Information Diffusion, Cluster formation and Entropy-based Network Dynamics in Equity and Commodity Markets

Stelios Bekiros and Duc Khuong Nguyen and Leonidas Sandoval Junior and Gazi Salah Uddin

IPAG Lab, IPAG Business School, Paris, France & Department of Economics, European University Institute, Florence, Italy, IPAG Lab, IPAG Business School, Paris, France & School of Public and Environmental Affairs, Indiana University, Bloomington, United States, Insper Instituto de Ensino e Pesquisa, Sao Paulo, Brasil, Department of Management and Engineering, Linköping University, Linköping, Sweden

September 2015

Information Diffusion, Cluster formation and Entropy-based Network Dynamics in Equity and Commodity Markets

Stelios Bekiros[a,b,*], Duc Khuong Nguyen[a,c], Leonidas Sandoval Junior[d], Gazi Salah Uddin[e]

[a] *IPAG Lab, IPAG Business School, Paris, France*

[b] *Department of Economics, European University Institute, Florence, Italy*

[c] *School of Public and Environmental Affairs, Indiana University, Bloomington, United States*

[d] *Insper Instituto de Ensino e Pesquisa, Sao Paulo, Brasil*

[e] *Department of Management and Engineering, Linköping University, Linköping, Sweden*

Abstract

This paper investigates the dynamic causal linkages among U.S. equity and commodity futures markets via the utilization of complex network theory. We make use of rolling estimations of extended matrices and time-varying network topologies to reveal the temporal dimension of correlation and entropy relationships. A simulation analysis using randomized time series is also implemented to assess the impact of de-noising on the data dependence structure. We mainly show evidence of emphasized disparity of correlation and entropy-based centrality measurements for all markets between pre- and post-crisis periods. Our results enable the robust mapping of network influences and contagion effects whilst incorporating agent expectations.

Keywords: Finance; commodity markets; correlation; transfer entropy; complex network; centrality
JEL Classifications: G1; G15

---

[*] ***Corresponding author***: Department of Economics, European University Institute, Via della Piazzuola 43, I-50133 Florence, Italy. Tel.: +39 0554685916 ; Fax: +39 0554685902.
E-mail addresses: stelios.bekiros@eui.eu (S. Bekiros), gazi.salah.uddin@liu.se (G.S. Uddin), duc.nguyen@ipag.fr (D.K. Nguyen), leonidassj@insper.edu.br (L. Sandoval Junior).

1. INTRODUCTION

The commodity futures markets have recently received much attention from both academics and investor community, following their increasing financialization which is characterized by an increase in trading activity as well as the number and nature of traders since the early 2000s. The 2008 Commodity Futures Trading Commission (CFTC) staff report documented an net amount of \$200 billion of commodity index investment as of June 2008, which is more than 10 times the level in 2003 and rose to about \$250 billion in 2009 (Irwin and Sanders, 2011). Major investors of commodities traded in the US markets include institutional investors, index funds, sovereign wealth funds, and retail investors holding ETFs (exchange-traded funds), ETNs (exchange-traded notes), and similar instruments. Beside their traditional role in hedging price risks of underlying commodities, commodity futures have also been increasingly viewed as a new asset class providing high potential for equity risk diversification, especially during times of financial crises and downturns in stock markets. The rationale behind this trend is that commodity futures offer high returns with relatively low volatility and low correlation with stocks and bonds.[1] Gorton and Rouwenhorst (2006) produce some stylized facts on the commodity markets after the 2000s and find that futures contracts have the same average returns as equities along with a negative correlation between bonds and equities whilst comparatively demonstrating lower volatility. Moreover, the risk factors that drive the dynamics of commodity returns may differ from those that affect stock and bond returns (e.g., Domanski and Heath, 2007; Dwyer et al., 2011). Empirical studies such as Arouri et al. (2011), Daskalaki and Skiadopoulos (2011), and Narayan et al. (2013) provide evidence of valuable diversification benefits from adding commodity assets into portfolios of traditional assets (bonds and stocks).

Recent developments in commodity futures markets witness, however, some structural shifts in their characteristics and linkages with equity markets. Cheng and Xiong (2014) document that the financialization has substantially affected commodity markets through risk sharing and information discovery mechanisms. Several studies including, among others, Büyüksahin et al. (2010), Büyüksahin and Robe (2011), Tang and Xiong (2012), and Büyükşahin and Robe (2014) uncover the increased correlation not only between commodity futures returns, but also between commodity futures and equity returns.[2] As a result of this stronger comovement, diversification benefits from

[1] The modern portfolio theory suggests that investors can improve their risk-adjusted return performance by allocating resources to imperfectly correlated assets, which is the case of equities and commodities. The reduction of regional and international diversification benefits due to higher stock market linkages and contagion risk during crisis periods (Forbes and Rigobon, 2002; Chan-Lau et al., 2004; Haldane, 2009) has caused investors to increase their holdings in alternative asset classes including commodity futures.

[2] Büyüksahin and Robe (2011) show that the equity-commodity link did not increase until 2008, but increases significantly amid the most severe episode of the global financial crisis 2008-2009. Büyükşahin and Robe (2014) reach similar conclusions based on a non-public dataset of trader positions in 17 U.S. commodity futures markets, two U.S. equity market indices (S&P's 500 and Dow Jones' Industrial Average Index), and the MSCI World equity market index. Tang and Xiong (2012) find increasing correlation between non-energy commodity futures and oil futures returns, which implies that prices of individual commodities are no longer driven by their own supply and demand conditions.

the inclusion of commodity futures into stock portfolios have been found to decrease, particularly during crisis periods where they are needed (Silvennoinen and Thorp, 2010; Daskalaki and Skiadopoulos, 2011). Large fluctuations in commodity prices over recent years have also been causes for concern among governments, policymakers and traders[3], which have significant impacts (either positive or negative) on equity markets (e.g., Baur and McDermott, 2010; Narayan and Sharma, 2011; Filis et al., 2011). Moreover, while the expansion of financial institutions' positions in commodity markets may improve risk sharing, it can increase the importance of common shocks and spark off substantial volatility spillover and contagious effects in case of important financial distresses due to limits to arbitrage as reported in Gromb and Vayanos (2010). For instance, the results of Adams and Glück (2015) suggest that volatility transmission between commodity and stock markets will remain high since commodities are now an investment style of professional investors.

It turns out from the above discussions that commodity futures and equity markets have become more interconnected to each other over time. If they form a financial network, their interactions could lead to the establishment of contagion channels and consequently amplify the shocks to the financial system. In their seminal study, Allen and Gale (2000) describe financial markets as complex networks and model financial contagion as an equilibrium phenomenon whereby liquidity shocks may spread through the network and contagious effects between financial institutions are smaller in a complete network structure than in an incomplete one. The complex dynamics of clusters and networks in financial markets have made them less diverse in the years that led to the financial crisis of 2008-2009, which thus witnesses the need for empirical research on financial networks, particularly after the global financial crisis (Allen and Babus, 2009). Upper (2011) provides a comprehensive survey on modern techniques and simulation methods to study contagion in financial networks. Other relevant studies on network modeling and contagion topic include Nagurney and Ke (2006), Barro and Basso (2010), and Veremyev and Boginski (2012).

The present study builds upon the works of Sandoval (2013, 2014) to investigate the dynamic financial networks and information transmission between commodity futures and equity markets. In Sandoval (2013), the temporal evolution of financial market clusters and networks is investigated based on the correlations between stock market indices with the aim to construct asset graphs in diverse time periods. Sandoval (2014) uses the measure of transfer entropy as opposed to correlation to examine the causal relationships among stocks of the 197 largest financial companies worldwide. Differently, we aim to illustrate two different types of networks (weighted and unweighted) based on daily data of the U.S. stock markets, represented by the S&P500 index and by ten S&P500 stock sector indices, and twelve liquid commodity futures grouped into three categories

[3] Oil price peaked at over US$140 per barrel, after reaching the US$100 per barrel mark for the first time in 2008, while gold prices have quadrupled over the 2001–2010 period (Baur and McDermott, 2010). Agricultural commodity prices based on the Chicago cash corn price rose to over $3.00/bushel to reach $7.20/ bushel in July 2008.

(energy, metals, and agriculture). These commodity futures are largely traded in the New York Mercantile Exchange (NYMEX) and the Chicago Board of Trade (CBOT).

Our work contributes to the relevant literature in four main aspects. First, it examines the financial networks of commodity futures and stock markets over the last 20 years, which covers several episodes of major financial crises and market crashes. Second, it introduces the measure of transfer entropy for analyzing the co-movements, dynamic causalities and clustering structure of the examined networks. This particular concept was developed in information science and is utilized to track the transfer/diffusion of information between commodity and stock markets in the first and second moment (volatility). Third, the dynamics of the complex equity-commodity networks are analyzed through the temporal evolution of the correlations and entropy. Finally, we attempt to reveal and rationalize the potential effects of commodity financialization process on the equity-commodity synchronization and their importance for investors, speculators, market makers and fund managers.

The remainder of this paper is organized as follows. Section 2 presents the novel methodology utilized to explore the equity-commodity market links via the use of dynamic networks and various centrality measures. Section 3 describes the data. Section 4 reports and discusses the empirical results in terms of asset graphs, causal interdependencies and temporal network dynamics. Section 5 provides the concluding remarks.

2. METHODOLOGY

2.1 Correlation and Transfer Entropy

The standard measure of relationship between variables is the well-known Pearson correlation coefficient. For two variables e_i and c_i, $i = 1,\dots,n$ with n sample size, it is defined as follows:

$$\rho_{ij} = \frac{\sum_{i=1}^{n}(e_i - \bar{e})(c_i - \bar{c})}{\sqrt{\sum_{i=1}^{n}(e_i - \bar{e})^2}\sqrt{\sum_{i=1}^{n}(c_i - \bar{c})^2}} \qquad (1)$$

where $\bar{e}$ is the average of e and $\bar{c}$ is the average of c. Although there are nonlinear measures such as the Spearman and the Kendall rank correlations, the results obtained are not substantially different for networks of financial variables as reported in Sandoval (2013). In our work, correlations between equity and commodity returns will be estimated over the whole period as well as in rolling samples in order to uncover possible differences in "normal" and "crisis" times. However, since correlations are unable to reveal directional causality relationships or the magnitude of impacts owing to the assumption of linearity and symmetry, we introduce a dynamic and non-symmetric measure called *Transfer Entropy*. The latter was developed by Schreiber (2000) and is based on the concept of *Shannon Entropy* as defined in information theory (Shannon, 1948).

Transfer entropy has been widely used in a variety of scientific fields such as the cellular automata in computer science (e.g., Lizier and Mahoney, 2013), the brain neural cortex in medicine (e.g., Vicente et al., 2011, Faes *et al.*, 2013), social networks (e.g,, Ver Steeg and Galstyan, 2012), causal influences and applied statistics (e.g., Barnett, 2009, Amblard and Michel, 2013, Liu et al., 2014), and in dynamical systems (e.g., Nichols et al., 2013, Prokopenko et al., 2013). Applications of transfer entropy to the analysis of financial market include, among others, Marschinski and Kantz (2002), Kwon and Yang (2008), Altiparmak and Dengiz (2009), Dimpfl and Peter (2012), Sandoval (2014), and Sandoval et al. (2015).

More formally, when the time series of variable X is a Markov process of k degree, the state i_{n+1} of X depends on the k previous states of the same variable i.e.,

$$p(i_{n+1}|i_n, i_{n-1}, \cdots, i_0) = p(i_{n+1}|i_n, i_{n-1}, \cdots, i_{n-k+1}) \qquad (2)$$

where $p(A|B)$ is the conditional probability of A given B, defined as $p(A|B) = p(A,B)/p(B)$. Hence, the conditional probability of state i_{n+1} of variable X on all its previous states is the same as the conditional probability of i_{n+1} on its k previous states. If variable X depends on another variable Y, we assume that the state i_{n+1} of X depends on the previous states j_n of variable Y. The Transfer Entropy from Y to X is defined as the "average information" contained in the *source* Y about the next state of the *destination* X that was not already contained in the destination's past. We assume that each element (observation) i_{n+1} of the time series X is influenced by the k previous states of the same variable and by ℓ previous states of Y, as depicted in Figure 1. The values of k and ℓ may vary, according to the data and the state "memory" used to analyze the transfer of entropy from one variable to another. The Transfer Entropy from Y to X is defined as

$$TE_{Y\to X}(k,\ell) = \sum_{i_{n+1}, i_n^{(k)}, j_n^{(\ell)}} \left[p\left(i_{n+1}, i_n^{(k)}, j_n^{(\ell)}\right) log_2 \frac{p\left(i_{n+1}|i_n^{(k)}, j_n^{(\ell)}\right)}{p\left(i_{n+1}|i_n^{(k)}\right)} \right] \qquad (3)$$

where i_n is the n element of X series, j_n of variable Y, $p\left(A,B\right)$ the joint probability of A and B and

$$p\left(i_{n+1}, i_n^{(k)}, j_n^{(\ell)}\right) = p(i_{n+1}, i_n, \dots, i_{n-k+1}, \dots, j_{n-\ell+1}) \qquad (4)$$

Eq. (4) shows the joint probability of state i_{n+1} with its $k+1$ predecessors, and predecessors of state j_n. This definition of Transfer Entropy assumes that events on a certain day may be influenced by events of k and previous days. We may safely assume based on the efficient market hypothesis and random walk behavior of stock prices that a short memory $k = \ell = 1$ i.e., previous day is significant, but this will be verified later by also calculating the case where $k = \ell = 2$. The calculation of Transfer Entropy may involve significant computational burden, particularly when the number of variables is high. Moreover, the measure is model-independent and asymmetric, and reveals directional causality relationships. It can be seen as a nonlinear version of the Granger causality, which is reduced to simple linear Granger causality in case of vector auto-regressive processes (Barnett, 2009). Similarly to the calculation of correlations, Transfer Entropy will be estimated over the whole period as well as in moving window samples.

[Please insert Figure 1]

2.2 Centrality measurement

On the basis of large matrices of correlations among assets or Transfer Entropy, a complex network can be formed and the corresponding nodes and edges can be specified. The edges represent the relationships between two nodes (e.g., commodity futures and equity indices). There exist two types of networks: weighted and unweighted. In weighted networks, all nodes are inter-connected and each edge is assigned a certain degree of connection intensity, i.e., each edge would comprise one element of the correlation or Transfer Entropy matrix. The unweighted network contains edges between nodes or not, whilst no intensity measure is assigned. Such networks may be obtained from large matrices by establishing thresholds above or below which edges are formed. For instance, if we consider the correlation matrix of all commodity futures examined in this study and filter out the correlations above a certain value *T,* which are represented as edges in the network, the rest of the relationships based on correlations below this value are ignored. Both representations incorporate different ways to measure the *centrality* of a node (Newman, 2010; Gómez et al., 2013). Centrality is an important concept in network theory, yet there is no unique definition. In financial networks, a node that is central may be important in the propagation or in the aversion of crises. Next, we describe the types of centrality measures utilized in this paper.

Networks can be directed or undirected, depending on the edges between them. If the edges w_{ij} and w_{ji} between two nodes i and j are always the same, then the network is undirected. Otherwise, it is directed. The networks obtained with correlation are undirected, while networks obtained with TE or ETE are directed. For undirected networks, one of the most important centrality measures is the so-called Node Degree (*ND*), which represents the total number of edges through which a node is connected with other nodes. This measure is extensively used in asset graphs where some nodes are not inter-connected to each other and varies according to the choice of the threshold as in Sandoval (2013). Another measure is Eigenvector Centrality (*EC*), which takes into account, aside from how many connections a node has, whether it is situated in a region of highly connected nodes. Moreover, Closeness Centrality (*CC*) measures the average distance - in terms of the total number of edges necessary to reach another node - of a certain node. Closeness Centrality assigns high values for nodes that are low on centrality and small values for nodes that are high on centrality. Instead, the Harmonic Closeness (*HC*) measure is calculated using the inverse of the distances from one node to all others, and gives high values for highly central nodes and small values for nodes that are not so central. Betweenness Centrality (*BC*) of a node is another type of measure that calculates how often a certain node is connected through the smallest distances (paths) among the other nodes. Finally, the Node Strength (*NS*) is independent of the use of thresholds in asset graphs, and takes into account the "strengths" of the connections, the degree of correlation or Transfer Entropy between the nodes. In particular, it measures the sum of the correlations or Transfer Entropy of a node to all others.

On the other hand, the aforementioned measures of centrality are rather inappropriate for directed networks. The latter have either ingoing edges to a node, outgoing edges from the node, or both. So, centrality measures often should be broken down into ingoing and outgoing ones. For example, a node that may be highly central with respect to all others is called *hub,* while a node that has many other nodes pointing at it is called *authority*. Node Degree (*ND*) is analyzed into two measures: In-Node Degree (ND_{in}), which measures the sum of all ingoing edges to a certain node, and Out-Node Degree (ND_{out}) representing the sum of all outgoing edges from a node. In a similar fashion, In-Eigenvector Centrality (EC_{in}), Out-Eigenvector Centrality (EC_{out}), In-Harmonic Closeness (HC_{in}) and Out-Harmonic Closeness (HC_{out}) are defined. Now Betweenness Centrality is calculated along directed paths only, and it is called Directed Betweenness Centrality, (BC_{dir}). Finally, the sum of the weights of all edges that end at a node demonstrates its In-Node Strength (NS_{in}), whilst the sum of the weights of all edges that begin at a node represents its Out-Node Strength (NS_{out}). Thereafter, we use all previously mentioned measures, except Closeness Centrality, to assess the centrality of each node depending on the type of the network.

3. DATA AND PRELIMINARY ANALYSIS

The dataset consists of daily observations of twelve continuous futures prices for three commodity groups: energy (crude oil, heating oil, and natural gas), metals (copper, platinum, gold and silver) and agricultural products (corn, wheat, cocoa, coffee and cotton).[4] Energy and metal futures contracts are traded in the New York Mercantile Exchange (NYMEX), whilst the agricultural commodities are widely traded in the Chicago Board of Trade (CBOT). We consider the U.S. stock indices for 10 industries (S&P sector indices): consumer discretionary, health, energy, financials, industrials, materials, technology, utilities, automobiles and consumer staples. These sectoral indices offer a robust view of the performance of the U.S. stock markets across economic sectors. The S&P 500 index is also used as a proxy for the U.S. equity markets as a whole. All data are obtained from Datastream International. The data span the period January 1995 to June 2015, which covers several episodes of major crises such as the Asian crisis of 1997-1998, the dot-com bubble burst in 2001, the 11th of September 2001 terrorist attack, the Gulf war in 2003, the commodity bull-bear cycle of

[4] Futures prices are derived from individual futures contracts based on the Type 0 and Type 2 roll methods as described in "Futures continuous series: Methodology and definitions.", Thomson Reuters Datastream, August 2010.

2006-2009, the Subprime crisis in 2007 and the global financial crisis of 2008-2009. We compute commodity futures and stock returns by taking the difference in the logarithm of two successive prices. Table 1 presents the descriptive statistics of the daily returns for the U.S. commodity futures and equity markets. Among the commodity futures, heating oil exhibits the highest average return followed by crude oil, while the lowest return is observed for coffee and cotton. Commodity futures with high return volatility as measured by standard deviation include those of the energy (natural gas in particular) and agriculture groups. Gold futures return is the least volatile. For equity, health and consumer discretionary sectors achieve the highest daily average return, while the most volatile sectors are financials and automobiles. Our diagnostic tests (Jarque-Bera, ARCH, and Ljung-Box) show that our return series exhibit departure from normality, conditional heteroscedasticity, and serial correlation in squared level. All return series are stationary according to the Augmented Dickey and Fuller (1981) and Phillips and Perron (1988) tests.[5] Table 2 reports the unconditional correlations between commodity futures and equity markets. They are low in general and positive, with the exception of the links between gold futures and the S&P500, between gold futures and some sector indices (consumer discretionary, health, financials, industrials, technology, automobiles, and consumer staples), and between natural gas futures and automobiles. The highest correlation (0.43) is, as expected found between crude oil futures and energy sector returns. This finding suggests that commodity futures, and particularly gold futures, could provide stock investors with good diversification benefits.

[Please insert Tables 1 and 2 and Figure 2]

An important aspect of empirical investigation is that we use log returns and their 1-period lagged values to explore the relations between present as well as past values of the commodity futures and equity markets. Accordingly, we deal with four interacting clusters over time, as depicted in Figure 2: one of commodity futures, one of equity indices, and two others made by their lagged counterparts. It is worth noting that these clusters are produced by defining a distance matrix between each of the variables (nodes) of the network and then by positioning each node as a point in two dimensions in such a way to preserve as best as possible the distances between nodes (see, Borg and Groenen, 2005). The particular distance measure being used is one minus the Spearman correlation between each pair of variables. We also tested other distance measures, but the results remain intact.

4. Empirical results

4.1 Heat maps for correlation

We first calculate the Pearson correlation matrix of the original variables plus their 1-period lagged counterparts and show the results in Figure 3. Both horizontal and vertical axes represent the varia-

[5] The results are not reported here for brevity, but can be made available on request.

bles in the same ordering described in Section 3, Table 1. Higher correlations are depicted by lighter tones, lower correlations are displayed in darker tones, while the main diagonal acquires the unitary correlation value. As we can see, the correlation matrix displays some structure with two stylized sectors: one representing the correlations among the original indices (left, topmost quarter) and another incorporating the correlations among the lagged variables (right, bottommost quarter). Moreover, there seems to be very low correlation between lagged and original variables (left, topmost and right, bottommost quarters), which indicates that past returns can predict only a small portion of actual returns. Along the main diagonal, there are clusters of bright spots that indicate stronger connection between variables, but it is still moderate and does not eliminate equity-commodity diversification benefits.

[Please insert Figure 3]

Figure 4 provides a breakdown of the correlation matrix into submatrices of correlations in order to better analyze the dynamic linkages between variables. Since the main purpose of this article is to study the commodity market and its relation with the equity indices, we do not explore in detail the correlations between equity indices and themselves for both actual and lagged return series. Figure 5 shows the five correlation submatrices involving commodity futures. In the first heat map where the correlation matrix between commodity futures is represented, we observe the formation of three main clusters with high correlation: one of energy commodities (crude oil and heating oil), one of metal commodities (platinum, silver and gold), and one of agriculture commodities (corn and wheat). The average correlation between commodity futures excluding autocorrelation (0.19) is low when compared with the one between equity indices (0.63). The heat map of correlation between commodity futures and equity indices witnesses the interest of holding commodity assets in diversified portfolios of stocks as most commodity-equity correlations do not exceed 0.2. Copper futures is the least attractive in terms of diversifying potential, while gold futures offer a good cushion for all stock sectors, except for energy sector which has a substantially higher correlation with gold futures. This result corroborates the findings of past studies on the safe haven role of gold for stock markets (e.g., Draper et al., 2006; Baur and McDermott, 2010).

[Please insert Figures 4 and 5]

The heat map of correlations between 1-period lagged and actual commodity futures returns show that all correlations are very low (below 0.07) and the highest correlation occurs between platinum and most lagged commodity futures. The last two heat maps uncover that correlations between lagged equity and actual commodity futures returns are significantly greater than those between lagged commodity futures and actual equity returns. This finding seems to suggest that past equity returns can be used as a predictor of commodity futures returns.

4.2 Statistical significance of correlations

Figure 6 displays the probability distributions of some return correlation submatrices. The histogram of the elements of the return correlation matrix between commodity futures is strongly tilted towards higher values and has some spikes at the right tail, together with a peak that corresponds to autocorrelations. The histogram of the return correlations between equity indices is even more tilted towards higher correlations. Histograms peaked at higher values of correlations are also obtained for the correlation matrices between commodity futures and equity returns, and the one between lagged equity and commodity futures returns. The histograms of the correlation matrices between lagged and actual commodity returns, and between lagged commodity and equity returns are mainly distributed close to zero. Lastly, the histogram of the correlation matrix between lagged and actual equity returns is slightly tilted to negative correlation values.

[Please insert Figure 6 and 7]

Based on the above probability distributions, we develop a new model to determine the characteristics of the networks that are not attributed to random connections. For this purpose, we consider the time series of each variable and reorder or re-shuffle all their elements randomly. This process, while maintaining the basic statistical attributes of the time series, yet destroys any possible correlation between them. A correlation matrix is thereafter calculated based on the new randomized series. Next, we build the probability distribution of the randomized series after 10,000 simulations and we consider the average of the correlation values obtained out of each simulation. The result for each correlation submatrix is a probability distribution that looks very much like a Gaussian distribution, which is depicted in Figure 7. So, correlations that are below -0.04 or above 0.04 can be considered as not being caused by random effects. This applies to all correlation submatrices and shows the statistical significance of almost all commodity-equity correlation coefficients.

4.3 Asset graphs for correlation

We now turn to *Asset Graphs*, which allow us to visualize dynamic networks while reducing the large amount of information contained in the correlation matrix. This tool has been used in a variety of works in finance (e.g., Onela et al., 2002,2003; Ausloos and Lambiotte, 2007; Sandoval, 2012,2014). In asset graphs, threshold values for the connections between nodes (commodity futures and equity indices in our case) are chosen and only nodes with connections (correlations) above (or below) such thresholds are represented, whereas nodes not connected to any other node are not considered at all. Our focus is on the connections between commodity futures only, and also on the connections between commodity futures and equity indices.

[Please insert Figure 8]

Figure 8 shows the asset graphs associated with thresholds ranging from $T=0.7$ to $T=0.2$. Above threshold $T=0.8$, there are no connections among commodity futures, whereas for threshold

T=0.8, the only connections are those between crude oil and heating oil, and between silver and gold. Next, for *T=0.6*, there is a connection between corn and wheat. For *T=0.5*, a cluster formed by platinum, silver and gold emerges, while for *T=0.4*, a connection between crude oil and the energy equity sector index is formed. For *T=0.3*, the metals cluster grows with the inclusion of copper and platinum and a connection between heating oil and the energy equity sector index forms. At this threshold level, a connection between copper and the materials equity index is also established. At *T=0.2*, a large cluster of metals and energy is formed, with connections to the energy, materials, and industrial equity sector indices, and also to the S&P 500. Copper is a major hub, connecting the metals and energy sub-clusters with the equity indices.

The results from asset graphs thus indicate that commodity futures form clusters within particular category and they cluster with other categories of commodity futures and equity markets at low correlation levels. This evidence confirms the findings of Sensoy et al. (2005) in that not all types of commodities are alike, with a high degree of convergence for energy and metal commodities and no sign of convergence for agricultural commodities. Similar to the conclusions of Domanski and Heath (2007), and Dwyer et al. (2011), it shows the potential of diversification benefits for equity investors from holding corn and wheat futures contracts, and energy and precious metal futures contracts (gold, platinum, and silver). Futures contracts on copper, which is an important industrial metal, are less interesting than others given their close connections with industrial activities.

4.4 Centrality measures for asset graphs based on correlation

As we have seen through the use of randomized data, an asset graph just above *T=0.1* can be set as the largest cluster which can be relatively indicated as "free of strong noise effects". We use this topology (*T=0.1*) to calculate some centrality measures and present the results in Tables 3 and 4. One exception is the Node Strength measure, which is calculated directly from the correlation matrix.

Table 3 shows some centrality measures obtained from the commodity-commodity correlation submatrix taken at threshold *T=0.1*. Many of the nodes are very much connected at this level and we thus have a large node degree for most commodity futures, except for natural gas which is decoupled from other energy, commodity and equity markets due to its regional nature and dependence on specific short-lived factors such as weather conditions, inventories and interruptions in supply (Aloui et al., 2014). Regarding the betweeness, harmonic closeness and eigenvalue centralities, we see a strong presence of agricultural commodities, with coffee, cocoa and wheat futures always among the top three positions. Node strength, which is independent of the threshold level, tells a different story. Results are closer to those obtained from node degree and the cluster analysis as expected, but positions change, with the energy and metals commodity futures appearing in the top positions, agricultural commodities appearing next, and natural gas again occupying the last position.

[Please insert Tables 3 and 4]

Table 4 reports the in and out node strengths of the networks concerning some of the remaining correlation submatrices. We do not consider more centrality measures because the networks built from those submatrices have much fewer connections above the threshold level and hence do not provide much relevant information for calculating the centrality measures. We must also have in mind that correlations between -0.04 and 0.04 can be associated with noise, and node strengths from about -0.5 to 0.5 are not statistically significant, since node strengths are calculated as the sums of lines or columns of the correlation submatrices (each with 11 or 12 lines or columns).

For the commodity-equity submatrix, we see that for in-ns, the energy and materials equity indices present higher values and that copper and crude oil have the largest values for out-ns. Gold appears with a negative value, but within the region that may be considered as due to noise. The equity-commodity correlation submatrix is just the transpose of the commodity-equity one, and we just have the exchange between in and out node strengths. There is no implied causation here. The lagged commodity–commodity submatrix leads to very small values of node strengths, well within what may be considered as not statistically significant. Also, although we are dealing with lagged variables, there is no causation implied since the transpose of this matrix (the commodity–lagged commodity submatrix) also gives the same results, with in and out node strengths interchanged. The same observations also apply to the lagged commodity–equities submatrix, as results are also not significant. However, the lagged equity–commodity correlation submatrix offers stronger results, with the metal commodity futures occupying the top in-ns positions while the materials, utilities, and energy equity sector indices have the highest out-ns values. This finding is indeed expected as these economic sectors are the most sensitive to price changes in commodity markets.

4.5 Transfer Entropy and Effective Transfer Entropy

We proceed by calculating the Transfer Entropy (TE) between the original variables and their lagged counterparts and by creating a causality matrix (in the sense of Granger causality). Figure 9 (first heat map) displays a heat map of the TE matrix, depicting with brighter shades the relations of higher value and with darker shades those of relatively low value. The TE was calculated using $k = \ell = 1$ in (3) and one might inquire whether the use of higher values of k and ℓ might offer better results. In order to test this, we calculated the TE using $k = \ell = 2$ in (3). The result is plotted as the second heat map in Figure 9. One may see there is no significant difference between the first and the second heat maps. The difference between each TE matrix is less than 6% of the maximum value of the first TE matrix. Calculations with $k = \ell = 2$ come at the expense of considerable added computational time, therefore we will use $k = \ell = 1$ throughout this article.

[Please insert Figure 9]

As it can be seen from the graph, there is some structure emerging particularly in the bottom left quadrant. However, we must first remove possible effects due to noise and take into account the fact that a high-entropy variable (i.e., high-volatility) naturally transmits more information to another one. Both effects can be taken into account when the TE matrix is re-calculated for randomized data, namely the time series are re-ordered in order to sustain their basic statistical properties, but all causal relations are destroyed. We perform the re-shuffling for a total of 1,000 simulations and take the average among all the calculated values of TE. The resulting TE for randomized data is plotted in the second heat map of Figure 9, where the darker shades in the main diagonal are still pertaining, as well as four lighter streaks that are also present in the matrix as in for the original TE matrix (note that the scales for the TE and randomized TE matrices are not the same, and are adjusted for better visibility). The first horizontal and vertical strips are due to the natural gas variable, and the second ones are produced by the lagged natural gas variable. Both series exhibit a much larger volatility (measured as the standard deviation of the time series) than the other variables.

Furthermore, we remove the randomized TE matrix from the original TE matrix, and thus obtain the Effective Transfer Entropy (ETE) matrix (Marschinski and Kantz, 2002), which is shown as the third heat map in Figure 9. The brighter strips have now disappeared, while the remaining brighter relations may be seen as representing some inherent structure among the variables. The main diagonal is darker as there is no ETE from one source to itself. The top left quadrant depicts the ETE pointing from the original variables to themselves, the bottom left quadrant shows the ETE from the lagged variables to the original ones, the top right quadrant illustrates ETE values from original variables to the lagged ones, whilst the bottom right quadrant demonstrates the ETE scores from the lagged variables to themselves. All but the bottom left quadrant show no sign of non-trivial ETE, i.e., only noise seems to be present in the structure. This means that most of the ETE-based causality dependencies are derived from lagged to original variables, in that the main influences are produced from one day to the next one.

[Please insert Figure 10]

We also break the ETE matrix into submatrices, and discuss in greater detail six of them (Figure 10): the ones depicting the ETEs from commodities to commodities, from commodities to equity indices, from equity indices to commodities, from lagged commodities to commodities, from lagged commodities to equity indices, and from lagged equity indices to commodities.

The ETEs from commodities to commodities (first heat map of Figure 10) are low. There is some ETE from heating oil, natural gas, copper, platinum and silver to crude oil, but the amount of transferred information is not much important. Coffee is the commodity futures that receives the least ETE from other commodity futures. The ETE matrix from commodities to equity indices (second heat map of Figure 10) presents some larger values and a clear transfer of entropy from copper and silver, and in a lesser way, from crude oil and platinum, to all equity indices. The ETE matrix from equity indices to commodities (third heat map in Figure 10) presents similar values as those of the previous submatrix. The diversified and automobiles equity sector indices are the ones that send most ETE to commodities and copper is the commodity futures that receives most ETE from equity indices.

The ETEs from lagged variables to original (unlagged) ones are typically much stronger than the ETEs between simultaneous variables. The fourth heat map of Figure 10 shows the ETEs from lagged commodity futures returns to actual commodity futures returns. We see a similar structure as the one based on the correlation between actual commodity futures turns, but there are very high ETE values from lagged commodity returns to their unlagged counterparts and exist some clusters (crude oil - heating oil, silver - gold - platinum, and corn - wheat). The fifth heat map of Figure 10 shows the ETE from lagged commodity futures returns to actual equity index returns (unlagged). The values are low again, but there are some larger values of ETE from crude oil and copper to most equity indices, with energy and materials being the two equity sector indices that receive most ETE. Energy, in particular, receives a good amount of information from crude and heating oil futures. The sixth heat map of Figure 10, which reveals the ETE from lagged equity index returns to actual commodity futures returns, shows that the energy equity sector transmits a good amount of information to crude oil and heating oil as well as to other commodity futures. The materials equity sector also sends some ETE to other commodity futures.

4.6 Assessing the statistical significance of Effective Transfer Entropy

Figure 11 shows the histograms for eight submatrices of the ETE matrix. The first four graphs show the histograms for ETE submatrix from commodities to commodities, from equity indices to equity indices, from commodities to equity indices, and from equity indices to commodities, respectively. They are all tilted towards higher values centered around different values for each one. The four last graphs show the histograms for ETE submatrices from lagged variables to unlagged ones. All signals are much stronger now and the distributions tend to have gaps between lower and larger values. The histogram from lagged commodities to commodities and from lagged equity indices to lagged equity indices have particularly large values, mostly due to the higher ETE from a lagged variable.

[Please insert Figure 11]

We do not plot here the randomized submatrices, since ETE has them already discounted, so in general, ETE values below zero mean Transfer Entropy values that are smaller than the expected for random data, and that any ETE above zero mean Transfer Entropy values above the expected for random data. Since randomized TE ranges from 0 to 0.0235, we may think as values smaller than 0.0235 as probably due to random effects.

4.7 Asset graphs for Effective Transfer Entropy

Under the same rationale as before, an analysis is conducted via the utilization of asset graphs. We shall do this analysis only on the lagged commodities – commodities submatrix of the ETE matrix, since the largest signals can be found there (with the exception of the ETEs from equity indices to equity indices, which are much higher). All ETE values above *0.3* account for the relationships between lagged variables and their original counterparts and those will not be analyzed thereafter. In Figure 11 we show the graphs entailing thresholds below $T=0.3$. All lines have no arrows because represent bidirectional ETE relations, which are always present in this sample. In Figure 11, the lagged commodities are not represented as separate nodes, but are merged with their unlagged counterparts. Comparing Figure 12 vs. Figure 8, one may observe a prominent correspondence between ETE-based connections vis-à-vis correlation-formed connections. For $T=0.3$ (not represented in the figure), the only ETEs (with the exception of ETEs form lagged variables to their unlagged counterparts) happened between Crude Oil and Heating Oil. For $T=0.2$ only, a new ETE relation appears between Silver and Gold and, for $T=0.15$, between Corn and Wheat. For $T=0.1$, Platinum Connects with Silver and Gold, and for $T=0.05$, we Copper connects to Silver, and we have ETEs being exchanged between Crude Oil, Heating Oil and the Energy equity index.

[Please insert Figure 12]

4.7 Centrality measures for asset graphs based on Effective Transfer Entropy

Next, we use the ETEs asset graphs above threshold $T=0.02$ to study some network centrality measures. The only submatrices with ETEs above this threshold value are those from lagged variables to unlagged ones. The submatrices from commodities to lagged commodities, and from equity indices to equity indices also present some values (from natural gas to lagged natural gas, from diversified to lagged diversified and lagged automobiles, and from automobiles to lagged diversified to lagged automobiles), although very small ones, close to the threshold value.

Table 5 presents some centrality measures for the submatrices corresponding to ETEs from lagged commodities to commodities, from lagged commodities to equity indices, and from lagged equity indices to commodities. The centralities for the submatrix corresponding to ETEs from lagged equity indices to equity indices are not shown, since that network is fully connected at this threshold level and thus offers no useful information.

[Please insert Table 5]

When considering the network formed by the asset tree based on the ETE submatrix from lagged commodities to commodities, we see that crude oil, copper, silver, and gold occupy the positions with highest in and out node degrees. In terms of node strength of the full ETE submatrix, crude oil, heating oil, silver, and gold occupy the highest positions of both in and out node strengths while copper falls to the eighth position in terms of both centralities. For the network formed from the ETE submatrix from lagged commodities to equity indices, we see that few nodes are connected, but energy and materials have highest in node degrees and lagged crude oil and lagged copper have the highest out node degrees. For node strength of the full submatrix, we have again energy and materials as receiving the most ETEs and crude oil, heating oil, and copper, all lagged, as sending the most ETEs. The opposite positions occur when dealing with the network based on the ETE submatrix from lagged equity indices to commodities, showing that the nodes that send most ETEs also receive most ETEs: crude oil, copper and heating oil are the major receivers, and lagged energy and lagged materials are the major senders. For node strength, crude oil, copper and heating oil are again the major receivers, and lagged energy and lagged materials sectors are the major senders.

5. Network Dynamics

In this section, we analyze the temporal (dynamic) dimension of the correlation and Effective Transfer Entropy linkages among the investigated commodities and equity indices. We utilize moving windows, each one covering a certain time period. Considering the trade-off between the "concentrated" influence of particular events that occur rarely during the examined period and require the use of narrow window lengths vs. the statistical problems due to small samples sizes, we adopted each window to correspond to a semester, with 126 observations for every window span. This could be seen as an ideal compromise toward not diluting too many extreme/peak events (e.g., crises) and at the same time not applying relatively small sample effects. Thereafter, we analyze the results from the moving correlation and ETE matrices.

In Figure 13, we plot the individual node strengths (NSs) for each of the variables in 10 submatrices of the correlation matrices calculated at each semester, from 1995 to mid-2015 in two different ways: as individual graphs and as heat maps, one way complementing the other. The NSs of commodities and commodities usually rise and fall together, particularly in the beginning of 2006 and after 2008, the year when the global financial crisis begun. In the case of equity indices and equity indices, we NSs are much higher, and they also rise in 2008, but not as steeply. There is also a peak in the late 2011, coinciding with the European sovereign debt crisis.

[Please insert Figure 13 and 14]

For the correlation submatrix of commodities and equity indices, the In and Out NSs rise during the crisis of 2008 and then slowly fall afterwards. This finding seems to suggest that the strong correlation between commodities and equities was mainly caused by the onset of the recent crises, but not the phenomenon of financialization of commodity markets since 2004 as suggested by several previous studies (Tang and Xiong, 2012; Cheng and Xiong, 2014). Inversely, it is consistent with the results of Büyüksahin and Robe (2011) in that the equity-commodity linkages did not increase until 2008, but only rise sharply during the global financial crisis 2008-2009. The remaining NSs are weaker in comparison to the first one, and we can only see that there is higher volatility after the crisis of 2008, although there is a clear peak in In NS from lagged commodities to equity indices in the first semester of 2010, and clear peak in Out NS from lagged equity indices to commodities in the first semester of 2008.

In Figure 14, the same is done to the ETE matrices, only now there are only In and Out NSs, since all submatrices are asymmetric. Not much can be seen from the evolution of the Out NS of the ETE from commodities to commodities, but there is a sharp drop in In NS in 2008 for crude oil and copper. For the In and Out NSs of the ETE from equity indices to equity indices, we can see a sharp drop of all indices during the crisis of 2008. This is common for the ETE from actual variables in times of high volatility, and shows that there is little spillover of information between return series during the crisis, probably because they behave like a block. The opposite situation happens when we consider the NSs of the ETEs from commodities to equity indices: although there is no rise of Out NS from commodities (except for natural gas and platinum), there is a clear peak in In NS to equity indices that happen precisely in the crisis of 2008. There is also a steep rise of Out NS from equity indices to commodities in the 2008 crisis, but little In NS to commodities at the same time. The ETEs from lagged variables to original (unlagged) ones are generally much higher. Both In and Out NSs rise during the crisis of 2008, falling quickly afterwards, what contrasts the fast rise and slow fall of the corresponding NSs for correlation. Platinum and silver show late peaks, only at the end of 2009. The In and Out NSs from lagged commodities to equity indices show some very diverse results. For Out NS from lagged commodities, there is a peak in early and late 2009 and another large peak in the end of 2011 for crude and heating oils, and also for copper. For In NS to equity indices, there are smaller peaks in 2009 and 2011. For the Out NS from lagged equity indices to commodities, we see a rise of energy sector before and during the crisis of 2008 and a peak in materials sector during the crisis. Another, smaller peak can be also be seen in 2011. The In NSs to commodities also rise during the crisis of 2008 and particularly during the European debt crisis in 2011, but mainly for the crude oil, heating oil, and copper futures. Looking at the In and Out NSs from lagged equity indices to equity indices, we see a very uniform behavior between NSs. Peaks occur in 2003, 2008, and in 2011, pinpointing some of the most severe episodes of recent crises. Lastly, the results from the heat maps of the ETE node strengths are in full accordance with the aforementioned findings.

6. Conclusions

We investigated the causal relationships between U.S. equity and commodity futures markets via the utilization of complex network theory, and particularly tools based on correlation and Transfer Entropy measures. This approach allows us to explore the commodity-equity linkages not only within same category of assets (i.e., energy futures, metal futures, agricultural futures, and equities) but also the clusters across different categories of assets (e.g., energy-metal futures, metal-agricultural futures, and equity-metal futures, etc.) extracted from various network topologies. Rolling estimations of large matrices are also implemented to analyze the temporal (dynamic) dimensions of equity-commodity networks based on correlation and Transfer Entropy. Furthermore, we conducted a simulation analysis using randomized time series for rolling windows of centrality measures to assess the impact of various time periods on the data dependence structure. We mainly show that commodity futures markets are not homogenous and only have strong connections within the same category. They are still decoupled from equity markets. The asset graphs based on the effective entropy transfer shows for example that commodity futures form three networks (crude oil-heating oil, silver-gold-copper-platinum, and corn-wheat), and that only energy equity sector is connected to energy futures (crude and heating oil). In terms of effective entropy transfer, financials, automobiles and energy equity sectors transmit the most information to commodity futures (particularly copper), while copper, silver, and crude oil futures to a lesser extent transfer the most information to equity markets. Finally, our results indicate that equity-commodity links only increased during recent crises, possibly due to increased holdings of commodity futures by financial investors to diversify way the risk of stock portfolios. Accordingly, while it does not seem to play an important role, the financialization of commodity markets helps facilitate the risk management and hedging strategies with commodity futures trading.

REFERENCES

Adams, Z., Glück, T., 2015. Financialization in commodity markets: A passing trend or the new normal? Journal of Banking & Finance 60, 93-111.
Allen, F., Babus, A., 2009. Networks in finance. In Kleindorfer, P., Wing, Y.,and Gunther, R. (eds.), The network challenge: strategy, profit, and risk in an interlinked world. Wharton School Publishing.
Allen, F., Gale, D., 2000. Financial contagion. Journal of Political Economy 108, 1-33.
Aloui, R., Ben Aïssa, M.S., Hammoudeh, S., Nguyen, D.K., 2014. Dependence and extreme dependence of crude oil and natural gas prices with applications to risk management. Energy Economics 42, 332-342. Altiparmak, F., Dengiz, B., 2009. A cross entropy approach to design of reliable networks. European Journal of Operational Research 199 (2), 542-552
Amblard, P.-O., Michel, O.J.J., 2013. The relation between Granger causality and directed information theory: A review. Entropy 15, 113–143.
Anton, H., Rorres, C., 2005. Elementary Linear Algebra with Applications, 9th edition. Wiley.

Arouri, M., Jouini, J., Nguyen, D.K., 2011. Volatility spillovers between oil prices and stock sector re-turns: implications for portfolio management. Journal of International Money and Finance 30, 1387-1405.
Ausloos, M. and Lambiotte, R., 2007. Clusters or networks of economies? A macroeconomy study through gross domestic product. Physica A 382, 16-21.
Barnett, L., 2009. Granger causality and transfer entropy are equivalent for Gaussian variables. Physical Review Letters 103, 238701.
Barro, D., Basso, A., 2010. Credit contagion in a network of firms with spatial interaction. European Journal of Operational Research 205 (2), 459-468
Baur, D.G., McDermott, T. K. 2010. Is gold a safe haven? International evidence. Journal of Banking and Finance 34, 1886-1898.Borg, I., and Groenen, P. 2005. Modern Multidimensional Scaling: theory and applications, 2nd edition, Springer-Verlag.
Büyükahin, B., Haigh, M.S., Robe, M.A., 2010. Commodities and equities: Ever a market of one? Journal of Alternative Investments, 12(3), 76-95.
Büyükşahin, B., Robe, M., 2011. Does paper oil matter? Energy markets' financialization and equity-commodity co-movements.Working paper, CFTC.
Büyükşahin, B., Robe, M.A, 2014. Speculators, commodities and cross-market linkages. Journal of International Money and Finance 42, 38-70.
CFTC, 2008. Staff report on commodity swap dealers & index traders with commission recommendations. September 2008. http://www.cftc.gov/PressRoom/PressReleases/pr5542-08
Chan-Lau, J. A., Mathieson, D. J., Yao, J.Y., 2004. Extreme contagion in equity markets. IMF Staff Papers 51, 386-408.
Cheng, I. H., Xiong, W. 2014. Financialization of commodity markets. Annual Review of Financial Economics 6, 419-441.
Daskalaki, C., Skiadopoulos, G., 2011. Should investors include commodities in their portfolio after all? New evidence. Journal of Banking and Finance 35, 2606-2626.
Dickey, D. A., Fuller, W. A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49, 1057-1072.
Dimpfl, T., Peter, F.J., 2012. Using transfer entropy to measure information flows between financial markets. Studies in Nonlinear Dynamics and Econometrics 17, 85–102.
Domanski, D., Heath, A., 2007. Financial investors and commodity markets. BIS Quarterly Review, 53-67.
Draper, P., Faff, R.W., and Hillier, D., 2006. Do precious metals shine? An investment perspective. Financial Analysts Journal 62, 98-106.
Dwyer, A., Gardner, G., Williams, T., 2011. Global commodity markets–price volatility and financialisation. Reserve Bank of Australia Bulletin, June, 49–57.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica 50, 987–1008.
Faes, L., Nollo, G., Porta, A., 2013. Compensated transfer entropy as a tool for reliably estimating information transfer in physiological time series. Entropy 15, 198–219.
Filis, G., Degiannakis, S., Floros, C., 2011. Dynamic correlation between stock market and oil prices: The case of oil-importing and oil-exporting countries. International Review of Financial Analysis 20(3), 152-164.
Forbes, K., Rigobon, R., 2002. No contagion, only interdependence: Measuring stock market comovements. Journal of Finance 57, 2223-2261.
Gómez, D., Figueira, J.-R., Eusébio, A., 2013. Modeling centrality measures in social network analysis using bi-criteria network flow optimization problems. European Journal of Operational Research 226 (2), 354-365
Gorton, G., Rouwenhorst, K.G., 2006. Facts and fantasies about commodity futures. Financial Analysts Journal 62, 47- 68.
Gromb, D., Vayanos, D., 2010. Limits of arbitrage: The state of the theory. National Bureau of Economic Research, Working paper series no.15821.

Haldane, A., 2009. Rethinking the financial network. Speech delivered at the Financial Student Association, Amsterdam.
Irwin, S.H., Sanders, D.R., 2011. Index funds, financialization, and commodity futures markets. Applied Economic Perspectives and Policy 33(1), 1–31.
Johnson, L. L., 1960. The theory of hedging and speculation in commodity futures. Review of Economics Studies, 27, 139–151.
Kwon, O., Yang, J-S., 2008. Information flow between composite stock index and individual stocks. Physica A 387, 2851-2856.
Liu, L.F., Hu, H.P., Deng, Y.S., Ding, N.D., 2014. An entropy measure of non-stationary processes. Entropy 16, 1493–1500.
Lizier, J.T., Mahoney, J.R., 2013. Moving frames of reference, relativity and invariance in transfer entropy and information dynamics. Entropy 15, 177–197.
Marschinski, R., Kantz, H., 2002. Analysing the information flow between financial time series – an improved estimator for transfer entropy. European Physics Journal B 2002, 30, 275–281.
Nagurney, A., Ke, K., 2006. Financial networks with intermediation: Risk management with variable weights. European Journal of Operational Research 172 (1), 40-63
Narayan, P. K., Narayan, S., Sharma, S. S., 2013. An analysis of commodity markets: What gain for investors? Journal of Banking and Finance 37(10), 3878-3889.
Narayan, P.K., Sharma, S., 2011. New evidence on oil price and firm returns. Journal of Banking and Finance 35, 3253–3262.
Newman, M.E.J. Networks: an introduction, 2010. Oxford University Press.
Nichols, J.M., Bucholtz, F., Michalowicz, J.V., 2013. Linearized transfer entropy for continuous second order systems. Entropy 15, 3186–3204.
Onnela, J.-P., Chakraborti, A., Kaski, K., 2003. Dynamics of market correlations: taxonomy and portfolio analysis. Physical Review E 68, 1-12.
Onnela, J.-P., Chakraborti, A., Kaski, K., Kertész, J., 2002. Dynamic asset trees and portfolio analysis. European Physics Journal B 30, 285-288.
Phillips, P. C., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335-346.
Prokopenko, M., Lizier, J.T., Price, D.C., 2013. On thermodynamic interpretation of transfer entropy. Entropy 15, 524–543.
Sandoval Jr., L., 2013. Cluster formation and evolution in networks of financial market indices. Algorithmic Finance 2, 3-43.
Sandoval Jr., L., 2014. Structure of a global network of financial companies based on transfer entropy. Entropy 16 (8), 4443-4482
Sandoval Jr., L., Mullokandov, A., Kenett, D.Y., 2015. Dependency relation among international stock market indices. Journal of Risk and Financial Management 8 (2), 227-265.
Schreiber, T., 2000. Measuring information transfer. Physical Review Letters 85, 461-464.
Sensoy, A., Hacihasanoglu, E., Nguyen, D.K., 2015. Dynamic convergence of commodity futures: Not all types of commodities are alike. Resources Policy 44, 150-160.
Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423, 623-656.
Silvennoinen, A., Thorp, S., 2013. Financialization, crisis and commodity correlation dynamics. Journal of International Financial Markets, Institutions and Money, 24, 42-65.
Tang, K., Xiong, W., 2012. Index investment and the financialization of commodities. Financial Analysts Journal 68(6), 54-74.
Upper, C., 2011. Simulation methods to assess the danger of contagion in interbank markets. Journal of Financial Stability 7(3), 111–125.
Ver Steeg, G., Galstyan, A., 2012. Information transfer in social media. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012; ACM: New York, NY, USA, 2012; pp. 509–518.

Veremyev, A., Boginski, V., 2012. Identifying large robust network clusters via new compact formulations of maximum k-club problems. European Journal of Operational Research 218 (2), 316-326
Vicente, R., Wibral, M., Lindner, M., Pipa, G., 2011. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. Journal of Computational Neuroscience 30, 45–67.

TABLE 1: DESCRIPTIVE STATISTICS

	Mean (%)	*Std. Dev.(%)*	*Skewness*	*Kurtosis*	*JB*	*Q^2(12)*	*ARCH(12)*
				Commodity futures			
Crude oil	0.010	1.000	-0.132	7.615	4760^{+}	1456.29^{+}	55.02^{+}
Heating oil	0.011	0.959	-0.637	9.380	9431^{+}	336.18^{+}	16.98^{+}
Natural gas	0.004	1.542	-0.031	12.202	18866^{+}	321.43^{+}	18.72^{+}
Copper	0.005	0.748	-0.222	7.328	4217^{+}	2337.27^{+}	73.95^{+}
Platinum	0.008	0.621	0.429	20.118	65445^{+}	122.47^{+}	7.57^{+}
Silver	0.009	0.825	-0.821	10.860	14363^{+}	641.72^{+}	29.84^{+}
Gold	0.009	0.465	-0.116	10.421	12280^{+}	551.84^{+}	25.50^{+}
Corn	0.005	0.776	-0.612	15.762	36618^{+}	47.94^{+}	3.24^{+}
Wheat	0.003	0.835	0.227	5.595	1546^{+}	492.33^{+}	22.05^{+}
Cocoa	0.008	0.807	-0.121	6.301	2441^{+}	227.50^{+}	12.58^{+}
Coffee	-0.002	1.041	0.116	7.455	4434^{+}	808.57^{+}	37.06^{+}
Cotton	-0.002	0.869	-1.087	22.147	82729^{+}	18.51^{+++}	1.32
				Equity markets			
S&P500	0.012	0.517	-0.247	11.577	16445^{+}	4839.60^{+}	153.41^{+}
Consumer discretationary	0.015	0.584	-0.110	10.280	11818^{+}	3703.41^{+}	119.26^{+}
Health	0.018	0.510	-0.135	9.584	9674^{+}	2407.78^{+}	88.78^{+}
Energy	0.014	0.684	-0.305	13.888	26493^{+}	5473.63^{+}	174.13^{+}
Financials	0.013	0.910	-0.040	17.797	48783^{+}	3840.63^{+}	121.70^{+}
Industrials	0.013	0.567	-0.342	8.882	7813^{+}	3853.55^{+}	122.97^{+}
Materials	0.009	0.645	-0.248	9.879	10598^{+}	4943.48^{+}	152.60^{+}
Technology	0.016	0.776	0.152	8.003	5597^{+}	2437.91^{+}	79.05^{+}
Utilities	0.006	0.499	-0.027	13.422	24199^{+}	4273.43^{+}	133.54^{+}
Automobiles	0.001	0.933	-0.088	12.114	18512^{+}	3893.01^{+}	131.98^{+}
Consumer staples	0.013	0.413	-0.127	11.756	17095^{+}	2491.76^{+}	90.39^{+}

Notes: JB, Q^2(12) and ARCH(12) refer to Jarque-Bera test for normality, the Ljung-Box test with 12-order serial autocorrelation in squared returns, and Engle (1982)'s test for conditional heteroscedasticity, respectively. $^{+}$, $^{++}$ and $^{+++}$ indicates rejection of the null hypotheses of normality, no autocorrelation and conditional homoscedasticity at the 1%, 5% and 10% levels.

TABLE 2: UNCONDITIONAL CORRELATION

	S&P500	*Cons-Dis*	*Health*	*Energy*	*Financials*	*Industrials*	*Materials*	*Technology*	*Utilities*	*Automobiles*	*Cons-Sta*
Crude oil	0.15	0.08	0.05	0.43	0.11	0.12	0.19	0.09	0.13	0.08	0.04
Heating oil	0.13	0.06	0.05	0.37	0.08	0.10	0.16	0.09	0.13	0.06	0.02
Natural Gas	0.03	0.01	0.01	0.16	0.01	0.01	0.04	0.00	0.06	-0.01	0.00
Copper	0.23	0.19	0.13	0.29	0.19	0.23	0.31	0.17	0.15	0.18	0.12
Platinum	0.09	0.06	0.03	0.17	0.06	0.08	0.17	0.07	0.08	0.06	0.04
Silver	0.07	0.03	0.02	0.20	0.03	0.07	0.20	0.04	0.05	0.04	0.02
Gold	-0.02	-0.06	-0.05	0.12	-0.06	-0.02	0.12	-0.03	0.00	-0.05	-0.04
Corn	0.10	0.08	0.05	0.16	0.06	0.09	0.15	0.08	0.09	0.07	0.06
Wheat	0.11	0.09	0.06	0.16	0.08	0.10	0.15	0.08	0.08	0.08	0.07
Cocoa	0.08	0.06	0.06	0.13	0.06	0.08	0.13	0.03	0.06	0.05	0.06
Coffee	0.07	0.07	0.05	0.10	0.06	0.07	0.11	0.05	0.04	0.06	0.05
Cotton	0.11	0.09	0.08	0.14	0.09	0.09	0.13	0.06	0.09	0.08	0.06

Notes: Cons-Dis and Cons-Sta refer to consumer discretionary and consumer staples sectors.

TABLE 3: CENTRALITY MEASURES FOR THE UNDIRECTED COMMODITIES NETWORK BASED ON CORRELATIONS

Degree	*Betweenness*	*Harmonic Closeness*	*Eigenvector*	*Strength*
Crude Oil (10)	Coffee (5)	Coffee (5)	Coffee (6)	Silver (3.98)
Heating Oil (10)	Wheat (4)	Cocoa (4)	Cocoa (5)	Gold (3.72)
Copper (10)	Cotton (4)	Wheat (3)	Wheat (4)	Crude Oil (3.59)
Platinum (10)	Cocoa (4)	Cotton (3)	Cotton (4)	Platinum (3.40)
Silver (10)	Copper (3)	Natural Gas (2)	Copper (3)	Heating Oil (3.39)
Gold (10)	Platinum (3)	Copper (1)	Platinum (3)	Copper (3.37)
Corn (10)	Silver (3)	Platinum (1)	Silver (3)	Corn (3.16)
Wheat (9)	Gold (3)	Silver (1)	Gold (3)	Wheat (3.05)
Cotton (9)	Corn (3)	Gold (1)	Corn (3)	Cotton (2.37)
Cocoa (9)	Natural Gas (2)	Corn (1)	Natural Gas (2)	Cocoa (2.30)
Coffee (8)	Crude Oil (1)	Crude Oil (1)	Crude Oil (1)	Coffee (2.25)
Natural Gas (2)	Heating Oil (1)	Heating Oil (1)	Heating Oil (1)	Natural Gas (1.99)

Notes: Variables classified in order of centrality, according to the diverse centrality criteria. The values of centralities appear in brackets.

TABLE 4: NODE STRENGTHS FOR DIRECTED NETWORKS BASED ON CORRELATIONS

Commodities - Equities		*Commodities* - Commodities*	
In-NS	*Out-NS*	*In-NS*	*Out-NS*
Energy (2.43)	Copper (2.19)	Platinum (0.46)	Wheat* (0.25)
Materials (1.86)	Crude Oil (1.48)	Cocoa (0.35)	Corn* (0.2)
S&P 500 (1.16)	Heating Oil (1.26)	Silver (0.13)	Platinum* (0.12)
Industrials (1.04)	Wheat (1.06)	Coffee (0.13)	Natural Gas* (0.08)
Utilities (0.95)	Cotton (1.01)	Gold (0.11)	Coffee* (0.08)
Diversified (0.78)	Corn (1.01)	Corn (0.06)	Cotton* (0.07)
Discretionary (0.76)	Platinum (0.91)	Corn (-0.01)	Heating Oil* (0.00)
Technology (0.73)	Cocoa (0.81)	Heating Oil (-0.02)	Cocoa* (0.00)
Automobiles (0.70)	Silver (0.75)	Wheat (-0.07)	Crude Oil* (-0.02*)
Health (0.51)	Coffeee (0.73)	Natural Gas (-0.12)	Gold* (-0.03)
Staples (0.50)	Natural Gas (0.32)	Copper (-0.14)	Silver* (-0.03)
	Gold (-0.08)	Crude Oil (-0.16)	Copper* (-0.06)
Commodities - Equities*		*Equities* - Commodities*	
In-NS	*Out-NS*	*In-NS*	*Out-NS*
Materials (-0.09)	Natural Gas* (-0.05)	Silver (1.15)	Materials* (0.81)
Automobiles (-0.11)	Gold* (-0.12)	Platinum (1.11)	Utilities* (0.70)
Utilities (-0.12)	Heating Oil* (-0.12)	Copper (0.82)	Energy* (0.69)
Energy (-0.16)	Copper* (-0.12)	Gold (0.54)	S&P 500* (0.69)
Technology (-0.16)	Cotton* (-0.13)	Coffee (0.52)	Discretionary* (0.63)
Industrials (-0.19)	Wheat* (-0.13)	Crude Oil (0.50)	Industrials* (0.62)
Health (-0.24)	Coffee* (-0.15)	Cocoa (0.43)	Automobiles* (0.60)
Diversified (-0.26)	Corn* (-0.19)	Heating Oil (0.41)	Technology* (0.54)
Discretionary (-0.27)	Cocoa* (-0.21)	Corn (0.34)	Staples* (0.51)
S&P 500 (-0.27)	Platinum* (-0.26)	Cotton (0.34)	Diversified* (0.48)
Staples (-0.30)	Crude Oil* (-0.34)	Natural Gas (0.25)	Health* (0.39)
	Silver* (-0.35)	Wheat (0.24)	

TABLE 5: NODE DEGREES FOR DIRECTED NETWORKS BASED ON CORRELATIONS

From Lagged Commodities to Commodities			
In-ND	*Out-ND*	*In-NS*	*Out-NS*
Crude Oil (5)	Crude Oil* (5)	Crude Oil (0.56)	Crude Oil* (0.57)
Copper (4)	Copper* (5)	Heating Oil (0.53)	Silver* (0.53)
Silver (4)	Silver* (5)	Silver (0.52)	Heating Oil* (0.52)
Gold (4)	Gold* (4)	Gold (0.47)	Gold* (0.46)
Heating Oil (3)	Platinum* (3)	Platinum (0.35)	Platinum* (0.34)
Platinum (3)	Heating Oil* (2)	Corn (0.31)	Corn* (0.31)
Natural Gas (2)	Natural Gas* (2)	Wheat (0.28)	Wheat* (0.28)
Corn (2)	Corn* (1)	Copper (0.24)	Copper* (0.24)
Wheat (1)	Wheat* (1)	Cotton (0.11)	Cotton* (0.11)
		Cocoa (0.10)	Coffee* (0.10)
		Natural Gas (0.10)	Natural Gas* (0.10)
		Coffee (0.09)	Cocoa* (0.09)
From Lagged Commodities to Equities			
In-ND	*Out-ND*	*In-NS*	*Out-NS*
Energy (3)	Crude Oil* (3)	Energy (0.28)	Crude Oil* (0.22)
Materials (2)	Copper* (3)	Materials (0.16)	Copper* (0.17)
Diversified (1)	Heating Oil* (1)	Diversified (0.10)	Heating Oil* (0.12)
Automobiles (1)		Automobiles (0.09)	Silver* (0.09)
		Industrials (0.08)	Corn* (0.08)
		S&P 500 (0.07)	Platinum* (0.08)
		Discretionary (0.07)	Wheat* (0.06)
		Utilities (0.06)	Gold* (0.06)
		Technology (0.06)	Cotton* (0.05)
		Health (0.03)	Coffee* (0.04)
		Staples (0.02)	Cocoa* (0.03)
			Natural Gas* (0.03)
From Lagged Equities to Commodities			
In-ND	*Out-ND*	*In-NS*	*Out-NS*
Crude Oil (3)	Energy* (3)	Crude Oil (0.23)	Energy* (0.29)
Copper (2)	Materials* (2)	Copper (0.16)	Materials* (0.16)
Heating Oil (1)	Diversified* (1)	Heating Oil (0.13)	Diversified* (0.13)
	Automobiles* (1)	Silver (0.09)	Automobiles* (0.10)
		Corn (0.08)	Industrials* (0.08)
		Platinum (0.08)	S&P 500* (0.07)
		Cotton (0.07)	Technology* (0.07)
		Wheat (0.06)	Consumer* (0.06)
		Gold (0.06)	Utilities* (0.06)
		Cocoa (0.05)	Health (0.04)
		Coffee (0.04)	Consumer* (0.03)
		Natural Gas (0.04)	

Notes: Variables classified in order of in and out node strengths for four correlation and three ETE submatrices submatrices. Lagged variables appear with an *.

FIGURE 1: SCHEMATIC REPRESENTATION OF TRANSFER ENTROPY

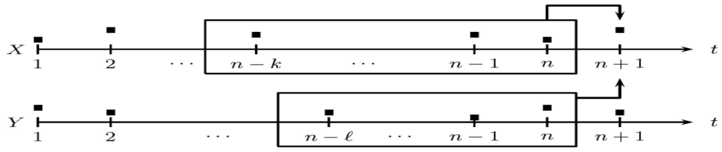
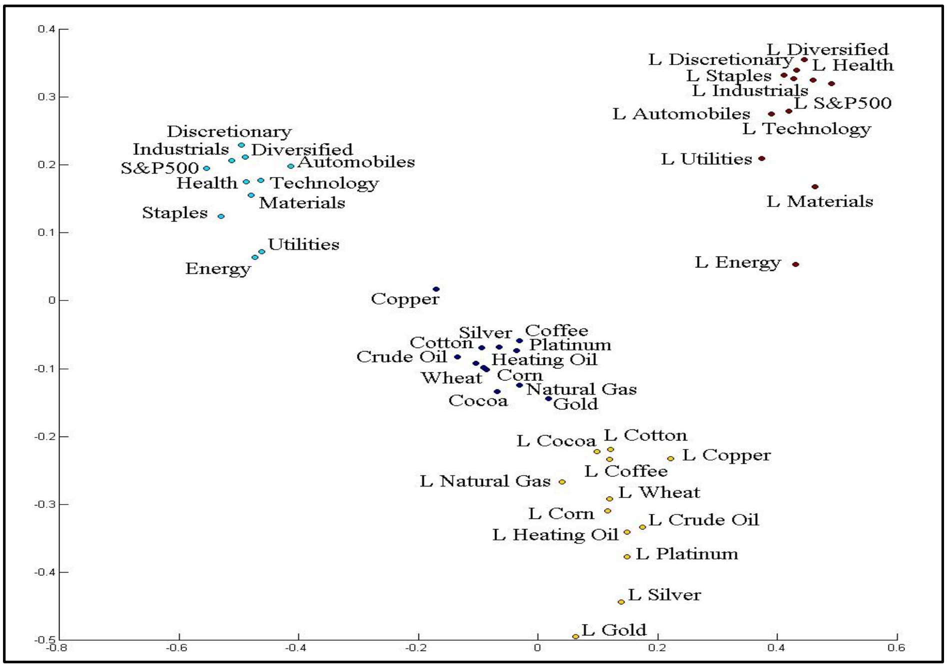
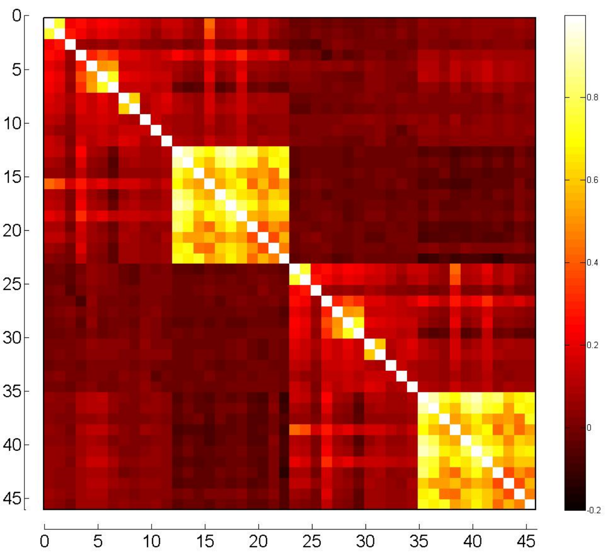


FIGURE 2: DISTANCE MAP OF COMMODITIES, EQUITY INDICES & THEIR LAGGED COUNTERPARTS



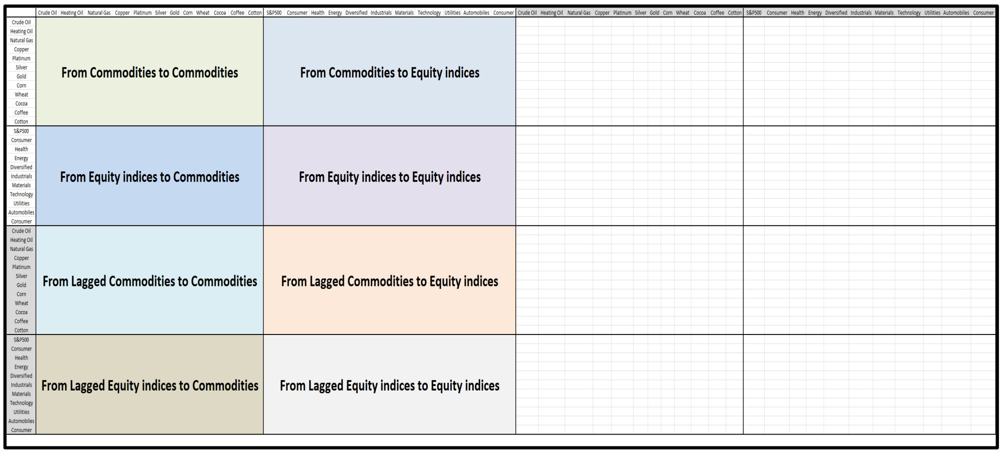
Notes: This figure shows four interacting clusters over time: one of commodity futures returns, one of equity index returns, and two others made by their 1-lagged values. They are produced by defining a distance matrix between each of the variables (nodes) of the network and then by positioning each node as a point in two dimensions. The distance measure is one minus the Spearman correlation between each pair of variables.

FIGURE 3: HEAT MAP OF THE CORRELATION MATRIX BETWEEN VARIABLES



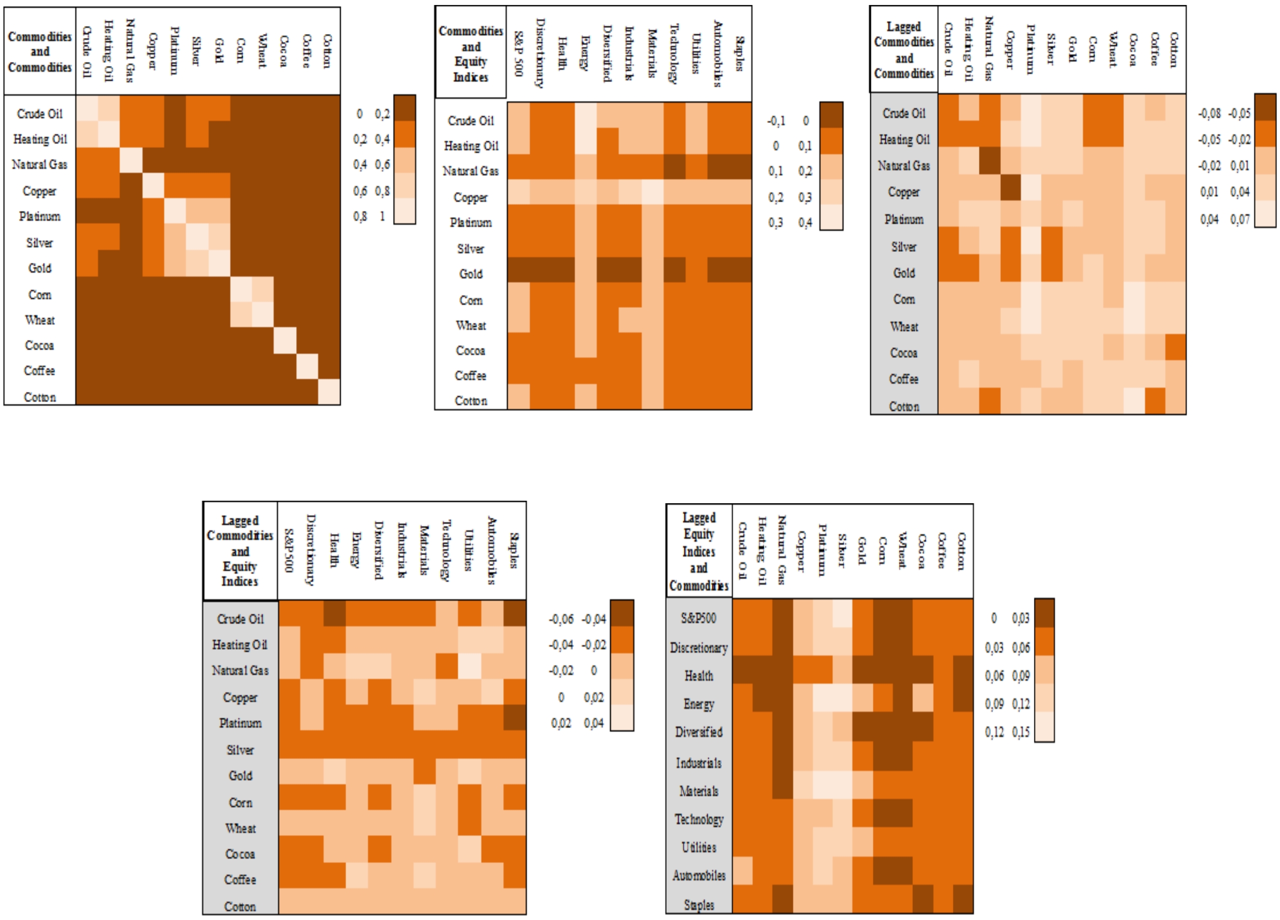
Notes: Brighter tones represent higher correlation, and darker tones represent lower correlation.

FIGURE 4: PARTITION OF THE CORRELATION MATRIX ACCORDING TO THE NATURE OF VARIABLES



Notes: Separation of the enlarged correlation and ETE matrices into eight separate matrices. The effects of unlagged variables to lagged ones (from the future to the past) are being ignored, as well as the effects from lagged to lagged variables, which are almost identical to the effects of variables on themselves, without lagging.

FIGURE 5: HEAT MAPS OF CORRELATION INVOLVING COMMODITIES AND EQUITY INDICES.



Notes: The first heat map shows the correlations between commodities; the second heat map shows the correlations between commodities and equity indices; the third heat map represents the correlations between lagged commodities and commodities; the fourth heat map shows the correlations between lagged commodities and equity indices; and the fifth heat map represents the correlations between lagged equity indices and commodities.

FIGURE 6: PROBABILITY DISTRIBUTION OF CORRELATION MATRICES

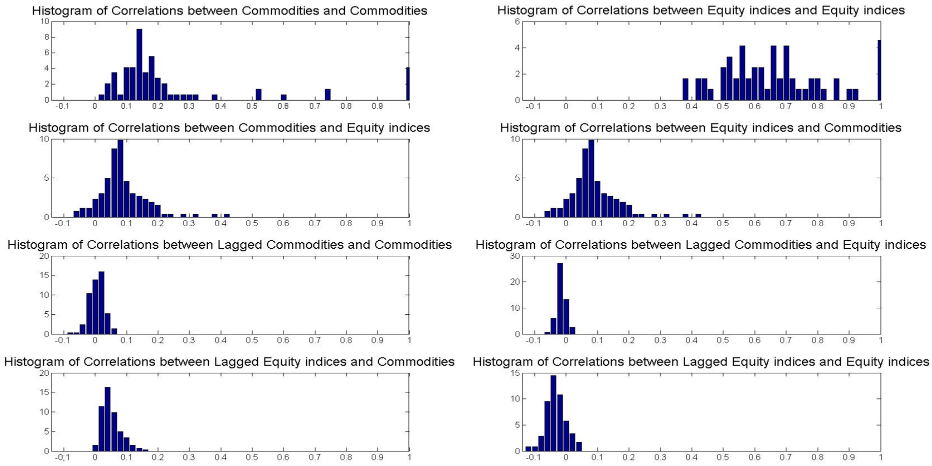
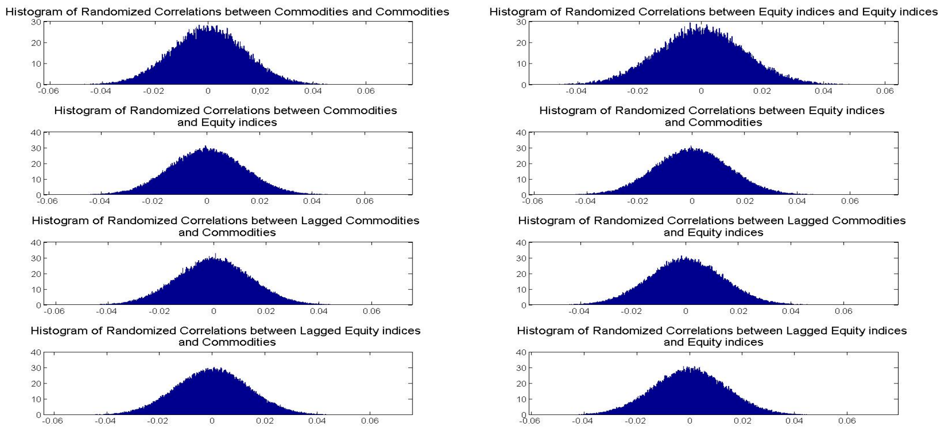


FIGURE 7: PROBABILITY DISTRIBUTION OF CORRELATION MATRICES BASED ON RANDOMIZED DATA



Notes: The probability distribution of the correlation matrices (Figure 6), and the same probability distribution obtained from 10,000 simulations of correlation matrices based on randomized data (Figure 7).

FIGURE 8: ASSET GRAPHS BASED ON THE CORRELATION MATRIX

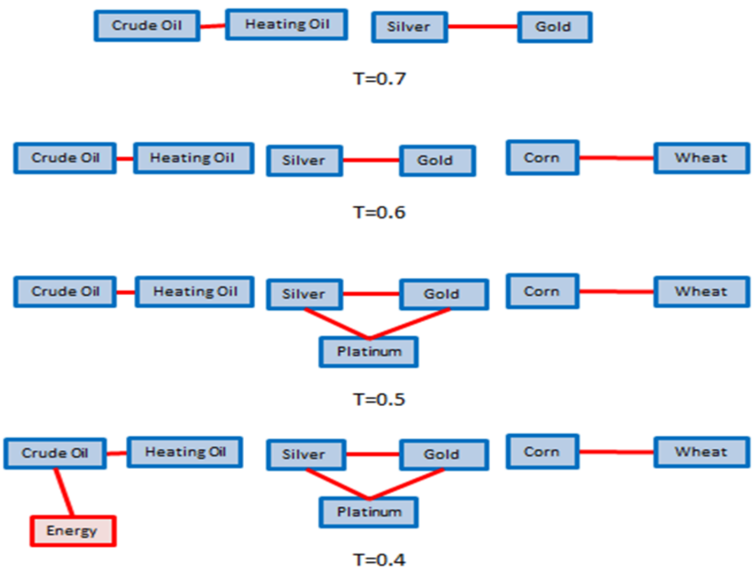
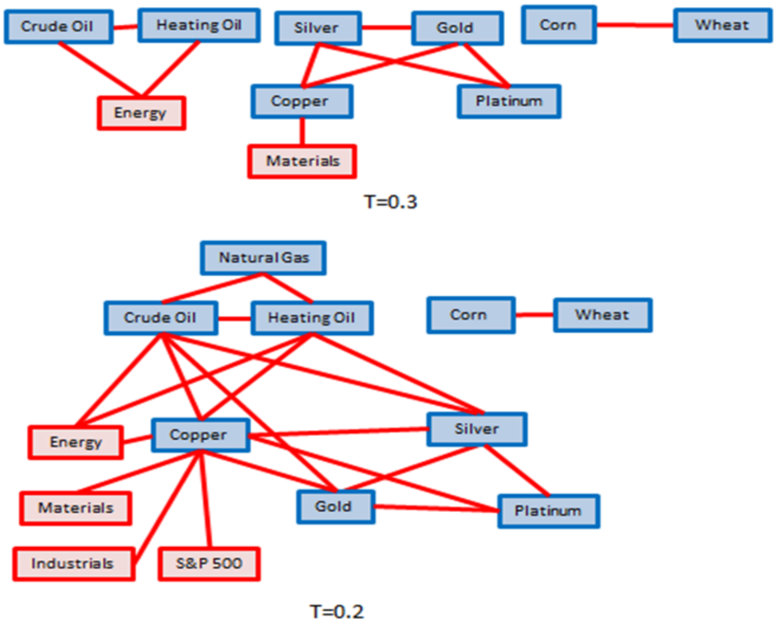
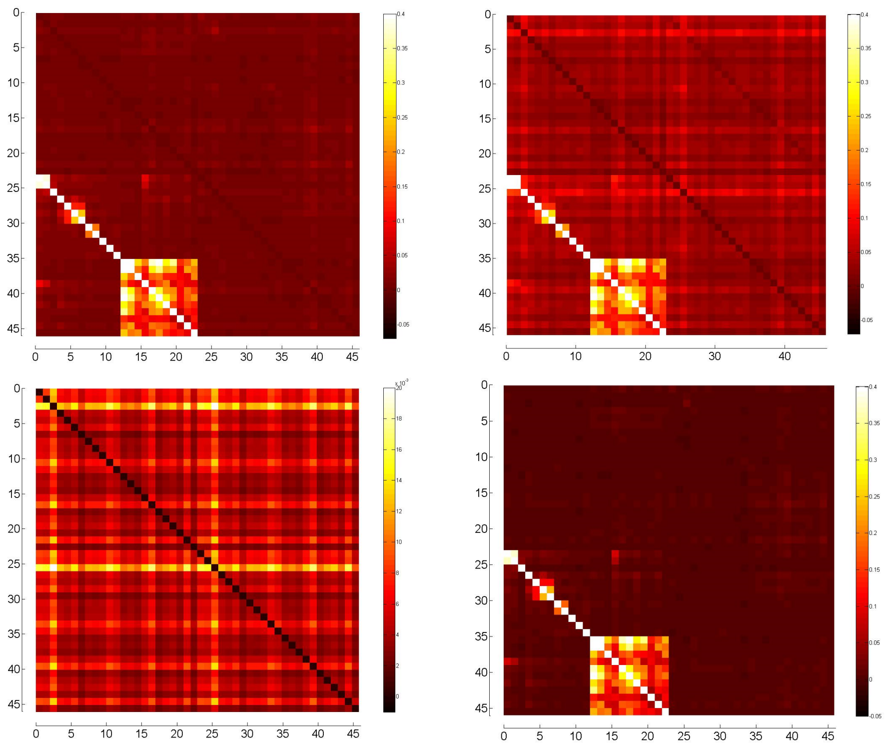


Figure 8 *(contd.)*: Asset graphs based on the correlation matrix *(contd.)*



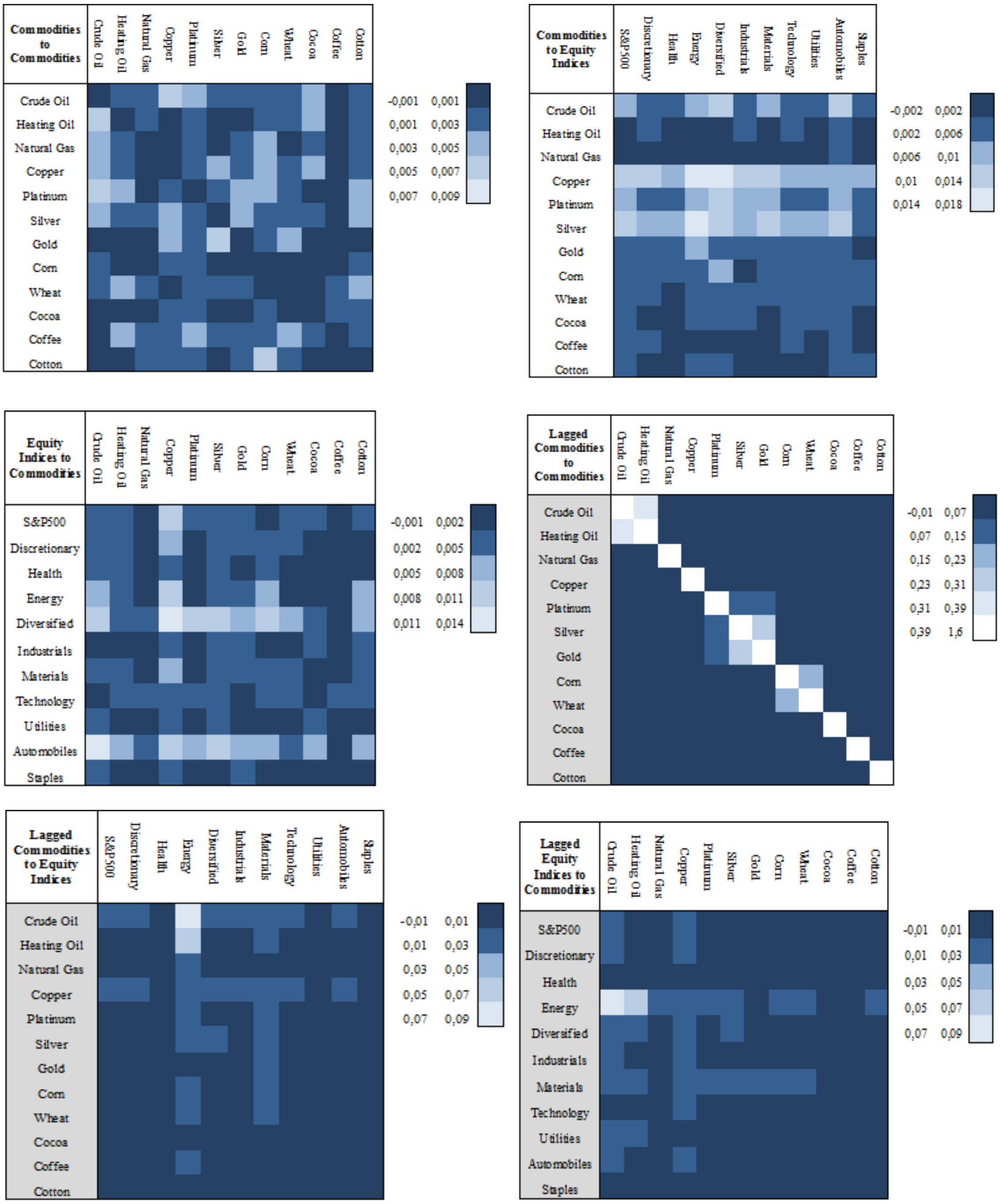
Notes: Asset graphs with thresholds from *T=0.7* to *T=0.2*. The nodes associated with lagged variables are not represented, since they have an almost exact structure as the ones associated with the original variables, and form a cluster detached from the first one. Only for *T=0.1*, there are connections between the original and lagged clusters. The network between commodities is also not represented.

Figure 9: Heat map of TE for $k = \ell = 1$ (1st figure), TE for $k = \ell = 2$ (2nd figure), of randomized TE (3rd figure), and of ETE (4th figure)



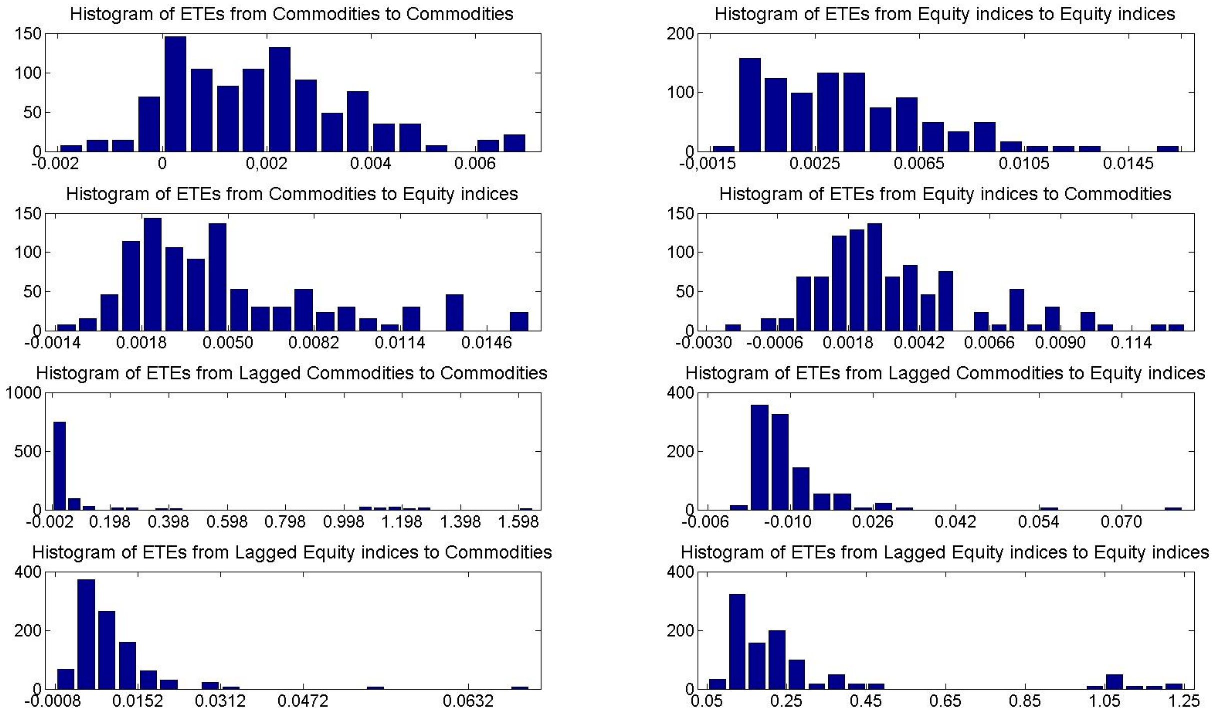
Notes: Brighter tones represent higher transfer entropy (or higher effective transfer entropy), and darker tones represent lower transfer entropy (or lower effective transfer entropy). The first graph represents TE for $k = \ell = 1$ and the second graph represents TE for $k = \ell = 2$. The third graph represents the average of the correlations calculated based on randomized time series. The fourth heat maps represents the Effective Transfer Entropy, obtained by subtracting the Transfer Entropy based on randomized data from the original Transfer Entropy matrix.

FIGURE 10: HEAT MAPS OF THE EFFECTIVE TRANSFER ENTROPY INVOLVING COMMODITIES AND EQUITY INDICES.



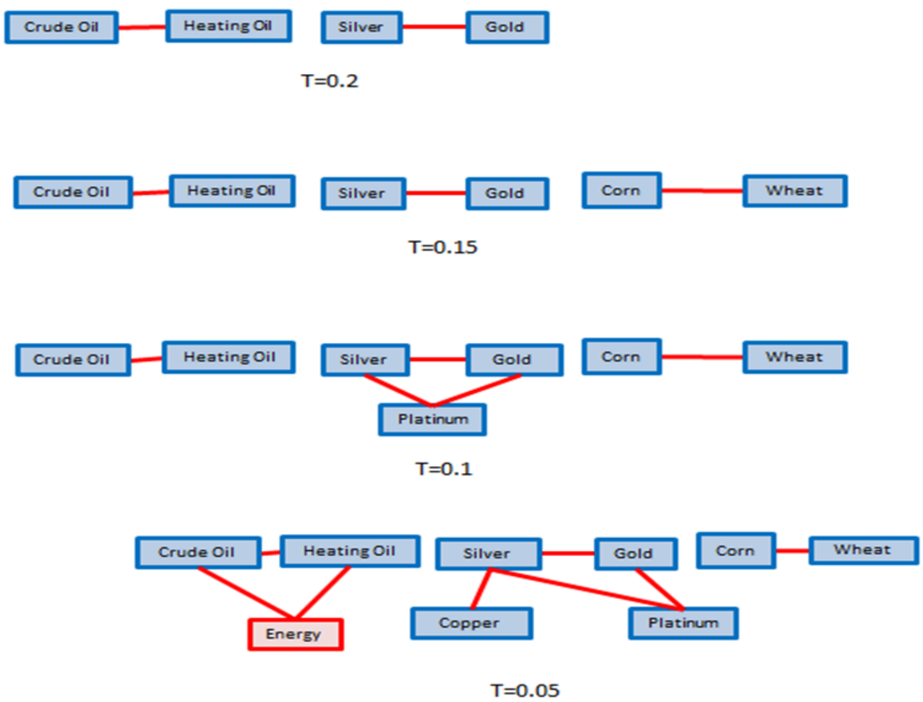
Notes: The first heat map shows the ETEs from commodities to commodities. The second heat map represents the ETEs from commodities to equity indices; the third heat map shows the ETEs from equity indices to commodities; the fourth heat map represents the ETEs from lagged commodities to commodities; the fifth heat map represents the ETEs from lagged commodities to equity indices; and the sixth heat map shows the ETEs from lagged equity indices to commodities.

FIGURE 11: PROBABILITY DISTRIBUTION OF EFFECTIVE TRANSFER ENTROPY MATRICES



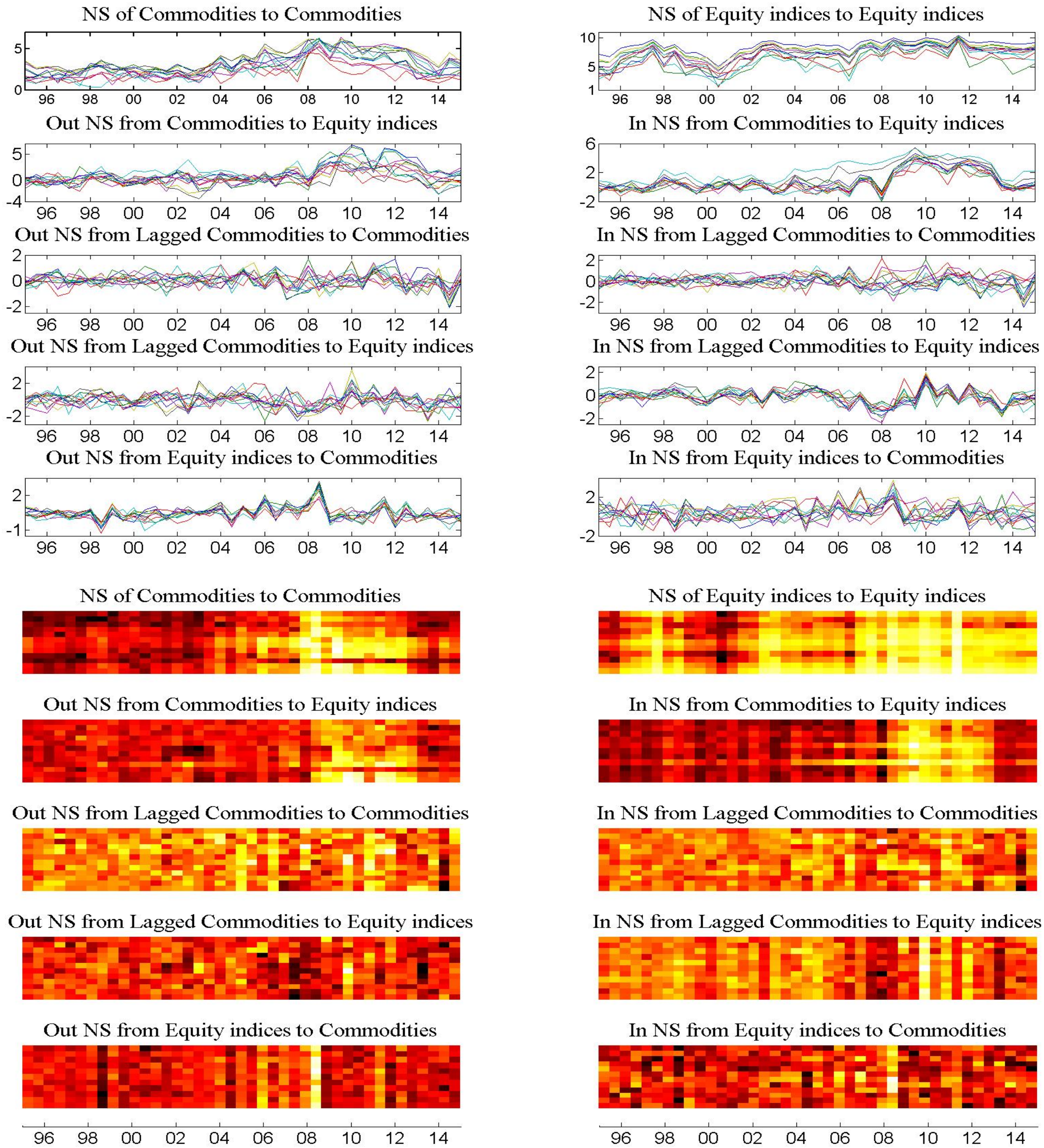
Notes: The probability distributions based on the effective transfer entropy matrices are depicted.

FIGURE 12: ASSET GRAPHS BASED ON THE EFFECTIVE TRANSFER ENTROPY MATRIX



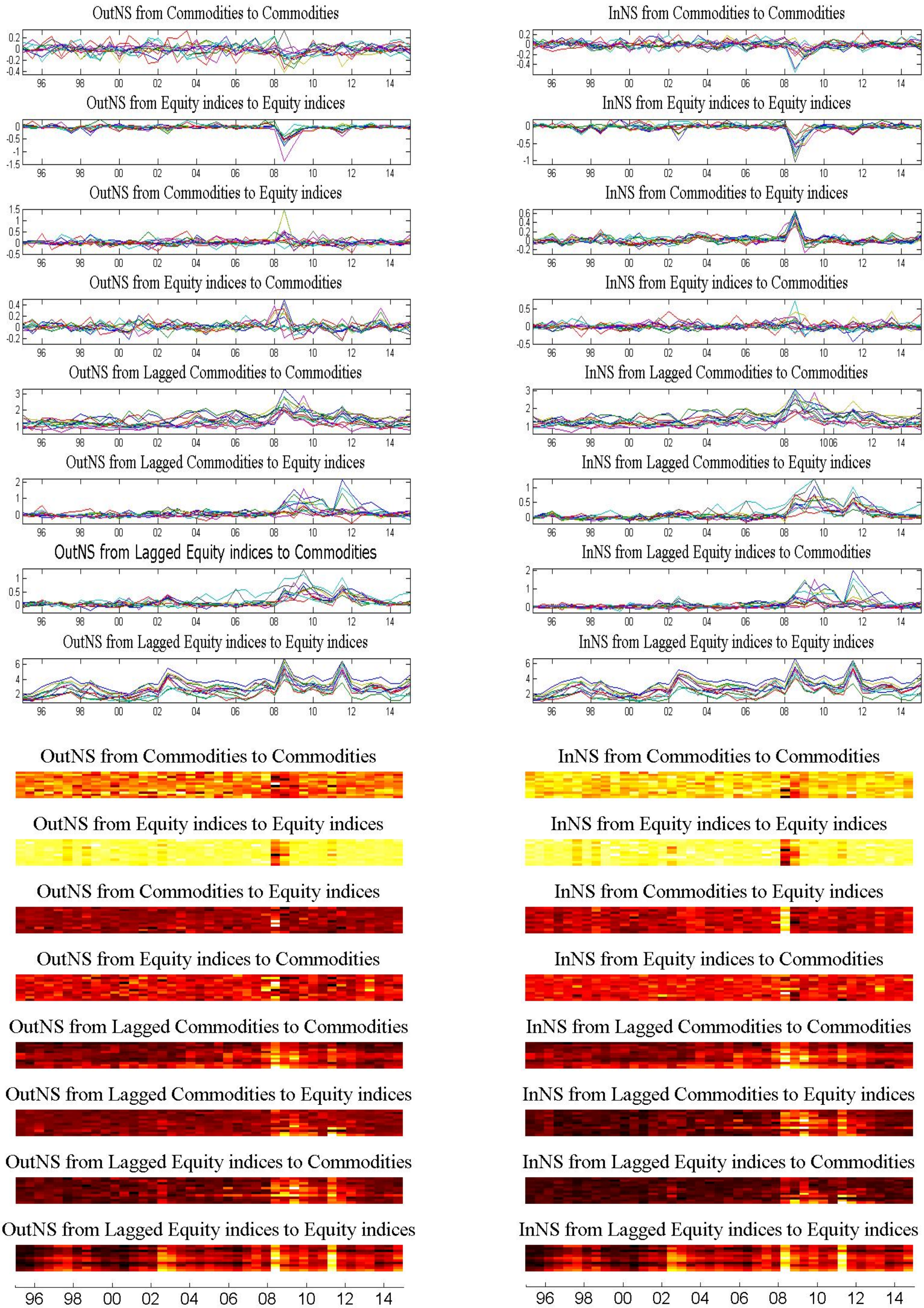
Notes: Asset graphs for the ETE matrix with thresholds from *T=0.2, 0.15, 0.1,* and *0.05*. The nodes associated with lagged variables are not represented, and all relations are assumed to be from a lagged variable to an original one.

FIGURE 13: NODE STRENGTHS AND HEAT MAPS OF CORRELATIONS BETWEEN VARIABLES IN TIME



Notes: The first ten pictures show the Node Strengths (NS) of the specified variables in time as individual graphs, and the last ten pictures show heat maps of these node strengths. For the correlation matrices of Commodities and Commodities and of Equity indices and Equity indices (first, second, eleventh and twelfth figures), there is just Node Strength, since the corresponding correlation matrices are symmetric. For the remaining figures, there are In and Out Node Strengths.

FIGURE 14: NODE STRENGTHS AND HEAT MAPS OF ETES BETWEEN VARIABLES IN TIME



Notes: The first sixteen pictures show the In and Out Node Strengths (InNS and OutNS) of the specified variables in time as individual graphs, and the last sixteen pictures show heat maps of these node strengths.