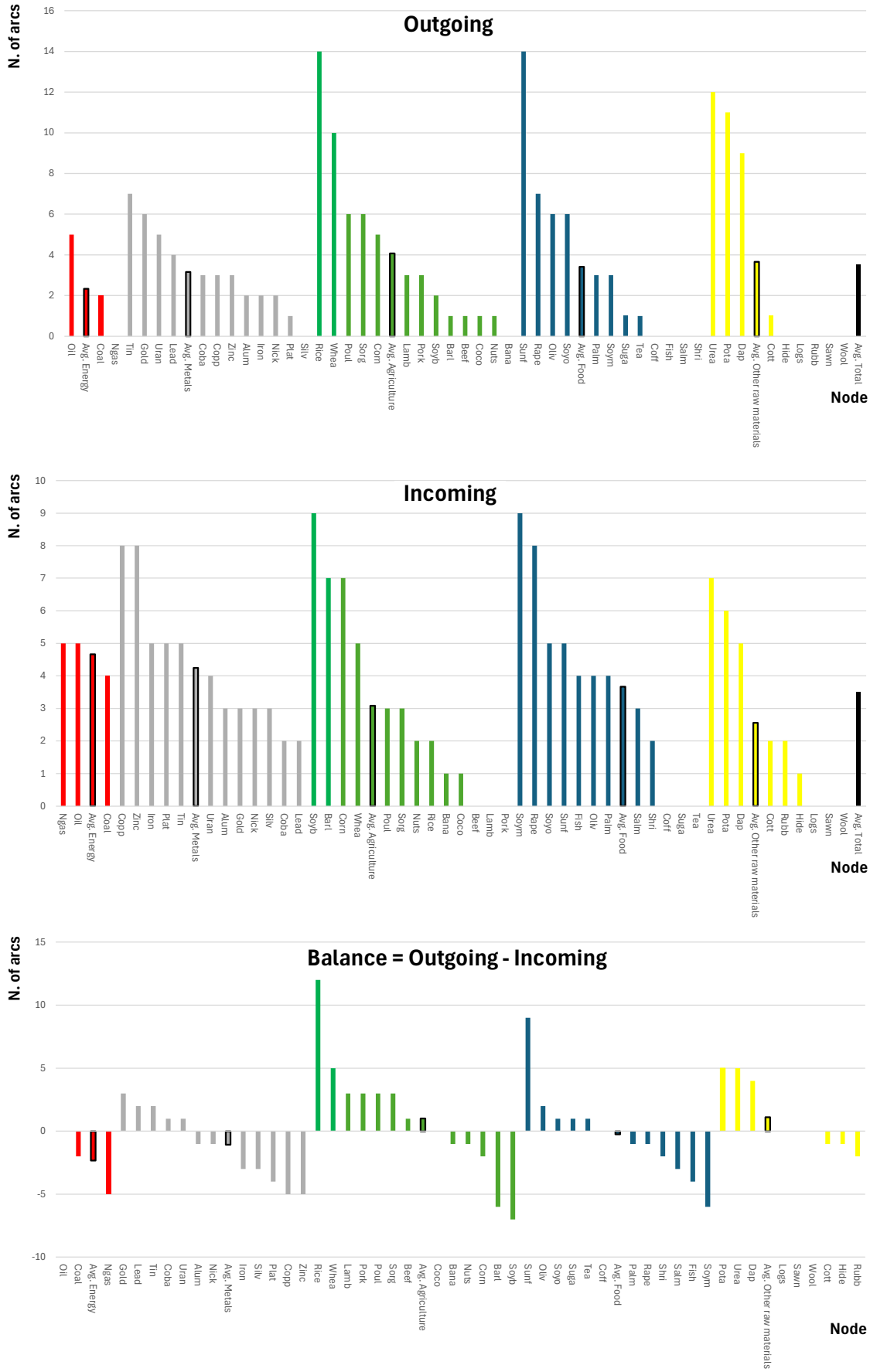
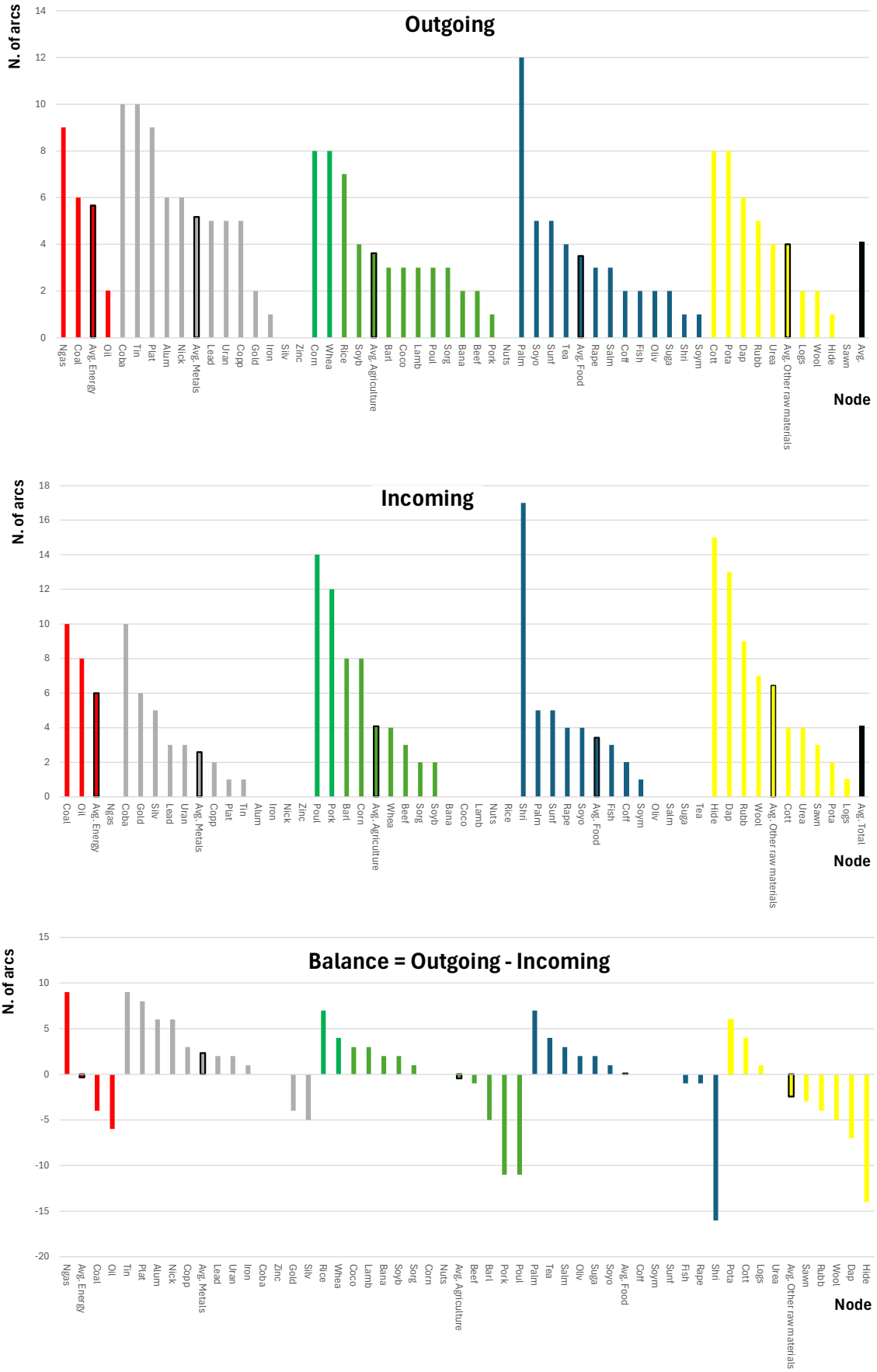


Figure 8 – Number of arcs per node in GCN5 by group and in descending order (price first differences).



REFERENCES

Acemoglu, D., Carvalho, V.M., Ozdaglar, A., Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica* 80(5): 1977–2016.

Adeosun, O.A., Olayeni, R.O., Tabash, M.I.,, Anagreh, S. (2023). Revisiting the Oil and Food Prices Dynamics: A Time Varying Approach. *Journal of Business Cycle Research* (in press). Available from: https://doi.org/10.1007/s41549-023-00083-3.

Aharon, D.Y., Aziz, M.I.A., Kallir, I. (2023). Oil price shocks and inflation: A cross-national examination in the ASEAN5+3 countries. *Resources Policy*, 82, 103573.

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. *The North American Journal of Economics and 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.

Banbura, M., Giannone, D., Reichlin, L. (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics* 25(1): 71-92.

Basu, S., Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models. *The Annals of Statistics* 43(4): 1535-1567.

Baum, C.F. (2005). Stata: The language of choice for time-series analysis? *The Stata Journal*, 5 (1), 46–63.

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

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 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. *Quarterly Journal of Economics* 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 Economics* 91, 104884.

Boakye, E.O., Heimonen, K., Junttila, J. (2024). Commodity markets and the global macroeconomy: evidence from machine learning and GVAR. *Empirical Economics* 67: 1919-1965.

Borgatti, S.P., Everett, M.G., Johnson, J.C. (2018). *Analyzing Social Networks* (2nd ed.). SAGE Publications, London.

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, 55-67.

Byrne, J.P., Sakemoto, R., Xu, B. (2020). Commodity price co-movement: heterogeneity and the time-varying impact of fundamentals. *European Review of Agricultural Economics*, 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. *Scientific Reports* 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 Economics* 30(6): 3063-3076.

Christensen, D.A., Schatzer, R.J., Heady E.O., English B.C. (1981). The effects of increased energy prices on U.S. agriculture: an econometric approach. CARD report No. 104. Center for Agricultural and Rural Development, Iowa State University.

Chudik, A., Pesaran, M.H. (2011). Infinite-dimensional vars and factor models. *Journal of Econometrics* 163(1): 4-22.

Corradi, V., Swanson, N.R. (2006). The effect of data transformation on common cycle, cointegration, and unit root tests: Monte Carlo results and a simple test. *Journal of Econometrics*, 132, 195–229.

Crain, S.J., Lee, H.J. (1996). Volatility in Wheat spot and Future Markets, 1950-1993: Government Farm Programs, Seasonality, and Causality. *The Journal of Finance*, 51, 325-343.

Devlin, W., Woods, S., Coates, B. (2011). Commodity price volatility, *Economic Roundup*, 1, 1-12.

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. *Journal of Econometrics* 182(1): 119-134.

Ding, S., Cui, T., Zheng, D., Du, M. (2021). The effects of commodity financialization on commodity market volatility. *Resources Policy*, 73, 102220.

Dolado, J.J., Lütkepohl, H. (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews* 15, 369-386.

Drachal, K., Pawłowski, M. (2024). Forecasting Selected Commodities' Prices with the Bayesian Symbolic Regression. *International Journal of Financial Studies* 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. *Journal of Econometrics* 154: 42–58.

Dvoskin, D., Heady, E.O. (1977). Commodity Prices and Resource Use Under Various Energy Alternatives in Agriculture. *Western Journal of Agricultural Economics*, 2, 53–62.

Eichler, M. (2012a). Causal inference in time series analysis. In: C. Berzuini, P. Dawid, Bernardinelli, L. (Eds.), *Causality: Statistical Perspectives and Applications*, John Wiley & Sons, Hoboken (NJ), pp. 327–354.

Enders, W. (2014). *Applied Econometric Time Series* (4th ed.). John Wiley & Sons. Hoboken (NJ).

Esposti, R. (2021). On the long-term common movement of resource and commodity prices. A methodological proposal. *Resources Policy* 72, 102010. DOI: 10.1016/j.resourpol.2021.102010.

Esposti, R. (2024a). Dating common commodity price and inflation shocks with alternative approaches. *Bio-based and Applied Economics* 13(2): 171-201. DOI:10.36253/bae-14060.

Esposti, R. (2024b). Who moves first? Resource Price Interdependence through time-Varying Granger Causality. *Natural Resource Modeling* 37(3), e12396. DOI: 10.1111/nrm.12396.

Esposti, R., Listorti, G. (2013). Agricultural Price Transmission across Space and Commodities during Price Bubbles. *Agricultural Economics*, 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) 2018/858, (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 or http://dx.doi.org/10.2139/ssrn.4557222.

Garzón, A.J., Hierro, L.A. (2022). Inflation, oil prices and exchange rates. The Euro's dampening effect. *Journal of Policy Modeling*, 44 (1), 130-146.

George, E.I., Sun, D., Ni. S. (2008). Bayesian stochastic search for var model restrictions. *Journal of Econometrics* 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.

Glynn, J., Perera, N., Verma, R. (2007). Unit-root tests and structural breaks: a survey with applications. *Journal of Quantitative Methods for Economics and Business Administration*, 3 (1), 63–79.

Hong, Y., Liu, Y., Wang, S. (2009). Granger causality in risk and detection of extreme risk spillover between financial markets. *Journal of Econometrics* 150(2): 271-287.

Kirikkaleli, D., Güngör, H. (2021). Co-movement of commodity price indexes and energy price index: a wavelet coherence approach. *Financial Innovation* 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. *Decisions in Economics and 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. *Journal of Revenue and Pricing Managemen*t: 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 Papers on Economic Activity* 2: 1-78.

Listorti, G., Esposti, R. (2012). Horizontal price transmission in agricultural markets: fundamental concepts and open empirical issues. *Bio-based and Applied Economics*, 1(1), 81-108.

Litterman, R.B. (1986). Forecasting with Bayesian vector autoregressions-five years of experience. *Journal of Business & Economic Statistics* 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. (2005). *New Introduction to Multiple Time Series Analysis*. Springer, Heidelberg.

Lütkepohl, H.. (1982). Non-causality due to omitted variables. *Journal of Econometrics* 19(2-3): 367-378.

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. Resources 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. *Scientific Reports* 4, 3655.

Miljkovic, D., Vatsa, P. (2023). On the linkages between energy and agricultural commodity prices, a dynamic time warping analysis. *International Review of Financial Analysis*, 90, 102834.

Morana, C. (2012). PC-VAR Estimation of Vector Autoregressive Models. *Open Journal of Statistics*, 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. *Journal of Climate* 19(7): 1182-1194.

Mutascu, M.I., Albulescu, C.T., Apergis, N., Magazzino, C. (2022). Do gasoline and diesel prices co-move? Evidence from the time-frequency domain. *Environmental Science and Pollution Research*, 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. *Journal of Agricultural and Applied Economics* 52(1):117-134.

OECD (2010). Developments in commodity price volatility. Working Party on Agricultural Policies and Markets, Trade and Agriculture Directorate, OECD, Paris.

Otero, J., Baum, C. (2021). Unit-root tests for explosive behavior. *The Stata Journal*, 21(4), 999–1020.

Pedace, R. (2013). *Econometrics for Dummies*. Hoboken, N.J.: John Wiley & Sons.

Peterson, H.H., Tomek, W.G. (2000). Implications of deflating commodity prices for time-series analysis. In: NCR-134 Conference on Applied Commodity Price Analysis. Forecasting, and Market Risk Management. Chicago, April 17-18.

Phillips, P.C.B., Shi, S. (2020). Real time monitoring of asset markets: bubbles and crises. In H. D. Vinod and C. R. Rao (Eds.), *Handbook of Statistics: Financial, Macro and Micro Econometrics Using R*. Volume 42. Amsterdam: Elsevier, 61-80.

Phillips, P.C.B., Shi, S., Yu, J. (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4), 1043-1077.

Phillips, P.C.B., Wu, Y., Yu, J. (2011). Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? *International Economic Review*, 52(1), 201-226.

Piot-Lepetit, I., M'Barek, R. (eds.) (2011). Methods to Analyse Agricultural Commodity Price Volatility. New York: Springer.

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.

Roch, F. (2019). The adjustment to commodity price shocks. *Journal of Applied Economics*, 22(1), 437–467.

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.

Sands, R., Westcott, P. (coordinators), Price, J.M., Beckman, J., Leibtag, E., Lucier, G., McBride, W., McGranahan, D., Morehart, M., Roeger, E., Schaible, G., Wojan, T.R. (2011). Impacts of Higher Energy Prices on Agriculture and Rural Economies. ERR- 123, U.S. Department of Agriculture, Economic Research Service, Washington D.C..

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. *Journal of Neuroscience* 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. *Resources Policy*, 74, 102298.

Shi, S., Hurn, S. , Phillips, P.C.B. (2020). Causal change detection in possibly integrated systems: Revisiting the money-income relationship. *Journal of Financial Econometrics*, 18 (1), 158–180 .

Shi, S., Phillips, P.C.B., Hurn, S. (2018). Change detection and the causal impact of the yield curve. *Journal of Time Series Analysis*, 39 (6), 966–987.

Shojaie, A., Fox, E.B. (2021). Granger causality: A review and recent advances. *arXiv: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. *Journal of the American Statistical Association* 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. *Physica A: Statistical Mechanics and its Applications* 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

Thoma, M. A. (1994). Subsample instability and asymmetries in money-income causality. *Journal of Econometrics*, 64, 279-306.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1): 267–288.

Toda, H.Y., Phillips, P.C.B. (1994). Vector autoregression and causality: a theoretical overview and simulation study. *Econometric Reviews*, 13, 259-285.

Uematsu, Y., Yamagata, T. (2025). Discovering the Network Granger Causality in Large Vector Autoregressive Models. *Journal of the American Statistical Association*, DOI: 10.1080/01621459.2025.2450836.

van Garderen, K. J. (2023). Forecasting Levels in Loglinear Unit Root Models. *Econometric Reviews* 42(9–10): 780–805.

Vatsa, P., Miljkovic, D. Baek, J. (2023). Linkages between natural gas, fertiliser and cereal prices: A note. *Journal of Agricultural Economics*, 74, 935–940.

Wang, D., Tomek, W.G. (2007). Commodity Prices and Unit Root Tests. *American Journal of Agricultural Economics*, 89 (4), 873-889.

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 Research Letters* 39: 101662.

Watts, D.J. (2004). The 'New' Science of Networks. *Annual Review of Sociology* 30: 243–270.

Yamada, H., Toda, H. Y. (1998). Inference in possibly integrated vector autoregressive models: Some finite sample evidence. *Journal of Econometrics*, 86, 55-95.

Yuan, M., Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68(1):49-67.

Zhang, D., Broadstock, D.C. (2020). Global financial crisis and rising connectedness in the international commodity markets. *International Review of Financial Analysis* 68, 101239.

Zhao, Z., Wen, H., Li, K. (2021). Identifying bubbles and the contagion effect between oil and stock markets: New evidence from China. *Economic Modelling*, 94, 780–788.

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.

ANNEX 1

Table A1 – Description of the commodity data used in the analysis (source: IMF).

Commodity Full Name	Code	Description	Unit and Market	Category
Coal	Coal	Australian thermal coal, 6,300	USD/metric ton, Australia	Energy
Natural Gas	Ngas	US Henry Hub natural gas	USD/MMBtu, US	Energy
Crude Oil	Oil	Brent crude oil	USD/barrel, Europe	Energy
Aluminum	Alum	LME aluminium, high grade	USD/metric ton, LME	Metals
Cobalt	Coba	LME cobalt, standard grade	USD/metric ton, LME	Metals
Copper	Copp	LME copper, grade A	USD/metric ton, LME	Metals
Gold	Gold	London gold fix	USD/troy ounce, London	Metals
Iron Ore	Iron	Iron ore fines, 62% Fe, CFR China	USD/dry metric ton, China	Metals
Lead	Lead	LME lead, standard grade	USD/metric ton, LME	Metals
Nickel	Nick	LME nickel, primary grade	USD/metric ton, LME	Metals
Platinum	Plat	London platinum fix	USD/troy ounce, London	Metals
Silver	Silv	London silver fix	USD/troy ounce, London	Metals
Tin	Tin	LME tin, standard grade	USD/metric ton, LME	Metals
Uranium	Uran	U3O8 uranium concentrate	USD/pound, Global	Metals
Zinc	Zinc	LME zinc, standard grade	USD/metric ton, LME	Metals
Bananas	Bana	Bananas, Central America	USD/metric ton, US	Agriculture
Barley	Barl	Feed barley	USD/metric ton, Global	Agriculture
Beef	Beef	Australian beef	USD/kg, Australia	Agriculture
Cocoa	Coco	Cocoa beans	USD/metric ton, Global	Agriculture
Corn	Corn	Yellow corn	USD/metric ton, US	Agriculture
Lamb	Lamb	New Zealand lamb	USD/kg, New Zealand	Agriculture
Nuts	Nuts	Tree nuts	USD/metric ton, Global	Agriculture
Pork	Pork	US pork	USD/kg, US	Agriculture
Poultry	Poul	US broiler chicken	USD/kg, US	Agriculture
Rice	Rice	Thai 5% broken rice	USD/metric ton, Thailand	Agriculture
Sorghum	Sorg	US sorghum	USD/metric ton, US	Agriculture
Soybeans	Soyb	US soybeans	USD/metric ton, US	Agriculture
Wheat	Whea	US hard red winter wheat	USD/metric ton, US	Agriculture
Coffee	Coff	Arabica coffee	USD/kg, Global	Food
Fish	Fish	Fishmeal	USD/metric ton, Peru	Food
Olive Oil	Oliv	Extra virgin olive oil	USD/metric ton, Europe	Food
Palm Oil	Palm	Malaysian palm oil	USD/metric ton, Malaysia	Food
Rapeseed Oil	Rape	Canadian rapeseed oil	USD/metric ton, Canada	Food
Salmon	Salm	Norwegian salmon	USD/kg, Norway	Food
Shrimp	Shri	Asian white shrimp	USD/kg, Asia	Food
Soybean Meal	Soym	Soybean meal, 48% protein	USD/metric ton, US	Food
Soybean Oil	Soyo	US soybean oil	USD/metric ton, US	Food
Sugar	Suga	World raw sugar	USD/kg, Global	Food
Sunflower Oil	Sunf	Sunflower oil, crude	USD/metric ton, Ukraine	Food
Tea	Tea	Average tea price	USD/kg, Global	Food
Cotton	Cott	Cotton A Index	USD/kg, Global	Other Raw Materials
Diammonium Phosphate	Dap	DAP fertilizer	USD/metric ton, Global	Other Raw Materials
Hides	Hide	Heavy steer hides	USD/piece, US	Other Raw Materials
Logs	Logs	Tropical logs	USD/cubic meter, Malaysia	Other Raw Materials
Potash	Pota	Potassium chloride	USD/metric ton, Global	Other Raw Materials
Rubber	Rubb	RSS3 rubber, Bangkok	USD/kg, Thailand	Other Raw Materials
Sawnwood	Sawn	Sawnwood, Cameroon	USD/cubic meter,	Other Raw Materials
Urea	Urea	Urea fertilizer, bulk	USD/metric ton, Global	Other Raw Materials
Wool	Wool	Australian wool, 64s	USD/kg, Australia	Other Raw Materials

ANNEX 2

Table A2 – P-value of the unit root (ADF) tests on the commodity price levels, first differences, logarithms and logarithms of the first differences (1980M1-2024M12) (in bold p-values>0.1).

	Price Levels	Price First Differences	Logarithm of Price Levels	Logarithm of Price First Differences
Coal	0.002	0.000	0.028	0.000
Ngas	0.000	0.000	0.006	0.000
Oil	0.045	0.000	0.080	0.000
Alum	0.002	0.000	0.002	0.000
Coba	0.000	0.000	0.001	0.000
Copp	**0.154**	0.000	**0.146**	0.000
Gold	**0.936**	0.000	**0.521**	0.000
Iron	0.048	0.000	0.084	0.000
Lead	0.068	0.000	0.091	0.000
Nick	0.002	0.000	0.011	0.000
Plat	0.085	0.000	0.082	0.000
Silv	**0.170**	0.000	**0.225**	0.000
Tin	**0.110**	0.000	**0.165**	0.000
Uran	0.038	0.000	0.074	0.000
Zinc	0.002	0.000	0.014	0.000
Bana	**0.114**	0.000	**0.142**	0.000
Barl	0.003	0.000	0.010	0.000
Beef	**0.372**	0.000	**0.294**	0.000
Coco	**0.524**	0.000	**0.168**	0.000
Corn	0.010	0.000	0.022	0.000
Lamb	0.001	0.000	0.000	0.000
Nuts	0.012	0.000	0.014	0.000
Pork	0.000	0.000	0.000	0.000
Poul	0.016	0.000	0.000	0.000
Rice	0.013	0.000	0.016	0.000
Sorg	0.027	0.000	0.029	0.000
Soyb	0.026	0.000	0.029	0.000
Whea	0.003	0.000	0.004	0.000
Coff	0.034	0.000	0.011	0.000
Fish	0.086	0.000	0.071	0.000
Oliv	0.007	0.000	0.025	0.000
Palm	0.015	0.000	0.011	0.000
Rape	0.013	0.000	0.013	0.000
Salm	0.007	0.000	0.011	0.000
Shri	0.000	0.000	0.000	0.000
Soym	0.040	0.000	0.033	0.000
Soyo	0.015	0.000	0.018	0.000
Suga	0.000	0.000	0.002	0.000
Sunf	0.010	0.000	0.027	0.000
Tea	0.003	0.000	0.004	0.000
Cott	0.000	0.000	0.000	0.000
Dap	0.002	0.000	0.042	0.000
Hide	0.004	0.000	0.004	0.000
Logs	0.006	0.000	0.011	0.000
Pota	0.001	0.000	0.010	0.000
Rubb	0.009	0.000	0.028	0.000
Sawn	0.030	0.000	0.035	0.000
Urea	0.001	0.000	0.016	0.000
Wool	0.001	0.000	0.001	0.000

ANNEX 3

Table A3 – Incoming and outgoing linkages (arcs) within GCN variants – Price Levels.

	Incoming					Outgoing				
	GCN1	GCN2	GCN3	GCN4	GCN5	GCN1	GCN2	GCN3	GCN4	GCN5
Coal	36	32	27	28	15	22	17	10	21	2
Ngas	33	32	28	14	4	24	16	12	28	13
Oil	28	27	17	22	15	25	14	6	22	6
Alum	16	9	4	10	3	22	17	17	28	8
Coba	24	15	9	9	9	28	24	21	22	7
Copp	18	13	6	23	11	41	34	25	9	6
Gold	15	8	5	3	3	33	26	12	18	11
Iron	24	14	5	11	4	31	26	22	26	15
Lead	12	8	5	32	4	36	28	23	24	13
Nick	8	5	2	34	10	30	20	20	21	10
Plat	16	14	7	39	10	33	27	16	21	9
Silv	16	12	10	20	11	29	24	17	17	7
Tin	25	20	12	9	7	34	29	27	20	10
Uran	16	14	3	18	2	26	19	15	21	13
Zinc	14	8	5	3	2	29	20	19	12	6
Bana	22	7	3	20	6	20	9	1	18	7
Barl	25	22	17	27	25	24	14	9	20	2
Beef	18	11	7	5	3	25	17	6	13	5
Coco	9	8	4	6	4	8	4	3	24	7
Corn	27	22	18	25	8	32	25	16	15	3
Lamb	11	6	2	33	1	10	5	1	30	9
Nuts	26	19	12	16	7	6	1	0	18	3
Pork	8	4	1	17	3	6	1	1	31	11
Poul	28	20	15	43	23	22	18	9	17	3
Rice	35	27	18	35	21	23	15	6	27	8
Sorg	27	20	18	19	11	25	18	12	15	8
Soyb	27	20	17	10	9	26	22	12	6	2
Whea	29	21	16	21	4	30	23	17	19	9
Coff	10	5	3	5	2	3	1	1	23	7
Fish	15	11	4	45	4	17	8	1	18	3
Oliv	11	4	0	14	2	20	7	7	29	9
Palm	28	18	14	22	3	32	26	22	30	15
Rape	33	22	16	22	6	27	20	15	13	5
Salm	20	14	3	30	5	15	8	7	33	12
Shri	17	7	3	4	2	2	1	0	28	9

Table A3 (continued)

Soym	32	22	12	20	9	22	16	7	23	6
Soyo	25	22	18	19	7	27	21	13	12	3
Suga	30	18	6	21	2	8	5	1	26	9
Sunf	36	30	22	20	17	28	23	18	24	4
Tea	17	7	1	23	3	1	1	0	30	3
Cott	22	18	9	23	8	17	14	6	24	9
Dap	36	32	28	31	7	29	27	15	18	7
Hide	19	12	5	40	0	1	1	0	27	10
Logs	5	2	0	24	16	3	2	2	21	7
Pota	31	28	27	7	7	27	19	9	26	7
Rubb	21	12	6	25	4	24	14	8	19	7
Sawn	3	1	1	23	7	9	4	0	10	3
Urea	34	26	23	39	8	25	14	8	15	7
Wool	11	3	2	25	4	12	7	1	22	3
Total	1049	752	496	1034	358	1049	752	496	1034	358
Avg.	21.4	15.3	10.1	21.1	8.8	21.4	15.3	10.1	21.1	8.8
Avg. Energy	32.33	30.33	24.00	21.33	11.33	23.67	15.67	9.33	23.67	7.00
Avg. Metals	17.00	11.67	6.08	17.58	6.33	31.00	24.50	19.50	19.92	9.58
Avg. Agriculture	22.46	15.92	11.38	21.31	9.62	19.77	13.23	7.15	19.46	5.92
Avg. Food	22.83	15.00	8.50	20.42	5.17	16.83	11.42	7.67	24.08	7.08
Avg. Other raw materials	20.22	14.89	11.22	26.33	6.78	16.33	11.33	5.44	20.22	6.67

Legend: GCN1= PairwiseGC (5% confidence level); GCN2 = PairwiseGC (1% confidence level); GCN3= PairwiseGC (1% confidence level+FDR); GCN4 = SparseVAR (GC – CV); GCN5) = SparseVAR (GC – BIC).

Table A4 – Incoming and outgoing linkages (arcs) within GCN variants – Price First Differences.

	Incoming					Outgoing				
	GCN1	GCN2	GCN3	GCN4	GCN5	GCN1	GCN2	GCN3	GCN4	GCN5
Coal	15	7	4	10	10	7	3	2	10	6
Ngas	7	5	5	9	0	7	1	0	12	9
Oil	13	10	5	9	8	15	8	5	5	2
Alum	9	7	3	0	0	4	2	2	9	6
Coba	9	4	2	11	10	11	4	3	13	10
Copp	10	8	8	2	2	6	4	3	9	5
Gold	14	8	3	1	6	12	6	6	3	2
Iron	14	8	5	0	0	12	3	2	7	1
Lead	12	5	2	6	3	10	5	4	8	5
Nick	11	5	3	1	0	13	4	2	9	6
Plat	11	5	5	0	1	7	2	1	13	9
Silv	9	6	3	23	5	6	0	0	0	0
Tin	11	8	5	4	1	17	7	7	17	10
Uran	5	5	4	5	3	15	6	5	6	5
Zinc	16	9	8	2	0	8	4	3	3	0
Bana	3	1	1	1	0	1	1	0	6	2
Barl	13	10	7	20	8	9	4	1	4	3
Beef	7	3	0	6	3	6	3	1	8	2
Coco	5	2	1	0	0	6	2	1	5	3
Corn	13	8	7	13	8	13	6	5	11	8
Lamb	6	1	0	7	0	8	3	3	5	3
Nuts	9	2	2	0	0	2	1	1	0	0
Pork	2	0	0	10	12	9	3	3	6	1
Poul	9	5	3	14	14	11	6	6	7	3
Rice	7	2	2	4	0	28	23	14	18	7
Sorg	16	7	3	6	2	12	6	6	5	3
Soyb	15	11	9	8	2	6	2	2	5	4
Whea	11	6	5	4	4	11	10	10	13	8
Coff	3	0	0	3	2	4	0	0	3	2
Fish	12	6	4	8	3	1	0	0	2	2
Oliv	7	5	4	0	0	13	6	6	6	2
Palm	10	5	4	11	5	11	3	3	16	12
Rape	13	9	8	3	4	15	11	7	7	3
Salm	11	7	3	5	0	7	4	0	9	3
Shri	9	5	2	17	17	6	3	0	4	1
Soym	20	10	9	1	1	7	3	3	6	1
Soyo	12	5	5	4	4	10	7	6	7	5

Table A4 (continued)

Suga	7	1	0	0	0	7	1	1	4	2
Sunf	7	6	5	5	5	26	21	14	8	5
Tea	1	0	0	3	0	3	2	1	5	4
Cott	8	4	2	6	4	7	3	1	10	8
Dap	18	6	5	17	13	23	18	9	9	6
Hide	6	2	1	24	15	4	0	0	1	1
Logs	1	1	0	7	1	2	0	0	4	2
Pota	18	8	6	5	2	25	22	11	14	8
Rubb	7	5	2	19	9	5	4	0	5	5
Sawn	2	0	0	4	3	2	0	0	3	0
Urea	9	8	7	20	4	17	12	12	6	4
Wool	9	0	0	11	7	5	2	0	3	2
Total	472	251	172	349	201	472	251	172	349	201
Avg.	9.6	5.1	3.5	7.1	4.1	9.6	5.1	3.5	7.1	4.1
Avg. Energy	11.67	7.33	4.67	9.33	6.00	9.67	4.00	2.33	9.00	5.67
Avg. Metals	10.92	6.50	4.25	4.58	2.58	10.08	3.92	3.17	8.08	4.92
Avg. Agriculture	8.92	4.46	3.08	7.15	4.08	9.38	5.38	4.08	7.15	3.62
Avg. Food	9.33	4.92	3.67	5.00	3.42	9.17	5.08	3.42	6.42	3.50
Avg. Other raw materials	8.67	3.78	2.56	12.56	6.44	10.00	6.78	3.67	6.11	4.00

Legend: GCN1= PairwiseGC (5% confidence level); GCN2 = PairwiseGC (1% confidence level); GCN3= PairwiseGC (1% confidence level+FDR); GCN4 = SparseVAR (GC – CV); GCN5) = SparseVAR (GC – BIC).

ANNEX 4

Table A5 - Topological properties across GCN variants: longest geodesic path and isolated nodes.

	Longest path	Isolated nodes
GCN1 (price levels)	Logs-Shri 3 steps; 1 connections-links (0.1%); Most: Logs (1), Shri(1)	0
GCN2 (price levels)	Oil-Logs, Nick-Logs, Plat-Logs, Tin-Logs, Uran-Logs, Bana-Coff, Corn-Logs, Lamb-Tea, Lamb-Logs, Lamb-Sawn, Prok-Coff, Pork-Shri, Pork-Sawn, Sorg-Logs, Whea-Logs, Coff-Logs, Coff-Sawn, Coff-Wool, Oliv-Logs, Salm-Logs, Shri-Logs, Hide-Logs, Salm-Sawn, Cott-Sawn, Hide-Sawn, Logs-Wool, Sawn-Wool, Urea-Sawn 3 steps; 28 connections-links (2.4%); Most: Logs (15), Sawn(8)	0
GCN3 (price levels)	Lamb-Swan, Shri-Sawn, Hide-Sawn, Sawn-Wool 5 steps; 4 connections-links (0.3%); Most: Sawn (4), Lamb, Hide, Wool (1)	0
GCN4 (price levels)	[Too many connections] 2 steps; 382 connections-links (32.5%); Most: Coff (17), Bana(16)	0
GCN5 (price levels)	Coal-Tea, Alum-Fish, Copp-Tea, Lead-Corn, Lead-Coff, Nick-Suga, Nick-Tea, Zinc-Corn, Bana-Corn, Barl-Sawn, Barl-Wool, Beef-Coff, Beef-Fish, Coco-Fish, Coco-Tea, Corn-Fish, Corn-Sawn, Fish-Suga, Fish-Hide, Oliv-Sunf, Oliv-Tea, Palm-Sawn, Rape-Tea, Rape-Sawn, Soym-Hide, Tea-Dap 3 steps; 26 connections-links (2.2%); Most: Tea (7), Fish(6)	0

Table A5 (continued)

	Longest path	Isolated nodes
GCN1 (price first differences)	Nuts-Coal, Tea-Lamb, Tea-Coal, Tea-,Ngas, Tea-Copp, Tea-Zinc, Tea-Corn, Hide-Pota, Logs-Coal, Logs-Ngas, Logs-Coba, Logs-Copp, Logs-Iron, Logs-Lead, Logs-Plat, Logs-Tin, Logs,Zinc, Logs-Bana, Logs-Beef, Logs-Corn, Logs-Cott, Logs-Dap, Logs-Hide, Pota-Hide, Sawn-Coal, Sawn-Coba, Sawn-Copp, Sawn-Iron, Sawn-Lead, Sawn-Nick, Sawn-Plat, Sawn-Tin, Sawn-Zinc, Sawn-Bana, Sawn-Beef, Sawn-Coco, Sawn-Lamb, Sawn-Cott, Urea-Logs 3 steps; 39 connections-links (3.3%); Most: Logs (16), Sawn (14)	0
GCN2 (price first differences)	Ngas-Suga, Ngas-Logs, Coba-Hide, Bana-Nuts, Bana-Tea, Coco-Tea, Lamb-Suga, Lamb-Tea, Lamb-Logs, Nuts-Pork, Nuts-Hide, Pork-Suga, Pork-Tea, Pork-Logs, Fish-Tea, Shir-Hide, Suga-Tea, Suga-Hide, Suga-Wool, Tea-Logs, Tea-Wool, Logs-Urea. 4 steps; 22 connections-links (1.9%); Most: Tea (8), Suga (6)	2 (Coff, Sawn)
GCN3 (price first differences)	Bana-Tea, Bana-Suga, Pork-Suga, Soym-Suga, Suga-Hide, Tea-Hide 5 steps; 6 connections-links (0.5%); Most: Suga (4), Bana, Hide. Tea (2)	4 (Coff, Logs, Sawn, Wool)
GCN4 (price first differences)	Oil-Coco, Alum-Coco, Coba-Coco, Copp-Uran, Gold-Soym, Gold-Suga, Gold-Tea, Gold-Sawn, Gold-Wool, Iron-Zinc, Iron-Bana, Nick-Coco, Nick-Barl, Nick-Salm, Uran-Suga, Uran-Tea, Zinc-Shri, Zinc-Soym, Zinc-Logs, Zinc-Sawn, Bana-Coco, Bana-Soyo, Barl-Suga, Coco-Soyo, Coco-Sawn, Lamb-Soym, Sorg-Suga, Soyb-Suga, Whea-Oliv, Coff-Soyo, Oliv-Wool, Soym-Sawn, Soyo-Tea, Soyo-Logs, Suga-Sunf 3 steps; 35 connections-links (3.0%); Most: Coco (7), Suga (6)	1 (Nuts)
GCN5 (price first differences)	4 steps; 7 connections-links (0.6%); Most: Iron (3), Suga (3)	2 (Zinc, Nuts)

Table A6 – Node-level farness centrality across GCN variants (maximum and minimum values 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	48	57	61	71	60	81	83	92	99	76	81
Ngas	48	60	63	70	63	80	86	104	106	80	90
Oil	48	57	66	78	61	78	73	83	96	83	91
Alum	48	65	71	81	65	87	84	90	104	88	101
Coba	48	65	68	75	72	83	82	97	112	78	83
Copp	48	**53**	58	69	67	81	83	91	92	85	93
Gold	48	60	65	87	76	83	75	85	99	**100**	94
Iron	48	58	68	75	62	77	75	93	98	90	**126**
Lead	48	58	66	75	59	83	80	94	102	84	91
Nick	48	64	74	80	55	83	78	94	106	89	100
Plat	48	60	67	80	51	79	85	95	100	82	94
Silv	48	61	66	75	68	80	83	98	125	72	103
Tin	48	55	63	70	73	79	76	86	91	75	89
Uran	48	65	73	84	67	81	79	90	101	88	96
Zinc	48	63	73	80	**82**	89	79	89	94	98	48
Bana	48	65	81	106	68	86	104	122	144	91	111
Barl	48	60	66	80	62	**70**	78	86	97	75	92
Beef	48	64	73	88	79	92	86	99	134	82	103
Coco	48	82	85	97	69	87	87	106	131	96	116
Corn	48	57	67	75	65	94	77	85	89	72	84
Lamb	48	78	88	109	54	86	85	106	119	85	102
Nuts	48	67	76	86	68	86	89	122	121	**48**	**48**
Pork	48	85	94	113	60	83	86	118	123	82	89
Poul	48	60	68	78	52	72	83	93	96	77	81
Rice	48	58	67	81	53	72	67	70	84	75	97
Sorg	48	58	68	75	65	77	75	86	97	87	98
Soyb	48	57	63	72	81	86	78	88	92	85	99
Whea	48	59	67	77	64	84	79	84	85	82	93
Coff	48	83	95	105	71	88	91	**48**	**48**	92	106
Fish	48	68	77	99	**51**	96	85	99	107	85	98
Oliv	48	69	86	95	60	88	77	91	91	93	113
Palm	48	56	62	70	56	79	81	95	107	73	82
Rape	48	56	67	81	69	87	78	84	86	85	96
Salm	48	69	79	93	55	79	79	90	119	83	102
Shri	48	79	91	118	65	85	84	103	118	76	79
Soym	48	55	67	86	62	84	75	88	93	92	121

Table A6 (continued)

Soyo	48	61	65	74	71	86	80	91	96	92	101
Suga	48	64	73	93	63	87	84	133	**164**	97	123
Sunf	48	58	62	72	61	77	69	73	82	82	91
Tea	48	78	89	122	58	**98**	103	**138**	159	91	102
Cott	48	66	72	88	61	81	85	96	116	80	86
Dap	48	54	**57**	**68**	61	86	**67**	75	87	74	79
Hide	48	76	85	104	52	88	91	132	147	71	84
Logs	48	**91**	**109**	138	59	75	**116**	136	48	87	124
Pota	48	58	62	70	69	83	67	71	87	79	88
Rubb	48	63	74	88	64	85	85	91	115	73	89
Sawn	48	85	99	**185**	69	92	112	48	48	92	122
Urea	48	59	68	75	56	82	74	79	77	73	105
Wool	48	77	90	111	62	91	84	121	48	85	96
Avg.	48	65	73	88	64	84	82	94	102	83	96
Avg. Energy	48	58	63	73	61	80	81	93	100	80	87
Avg. Metals	48	61	68	78	66	82	80	92	102	86	93
Avg. Agriculture	48	65	74	87	65	83	83	97	109	80	93
Avg. Food	48	66	76	92	62	86	82	94	106	87	101
Avg. Other raw materials	48	70	80	103	61	85	87	94	86	79	97

ANNEX 5

Figure A1 – Topology of selected synthetic networks with the 49 nodes (commodity prices).