Les Cahiers de la Chaire / N°49

Systemic Risk in Energy Derivative Markets: A Graph-Theory Analysis

Delphine Lautier & Franck Raynaud

CHAIRE
Finance & Développement Durable

Systemic risk in energy derivative markets: a graph-theory analysis

Delphine Lautier [*1,2] and Franck Raynaud [†1,2]

[1] *University Paris-Dauphine*

[2] *DRM-Finance, UMR CNRS 7088*

Abstract

This article uses graph theory to provide novel evidence regarding market integration, a favorable condition for systemic risk to appear in. Relying on daily futures returns covering a 12-year period, we examine cross- and inter-market linkages, both within the commodity complex and between commodities and other financial assets. In such a high dimensional analysis, the graph theory enables us to understand the dynamic behavior of our price system. We show that energy markets - as a whole - stand at the heart of this system. We also establish that crude oil is itself at the center of the energy complex. Further, we provide evidence that commodity markets are becoming more integrated over time.

[*]D. Lautier is also member of the Fime laboratory and associate research fellow at Mines ParisTech.

[†]F. Raynaud is currently holding a postdoc position at University Paris-Dauphine.

The financial support of the French Energy Council and the Finance and Sustainable Development chair is gratefully acknowledged. We also thank the very helpful comments provided by the editor and the reviewers of the Energy Journal and the participants at the 4th Financial Risks International Forum (Paris, March 2011).

1 Introduction

This article examines the integration of organized derivative markets, because integration is a favorable condition for systemic risk to appear in. Concerns about such a phenomenon have recently grown, notably among energy commodities. These markets are supposed to be more and more integrated, both in regard to each other and to other markets. For some months now, fluctuations in the prices of energy products have often been invoked to explain corresponding fluctuations in soft commodities like soy, corn, or wheat. Furthermore, because commodities are nowadays considered a new class of assets, investors use them for diversification purposes. Therefore, the price fluctuations recorded in commodity markets might be, at least partially, explained by external events like the fall in stock prices or in interest rates.

In the context of our study, systemic risk is associated with the propagation of price shocks in the financial system. Organized markets are indeed characterized by the presence of a clearing house, whose most important economic function is the management of credit risk through the mechanisms of initial margins and margin calls. So, in our case, the event of importance is not the default of an economic entity and its consequences but rather the propagation of a price shock that, through a contagion phenomenon, could impair the performances of the two main financial services offered

by derivative markets, namely hedging against price fluctuations and price discovery[1]. Market integration and systemic risk are intimately linked to each other. Co-movements indeed are the first evidence of the risk of a "domino pattern." This pattern can destabilize several markets and even the entire financial system and cause governments and market monitoring institutions to act. The propagation of a price shock, however, does not call for the same crisis management and / or prevention policy when a contagion remains local or spreads turmoil into other markets. This is the reason why, in this article, we propose a holistic approach for systemic risk. We study integration simultaneously in three dimensions: space, time, and the maturity of the transactions. Such an analysis is crucial: it accounts for the eventuality that a price shock that occurs on a specific asset's physical market can spread, not only through its own futures market (at a local level), but also into other physical and / or paper markets, and *vice versa* (at a global level). To the best of our knowledge, this is the first time that such an approach has been envisaged.

A full comprehension of systemic risk can only be made through a large scale analysis that requires the manipulation of a huge amount of data. In our case, analyzing integration in three dimensions (3-D), on the basis of 14 derivative markets (six energy commodities, four agricultural commodities, and four financial assets), over a 12-year period, led us to setup a database containing more than 750, 000 prices. To accomplish this task, we rely on methods initially designed for statistical physics. These methods help to understand the behavior of complex systems. Well established in their original context, they are also used in finance and management[2]. These methods incited us to consider all futures prices, quoted in different places and with different maturities, as

[1]This definition is inspired by the one given by the Bank of International Settlements in its latest annual report (chapter 6, page 83).

[2]See, for example, the special issue on Complex Systems in *Management Science* (2007).

a complex dynamic system. Moreover, this consideration led us to a set of tools that proved very useful for the study of systemic risk: graph theory.

A graph is a mathematical representation of pairwise relations within a collection of discrete entities. Through this prism, the nodes in our graph are the daily price's returns and the links stand for distances, the latter being computed as a function of the correlations between the returns. This representation allows us to analyze, in the first part of the study, the integration of the markets and its evolution, thanks to the structure of the connections between the futures contracts. What is especially interesting here is that we can consider, simultaneously, all possible pairs of assets.

The dimension of the fully connected graph being high, we rely on a specific type of graph in the second part of the empirical study: a Minimum Spanning Tree (MST). A MST provides a way to extract the most important information contained in the initial graph. It is unique and corresponds to the shortest path covering all the nodes of the graph without loops. Such a tree is thus especially interesting for the study of systemic risk: it can be assimilated into the shortest and most probable path for the propagation of the price shock. To the best of our knowledge, this is the first time that this tool has been used this way.

The visualization of the MST and the computation of some specific measures, like allometric coefficients, make possible the analyzation of the organization of the trees. Two extreme configurations are used as references in this article. A chain-like organization signifies that, when it appears at one extremity of the price system, only one way exists for the price shock to propagate; before reaching the other extremity of the graph, the shock will have to cross each node. On the other hand, in a star-like organization, the paths for the transmission of fluctuations are less easy to predict. Here, the node located at the centre of the star is of crucial importance; whenever a

shock arises at this point, it might disseminate to the whole system! We first examine the MST according to these two ideal types of organizations. Then, given the time dependency of correlation-based graphs, we study their evolution over time and their robustness. Our first main results lie in the economic meaningfulness of the graphs. As we rely on a methodology that is unusual in finance and economics, this is particularly important. In the spatial as well as in the 3-D analyses, the trees are organized into sub-trees corresponding to the three sectors of activity under examination: energy commodities, agricultural products, and financial assets.

The second set of results, interesting for regulatory purposes, shows that energy products promote the connection between the different sectors. Moreover, crude oil stands at the centre of the energy complex. This commodity is thus at the heart of all concerns. A third category of results concerns the evolution of integration over time. In commodity markets, both spatial and maturity dimensions tend to be more integrated. Thus, the conditions for the appearance of systemic risk increase.

In Section 2 of this paper, we review the literature related to this article. Section 3 explains the data. Section 4 focuses on the methodology adopted for the study. In Section 5, we present the empirical results. Meanwhile, as traditional measures of statistical significance are not suitable given the choice of our methodology, we discuss their robustness. Section 6 presents the conclusions and policy implications.

2 Literature review

Our analysis has a relation to the different trends in the literature: graph theory, cross-market linkages, and intra-market linkages.

Our use of graph theory relies, first of all, on recent methods that originate in statis-

tical physics. In the last few years, many theoretical and numerical tools have been developed to investigate the behavior of complex dynamical systems in various areas. Among others, Albert, Jeong, and Barabàsi (2000) examine the tolerance of complex networks to errors and attacks. More recently, Buldyrev (2010) studies the catastrophic cascade of failures in interdependent networks. In contrast with these papers, our article focuses on financial markets. Several studies also investigate this domain. The first is Mantegna (1999), who uses MSTs also but applies them to the analysis of the cross-correlations of stock returns. Like Miceli and Susinno (2003), we rely on the filtering approach of the MST to construct a correlation-based classification. However, in our case, the economic entities under scrutiny are derivative markets (in the spatial as well as in the 3-D approaches) and/or futures contracts (in the maturity dimension). Instead, these two authors focus on banks and hedge funds. Further, as was done in a series of studies (see, for example, Onnela, Chakraborti, Kaski, Kertész, and Kanto (2003)), we take into account the time dependency of our correlation-based graphs, and we examine the robustness of the MST characteristics over time.

The graph theory has also been used in very few studies in finance and economics. Haigh and Bessler (2004) investigate spatial relations between markets on the basis of directed acyclical graphs. While this method is very interesting because it enables a causality analysis, it becomes very difficult when undertaking large scale studies. Their graphs indeed comprise no more than three nodes, whereas we are dealing with up to 215 in the 3-D analysis. More recently, Bech, Chapman, and Garratt (2010) examine the relations between banks according to their liquidity holdings. They rely on fully connected graphs and on a classification method similar to Google's PageRank procedure in order to give some explanation as to the functioning of the Canadian payment system. Cohen-Cole, Kirilenko, and Patacchini (2011) are the closest to our

work. They also analyze the topology of their graph, which is made of individual traders. As their graph is mainly a star-like one -which is not our case- a lot of attention is devoted to the centrality measure. They also compensate for the inability of traditional statistical measures through the use of different procedures. However, the latter are suited to their agent-based analysis and could not be used in our case.

Our article is also related to previous works on cross-market linkages. The question of whether commodity markets move in sync with one another and with other asset markets has received a lot of attention. Pindyck and Rotenberg (1990) study the herding behavior of investors on commodity derivative markets and show that the persistent tendency of commodity prices to move together can not be totally explained by the common effects of inflation, exchange rates, interest rates and other macroeconomic variables. Focusing on spatial integration, Jumah and Karbuz (1999) propose another approach: their study centers on the relations between the prices of raw materials negotiated in different places. These authors initiate several works on spatial integration based on the methodology of co-integration. The empirical tests generally conclude that commodity markets are more and more spatially integrated. We complement their analysis of the commodity markets by adding the maturity dimension.

More recently, co-movement between commodity price returns has been investigated in order to see whether or not speculative activities influence commodity prices. Korniotis (2009) compares the synchronization in the prices of exchange- and non-exchange traded metals, whereas Tang and Xiong (2011) focus on commodities that are in indexes and commodities that are "off-index", like Chinese commodities. The former shows that exchange- and non-exchange traded commodity prices exhibit similar structural breaks and price paths in the past decade, whereas the latter find that the increasing correlation in commodity prices is more pronounced for exchange-traded

assets. However, both conclude that cross market linkages become more intense. We corroborate their findings, in the sense that we also observe an increase in the correlations among commodity markets. Moreover, this trend started long before the 2008 crisis.

Buyukşahin, Haigh, and Robe (2010) extend this analysis of cross-market linkages with two kinds of investigations. First, they look at changes in the extent to which different groups of commodities (essentially energy, agricultural products, and metals) move in synchronization with each other. Second, they examine the co-movement between commodity markets and more traditional assets. Among other results, on the basis of a nonpublic trader-level database from the CFTC, they observe that hedge funds that trade in both equity and commodity markets help explain long-term linkages between these two categories of assets. Their portfolio approach to cross-commodity linkages complements contemporaneous work by Chong and Miffre (2010) and precede the analysis performed by Stoll and Whaley (2010) and by Tang and Xiong (2011). In the same vein, we analyze what happens between different groups of commodities, and between commodities and more traditional assets like equities, interest rates, exchange rates, and gold.

Stoll and Whaley (2010), like Tang and Xiong (2011), use publicly available data to ask whether the arrival of index traders in commodity futures markets brought about an increase in the co-movements between various commodities. Whereas the first find that commodity index flows have little impact on futures prices, the latter arrive at the opposite conclusion. In two recent papers, relying on their nonpublic dataset, Buyuksahin and Robe (2010), and Buyuksahin and Robe (2011) contribute to this debate by showing that the composition of the open interest helps explain the joint distribution of commodity and equity returns. They indeed show that some (more pre-

cisely, hedge funds) but not all types of traders (and more precisely not index traders) affect the correlations in price returns. Moreover, in the specific case of energy markets, Buyuksahin and Robe (2011) find considerable changes in the make-up of the open interest of energy futures between 2000 and 2010. They show that these fluctuations help explain the co-movements between returns in energy and equity markets. Unlike these papers, that try to find an explanation to the reasons explaining market integration, we rather focus on the way to appreciate integration in a very large scale analysis that includes the maturity dimension.

Indeed, integration also has a temporal dimension that has not been greatly explored until now. As early as 1992, Bradley and Lumpkin (1992) examine this question in the case of Treasury securities, with maturities ranging from 3 months to 30 years. Lautier (2005) investigates cross-maturity linkages in the term structure of commodity prices between 1999 and 2002. She examines the segmentation hypothesis, in the sense of Modigliani and Sutch (1966) and the propagation of price information along the price curve in the crude oil market. She shows that temporal integration progresses over time. Buyuksahin, Haigh, Harris, Overdahl, and Robe (2009) confirm this result and extend the analysis over a longer period of time. Moreover, they link this progression to the trader composition of futures market activity in the crude oil market. In this paper, focusing on the correlation of price returns, we extend these works to a large number of derivative markets. Such an analysis is very helpful when building term structure models. The tendency for segmentation in the price curve to disappear when transaction volume grows confirms the appropriateness of imposing common stochastic processes on such models whatever maturity is concerned (as was done, in the case of commodities, by Schwartz (1997), for example).

3 Presentation of the database

We select futures markets corresponding to three sectors: energy, agriculture, and financial assets. On the basis of the Futures Industry Association's reports, we retain those contracts whose characteristics are large transaction volumes over long time periods. In the absence of reliable spot data for most commodity markets, we always approximate the spot with the nearest futures prices, the latter thus linking our analysis to physical markets. We use Datastream to collect settlement prices on a daily basis. We rearrange the futures prices in order to reconstitute daily term structures; that is, the relation linking, at a specific date, several futures contracts with different delivery dates. Table (1) summarizes the characteristics of our database.

In a derivative market, the maturities of contracts usually rise through time. Indeed, the growth in the transaction volumes results in the introduction of new delivery dates. Thus, in order to have continuous time series, we have to remove some maturities from the database. Moreover, when performing spatial and 3-D analyses, we have to retain the longest common time period for all underlying assets, between 2000 and 2011. We also have to take away all observation dates that are not shared by all markets. Once these selections have been carried out, our database still contains more than $750,000$ prices.

Underlying assets	Exchange-Zone	Period	Maturities	Records
Light crude	CME-US	1998-2011	up to 84	3343
Brent crude	ICE-Eu	2000-2011	up to 18	2923
Heating oil	CME-US	1998-2011	up to 18	3227
Gasoil	ICE-Eu	2000-2011	up to 12	2950
Nat. gas (US)	CME-US	1998-2011	up to 36	3336
Nat. gas (Eu)	ICE-Eu	1997-2011	up to 9	3698
Wheat	CME-US	1998-2011	up to 15	3412
Soy bean	CME-US	1998-2011	up to 14	3370
Soy oil	CME-US	1998-2011	up to 15	3447
Corn	CME-US	1998-2011	up to 25	2960
Eurodollar	CME-US	1997-2011	up to 120	3689
Gold	CME-US	1998-2011	up to 60	3060
Exchange rate USD/EUR	CME-US	1999-2011	up to 12	3239
Mini SP500	CME-US	1997-2011	up to 6	3611

Table 1: Main characteristics of the collected data: nature of the underlying asset, trading place of the futures contract, localization of the exchange, time period, longest maturity (in months) and number of records per maturity. CME stands for Chicago Mercantile Exchange, ICE for Inter Continental Exchange, NYSE LIFFE for New York Stock Exchange - London International Financial and Futures Exchange. US stands for United States and Eu for Europe.

4 Methodology: Minimum Spanning Trees and the analysis of integration

In order to study the integration of derivative markets, we rely on graph-theory. Among the different tools this method provides, we select those that allow us to analyze market integration by using a 3-D approach. We first focus on the synchronous correlations of price returns. Having transformed these correlations into distances, we are able to draw a fully connected graph of the price system, where the nodes (vertices) of the graph represent the time series of futures prices. In order to filter the information contained in the graph, we then rely on MST (Mantegna (1999)). This tree can

be defined as the one providing the best arrangement of the network's different nodes.

4.1 Synchronous correlation coefficients of prices returns

The first step towards the analysis of market integration is, in our case, the computation of the synchronous correlation coefficients of price returns defined as follows:

$$\rho_{ij}(t) = \frac{\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle}{\sqrt{\left(\langle r_i^2 \rangle - \langle r_i \rangle^2\right)\left(\langle r_j^2 \rangle - \langle r_j \rangle^2\right)}}, \tag{1}$$

where i and j correspond to two different time series of futures returns. The daily logarithm price differential stands for price returns r_i, with $r_i = (\ln F_i(t) - \ln F_i(t - \Delta t))/\Delta t$, where $F_i(t)$ is the price of the futures contract at t. Δt is the lag between two consecutive trading days, and $< . >$ denotes the statistical average performed over time on the trading days of the study period.

For a given time period and a given set of data, we thus compute the matrix of $N \times N$ correlation coefficients C for all the pairs ij. C is symmetric with $\rho_{ij} = 1$ when $i = j$. Thus, $N(N-1)/2$ coefficients characterize C.

4.2 From correlations to distances

In order to use graph theory, we need to introduce a metric. The correlation coefficient ρ_{ij} cannot be used as a distance d_{ij} between i and j, because it does not fulfill the three axioms that define a metric (Gower (1966)): (1) $d_{ij} = 0$ if and only if $i = j$; (2) $d_{ij} = d_{ji}$, and (3) $d_{ij} \leq d_{ik} + d_{kj}$.

A metric d_{ij} can be extracted from the correlation coefficients through the following

non linear transformation:

$$d_{ij} = \sqrt{2\,(1 - \rho_{ij})}. \tag{2}$$

A distance matrix D is thus extracted from the correlation matrix C according to Equation (2). Both, C and D are $N \times N$ dimensional. Whereas the coefficients ρ_{ij} can be positive for correlated returns or negative for anti-correlated returns, the quantity d_{ij} that represents the distance between price returns is always positive. This distance matrix corresponds to the fully connected graph: it represents all the possible connections in the price system.

4.3 From fully connected graphs to Minimum Spanning Trees

A graph gives a representation of pairwise relations within a collection of discrete entities. A simple connected graph represents all the possible connections between N points with $N(N-1)/2$ links (or edges). Each point of the graph constitutes a node (or a vertex). The graph can be weighted in order to represent the different intensities of the links and / or nodes. In our case, these weights represent the distances between the nodes.

In order to understand the organizing principles of a system through its representation as a graph, the nodes need to be spanned. However, there are a lot of paths spanning a graph. For a weighted graph, the MST is the one spanning all the nodes of the graph without loops. This MST also has less weight than any other tree.

Through a filtering procedure that reduces the information space from $N(N-1)/2$ to $N-1$, the MST highlights the most relevant connections in the system. In our study, the MST provides the shortest path to linking all nodes. It discloses the underlying mechanisms of systemic risk: the MST represents the strongest links in terms of the

correlations of price returns. Thus, because this tree is unique, it can be considered the easiest path for the transmission of a price shock.

5 Empirical results

The first information that a MST provides is the kind of arrangement that exists between the vertices: its topology. Therefore, this section focuses on this topology and its consequences for systemic risk. Then, the focus switches to the dynamic behavior of the price system. Because traditional measures of statistical significance are not suitable for our methodology, we use the robustness of the different topologies and their economic meaning to measure significance.

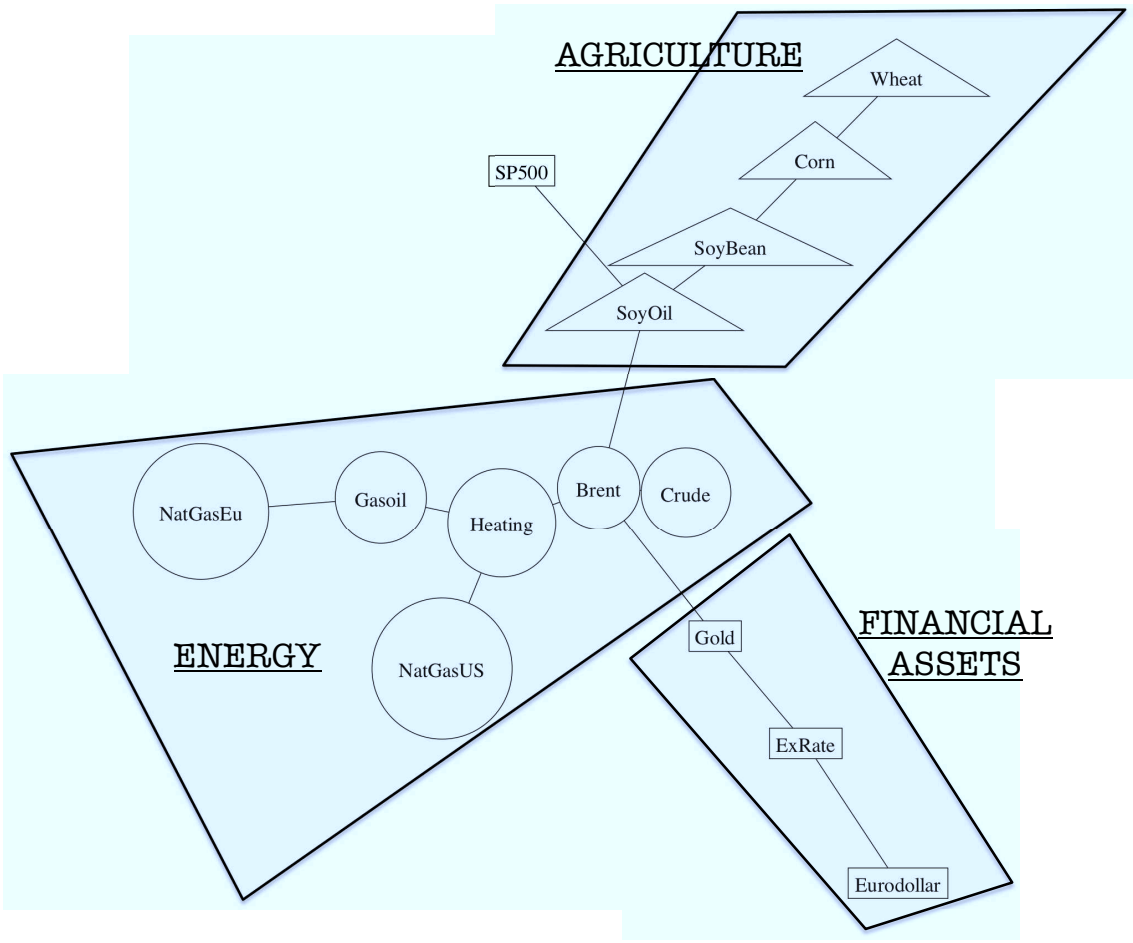
5.1 Topologies of the MST and their consequences for systemic risk

The first step in studying the MST lies in their visualization. Then, we use allometric coefficients to determine whether the MST are totally organized, totally random, or are situated somewhere between these two extreme topologies. In this first part of the study, we consider the whole time period as a single window and perform a static analysis.

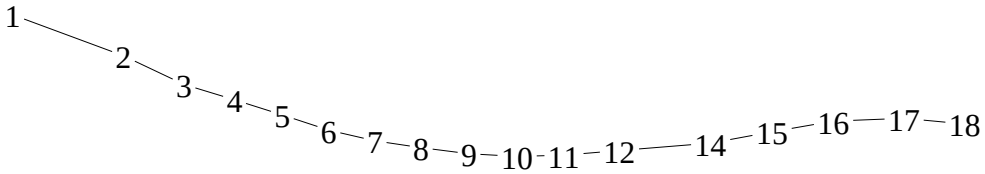
5.1.1 The emerging taxonomy in the three dimensions

Figure (1) presents the MST obtained for the spatial and maturity dimensions.

As far as the spatial dimension is concerned, all three sectors can be identified. Energy comprises American as well as European markets and is situated between agriculture



(a)



(b)

Figure 1: Static Minimum Spanning Trees built from the correlation coefficients of the prices returns. (a): MST in the spatial dimension (April 2001 - April 2011). (b): MST of the Brent crude in the maturity dimension (April 2000 - April 2011). The curvature only eases the visualization.

(on the top) and financial assets (on the bottom). The most connected node in the graph is the Brent, which makes it the best candidate for the transmission of price fluctuations in the tree (actually, the same could have been said for the Crude (Light crude), as the distance between these products is very short). Further, the energy sector is the most integrated of the three sectors because the distances between the nodes are short. The link between the energy and agricultural products passes through soy oil, which can be used for fuel. The link between commodities and financial assets passes through gold, which can be seen as a commodity but also as a reserve of value. The only surprising link comes from the S&P500, which is more correlated to soy oil than to financial assets.

Such a star-like organization leads to specific conclusions regarding systemic risk. A price move in the energy markets, situated at the heart of the price system, will have more impact than a fluctuation affecting peripheral markets such as interest rates or wheat.

Things are totally different in the maturity dimension. The results are illustrated by the example of the Brent crude, depicted by Figure (1)-b. For all contracts, the MSTs are linear and the maturities are regularly ordered from the first to the last delivery dates.

The results obtained in the maturity dimension give rise to three remarks. Firstly, the linear topology is meaningful from an economic point of view, as it reflects the presence of the Samuelson effect. In derivative markets, the movements in the prices of the prompt contracts are larger than the other ones. This difference results in a decreasing pattern of volatilities along the price curve and leads to higher correlations between the maturities that are the closest to each other. Secondly, this type of organization impacts the possible transmission of price shocks. The most likely path

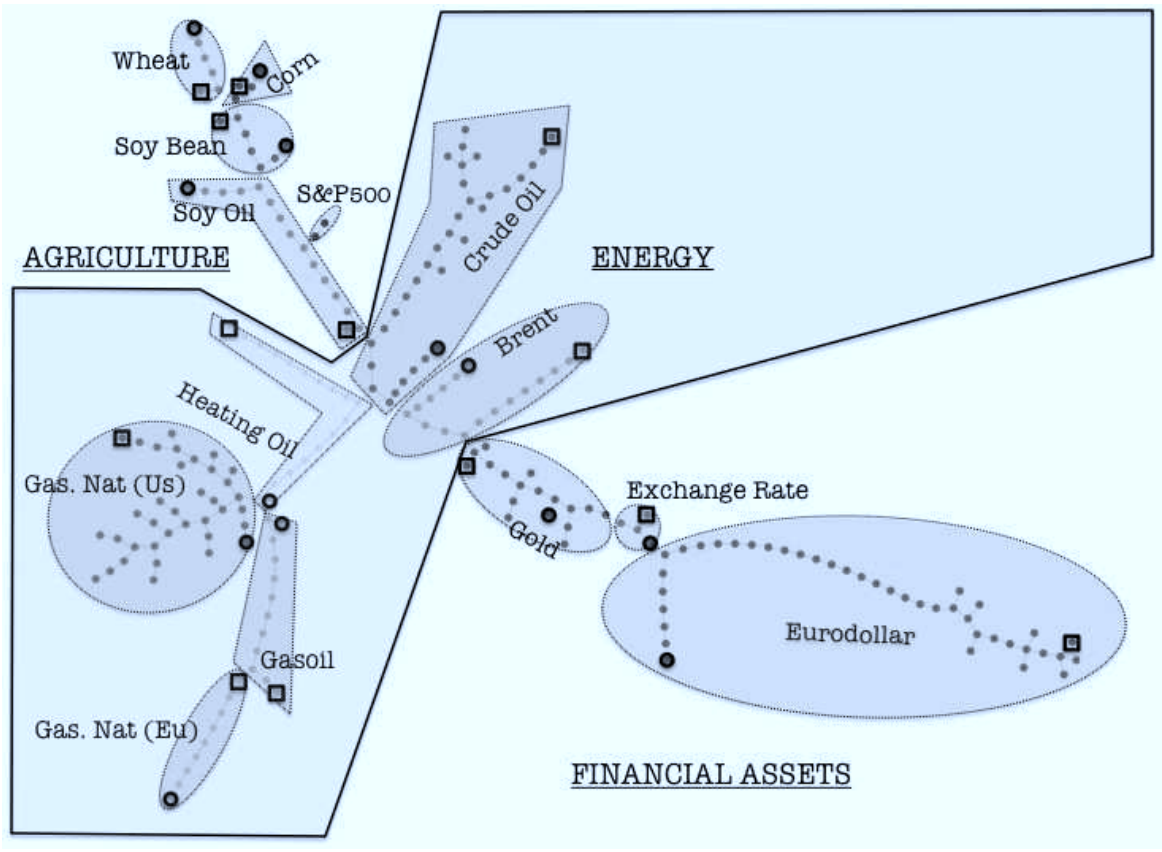


Figure 2: Static MST in 3-D (2000-2011). Each futures contract is enclosed in a shaded area with its name. The first and last maturities are respectively represented by a bold circle and a bold square. The distance between the nodes is set to unity.

for a shock is indeed unique and passes through each maturity, one after the other. Thirdly, the short part of the curves are less correlated with the other parts. This phenomenon can result from price shocks emerging in the physical market with the most nearby price being the most affected; it could also reflect noises introduced on the first maturity by investors in the derivative market.

Figure (2) represents the 3-D static MST. Its shape brings to mind the spatial dimension. However, it is enhanced by the presence of the different maturities available for each market. These maturities have a clear, linear organization. Again, the tree shows

a clear separation between the sectors. Three energy contracts, the crude oil (Light crude), the Brent and the Heating oil, are at the center of the graph. They are the three closest nodes in the graph. Whereas the maturities of each market primarily have a linear organization, the American natural gas behaves differently and displays an atypical topology with numerous ramifications.

It is interesting to see which maturities connect two markets or sectors. Economic reasoning suggests that two kinds of connections should exist: with the shortest and / or with the longest part of the curves. In the first case, the price system would be essentially driven by underlying assets; in the second, it would be dominated by derivative markets. However, a closer analysis of the 3-D trees does not provide evidence of either kind. Furthermore, the analysis of the trees at different periods does not lead to the conclusion that there is something like a pattern in the way connections occur.

5.1.2 Where does our prices system stand, between order and disorder?

The computation of the allometric coefficients of a MST provides a means of quantifying where this tree stands between two asymptotic topologies: star-like trees that are symptomatic of a random organization, and chain-like trees that show a strong ordering in the underlying structure.

Banavar (1999) developed the first model for the allometric scaling of a spanning tree. The first step of the procedure consists of initializing each node of the tree with the value of one. Then the root or central vertex of the tree must be identified. In what follows, the root is defined as the node that has the highest number of links attached to it. Starting from this root, the method consists of assigning two coefficients A_i and

B_i to each node i of the tree:

$$A_i = \sum_j A_j + 1 \text{ and } B_i = \sum_j B_j + A_i, \tag{3}$$

where j stands for all the nodes connected to i in the MST. The definition of the allometric scaling relation is the relation between A_i and B_i:

$$B \sim A^{\eta}, \tag{4}$$

where η is the allometric exponent. It represents the degree or complexity of the tree and stands between two extreme values: 1^+ for star-like trees and 2^- for chain-like trees.

Table (2) summarizes the allometric properties of the MSTs for each dimension, in the static as well as in the dynamic analyses. The top section corresponds to the maturity dimension with the information concerning the spatial and 3-D analyses at the bottom. In each case, we reproduce the exponents and their corresponding 95% confidence interval (CI). The error values are negligible in both cases, which confirms the robustness of the topologies: the economic meaningfulness found in the static analysis is stable. Also, Table (2) shows that the dynamic allometric exponents are consistent with the static ones.

Within the maturity dimension, the coefficients tend towards their asymptotic value: $\eta = 2^-$. However, they are a bit smaller than 2, due to finite size effects (there is a finite number of maturities). Such a result is probably due to arbitrage operations. When performed on the basis of contracts having the same underlying asset, such operations are easy and rapidly undertaken, thus resulting in a perfect ordering of the

Maturities	Static	$CI_{95\%}$	Dynamic	$CI_{95\%}$
Light crude	$1,994$	$1,9058 - 2.0822$	$1,91$	$1,8904 - 1.929$
Brent crude	$1,889$	$1,883 - 1,894$	$1,888$	$1,88 - 1,895$
Heating oil	$1,899$	$1,891 - 1,906$	$1,886$	$1,874 - 1,898$
Gasoil	$1,88$	$1,874 - 1,885$	$1,845$	$1,835 - 1,854$
Nat. gas (US)	$1,75$	$1,677 - 1,822$	$1,796$	$1,745 - 1,847$
Nat. gas (Eu)	$1,874$	$1,87 - 1,877$	$1,832$	$1,83 - 1,834$
Wheat	$1,864$	$1,609 - 2,118$	$1,761$	$1,694 - 1,827$
Soy bean	$1,848$	$1,661 - 2,034$	$1,68$	$1,623 - 1,736$
Soy oil	$1,889$	$1,883 - 1,894$	$1,856$	$1,832 - 1,879$
Corn	$1,88$	$1,874 - 1,885$	$1,772$	$1,731 - 1,813$
Eurodollar	$1,927$	$1,817 - 2,036$	$1,846$	$1,806 - 1,885$
Gold	$1,732$	$1,552 - 1,912$	$1,826$	$1,788 - 1,863$
Spatial	$1,493$	$1,383 - 1,602$	$1,621$	$1,574 - 1,668$
3D	$1,757$	$1,712 - 1,802$	$1,85$	$1,673 - 2,023$

Table 2: Allometric properties of the trees. Static and dynamical exponents (with their 95% confidence interval) first in the maturity dimension, then in the spatial dimension and in 3-D.

maturity dates.

With concern for the spatial dimension, the exponents indicate that even if Figure (1)-a exhibits a star-like organization, the shape of the MST is rather complex and stands exactly between the two asymptotic topologies. There is an ordering of the tree, which is well illustrated by the agricultural sector, which forms a regular branch. Finally, even if the topologies of the spatial and 3-D trees seem similar, they are quantitatively different. The allometric exponent for the 3-D tree is higher: the best fit from our data gives an exponent close to 1.757 as compared to the value of 1.493 in the spatial case. Thus, the topology in 3-D merges the organization in sectors induced by the spatial dimension and the chain-like organization arising from the maturity dimension.

5.2 Dynamical analysis of integration

Because it is based on correlation coefficients, our study of market integration is intrinsically time dependent. On the basis of the fully connected graph, we first examine the dynamic properties of the correlation coefficients, as well as the node's strength, which provides information on how close a given node is to all others. We then turn to the MSTs. In order to study the robustness of their topology, we compute their length that shows the state of the system at a specific time. Survival ratios also indicate how the topology of the trees evolves over time. Further, we propose a deeper investigation into the connections between markets in the 3-D analysis.

In what follows, we retain a rolling time window of $\Delta T = 480$ consecutive trading days.

5.2.1 Evolution of the correlation coefficients and their variances

In order to examine the time evolution of our system, we investigate the mean correlations of the returns and their variances (Sieczka and Holyst (2009)). The mean correlation $C^T(t)$ for the correlation coefficients ρ_{ij}^T in a time window $[t - \Delta T, t]$ can be defined as follows:

$$C^T(t) = \frac{2}{N(N-1)} \sum_{i<j} \rho_{ij}^T(t), \tag{5}$$

The variance $\sigma_C^2(t)$ of the mean correlation is given by:

$$\sigma_C^2(t) = \frac{2}{N(N-1)} \sum_{i<j} \left(\rho_{ij}^T(t) - C^T(t)\right)^2. \tag{6}$$

Where i and j correspond to two different time series of future returns. Figure (3) represents the mean correlation and its variance in the spatial dimension. It shows

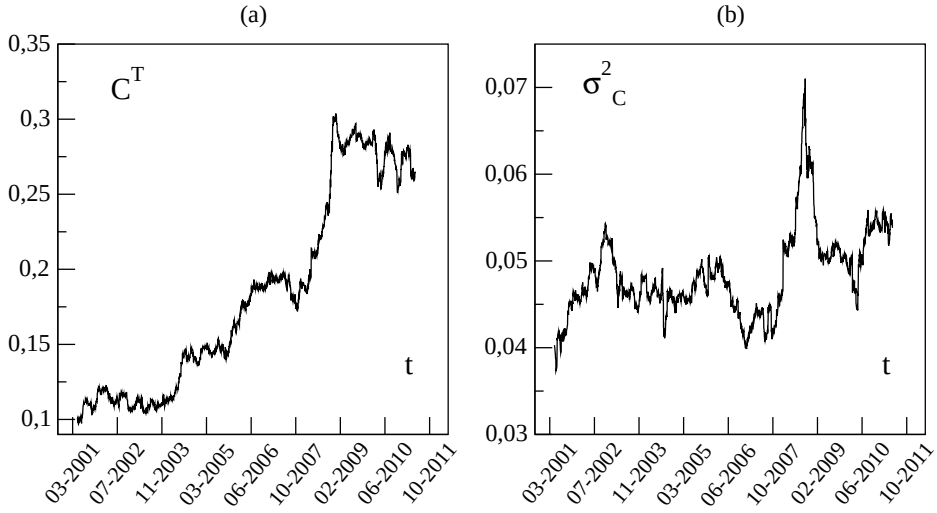


Figure 3: Correlation coefficients of the price returns in the spatial dimension, for all markets $2001 - 2011$. (a): Mean; (b): Variance.

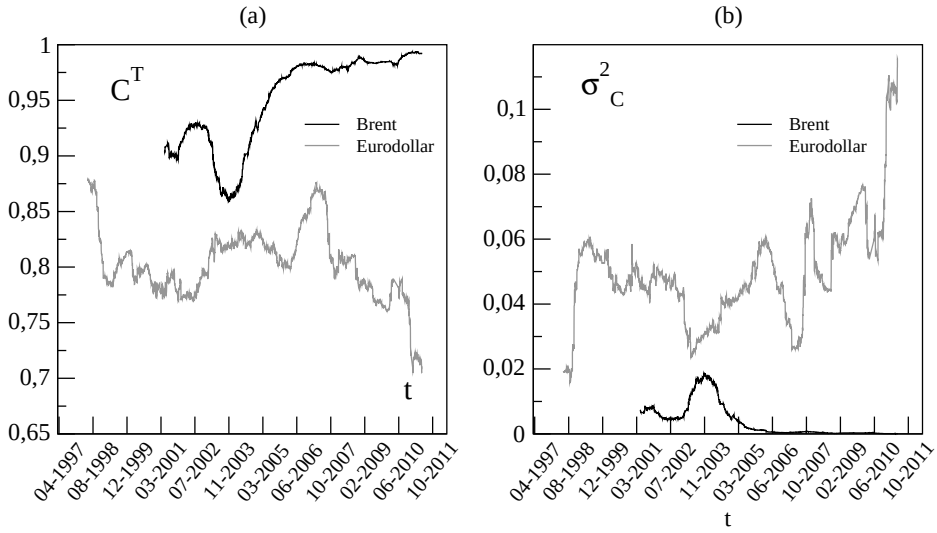


Figure 4: Correlation coefficients of the price returns in the maturity dimension for the Eurodollar (gray lines) and the Brent crude (black lines) $1998 - 2011$. (a): Mean; (b): Variance.

that the mean correlation increases over time with a huge rise from 2007 to the end of 2008. The variance is characterized by a peak in $2007-2009$, and reaches its maximum at the end of September 2008 just after the Lehman Brothers' bankruptcy.

Underlying assets	Mean	Min	Max	Variance	$\rho_{mean-variance}$
Light crude	$0,941$	$0,863$	$0,979$	$0,94\,10^{-3}$	-0.98
Brent crude	$0,952$	$0,859$	$0,99$	$0,17\,10^{-2}$	$-0,966$
Heating oil	$0,949$	$0,875$	$0,992$	$0,23\,10^{-2}$	$-0,953$
Gasoil	$0,956$	$0,891$	$0,991$	$0,65\,10^{-3}$	$-0,943$
Nat. gas (US)	$0,629$	$0,393$	$0,855$	$0,0255$	$-0,964$
Nat. gas (Eu)	$0,289$	$-0,16910^{-3}$	$0,769$	$0,097$	$-0,916$
Wheat	$0,926$	$0,814$	$0,993$	$0,21\,10^{-2}$	$-0,943$
Soy bean	$0,913$	$0,769$	$0,974$	$0,37\,10^{-2}$	$-0,794$
Soy oil	$0,948$	$0,826$	$0,997$	$0,23\,10^{-2}$	$-0,963$
Corn	$0,85$	$0,709$	$0,902$	$0,34\,10^{-2}$	$-0,96$
Eurodollar	$0,803$	$0,705$	$0,878$	$0,69\,10^{-2}$	$-0,826$
Gold	$0,984$	$0,939$	$0,996$	$0,13\,10^{-3}$	$-0,883$

Table 3: Characteristics of the correlation coefficients of the price returns in the maturity dimension. Mean correlation coefficients, min and max, variance and correlation between the mean and the variance.

We then examine the maturity dimension. First, we focus on the statistical properties of the correlation coefficients of two futures contracts, represented by Figure (4). They are very different for these contracts. The maturities of the Brent crude oil are more and more integrated over time: at the end of the period, the mean correlation is close to one. Such a trend does not appear for the Eurodollar contract. This is consistent with the peripheral position of the interest rate market in the correlation landscape. As far as crude oil is concerned, the level of integration becomes so strong that the variance decreases and exhibits an anti-correlation with the mean correlation. The result is totally different in the spatial case: the mean correlation and its variance are correlated in $2007-2008$ (Onnela et al. (2003) also observe this positive correlation during price growth and financial crises).

Table (3) summarizes the statistical properties of the mean correlations and variances for the 14 markets in the maturity dimension. The table confirms that, for almost

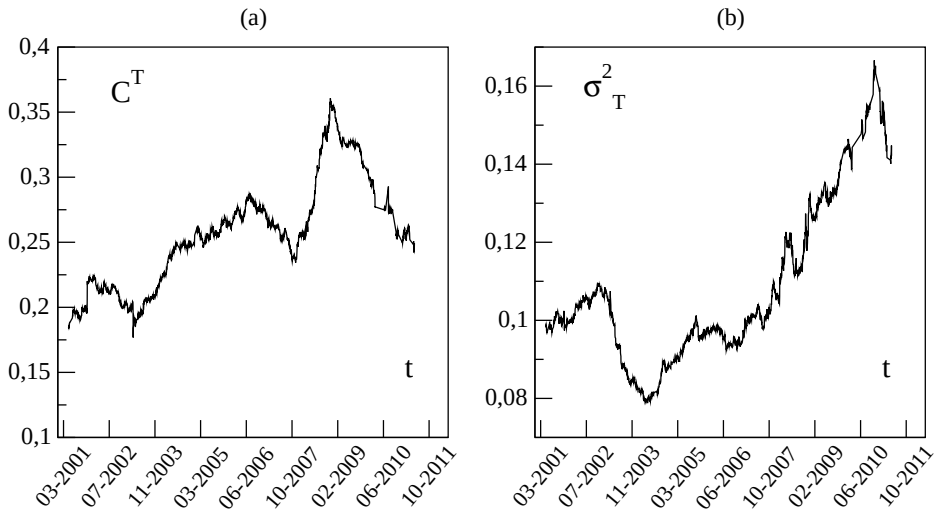


Figure 5: Correlation coefficients of the price returns in 3-D, $2001 - 2011$. (a): Mean; (b): Variance.

every contract, the mean correlation is very high and anti-correlated with its variance. However, the two natural gases exhibit more specific figures. Their correlation level is quite low compared with other markets, especially for European Natural Gas. Meanwhile, the variance is high.

Merging space and maturity, we also observe an important rise in the mean correlation and variance, as shown in Figures (5)-a and (5)-b. Moreover, these values have a correlation which likens the spatial tree to the 3-D tree, with a maximum correlation and variance at the end of 2008 (followed by a decrease in the mean correlation) and 2010 respectively. So, as our price system becomes more and more integrated, it becomes less stable.

5.2.2 How does markets closeness evolve?

The node strength, calculated for each node i, indicates the closeness of one node i to all others in the fully connected graph. It is defined as follows:

$$S_i = \sum_{i \neq j} \frac{1}{d_{ij}}. \tag{7}$$

In our case, the node strength provides information on the intensity of the correlations linking a given node to the others. When S_i is high, the node is close to all others. For the sake of simplicity, we use this measure in the spatial dimension only. As far as the maturity dimension is concerned, it was indeed not easy to represent the nodes strength for all futures contracts.

Figure (6) represents the time evolution of the nodes strength in the spatial dimension. The figure has been separated into four panels: the energy sector is at the top, with American products on the left and European ones on the right; the agricultural sector is at the bottom left and financial assets are at the bottom right.

Figure (6) shows that, at the end of the period, out of all the assets studied, the two crude oils and Heating oil show the greatest nodes strength. However, since 2010, the American node strength has decreased, which indicates a difference in the connectivity of the two crude oils. The petroleum products are followed by soy oil, other agricultural assets, the S&P500 contract, gold, the exchange rate USD/EUR, and the gasoil. A remarkable evolution is the sharp rise in the equity connectivity in the post-Lehman period, as opposed to $2001-2007$. This finding corroborates those of Buyukşahin et al. (2010), Buyuksahin and Robe (2010) and Tang and Xiong (2011). Finally, the more distant nodes are those representing the Eurodollar and the two natural gases.

As far as the time evolution of the node strength is concerned, the sectors exhibit

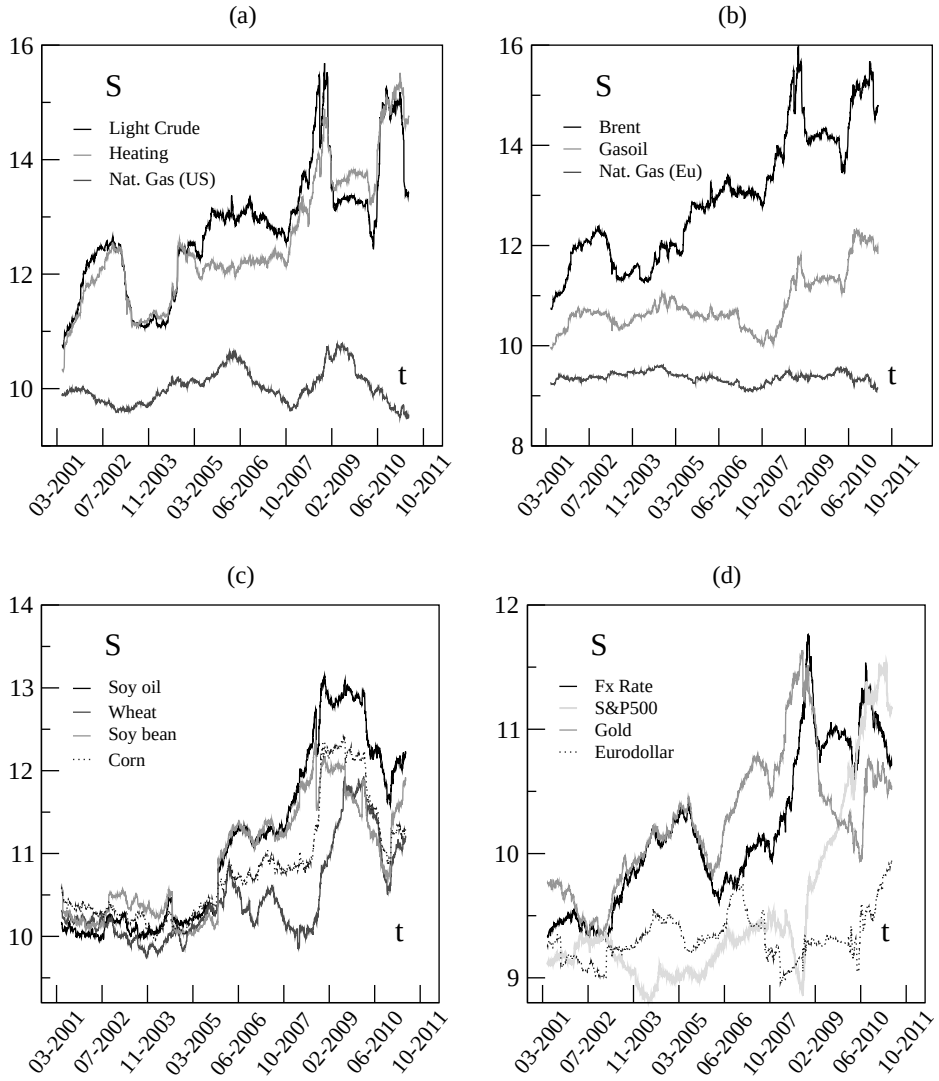


Figure 6: Nodes strength of the markets in the spatial dimension $2001-2011$. (a): American energy products; (b): European energy products; (c): Agricultural products; (d): Financial assets.

different patterns: the integration movement, characterized by an increase in this measure, emerges earlier for the energy sector than for the agricultural sector. However, it decreases for energy at the end of the period (especially for the Light crude oil). The nodes strength of the agricultural products is characterized by a plateau from the middle of 2009 to the beginning of 2010, followed by a drawdown until the

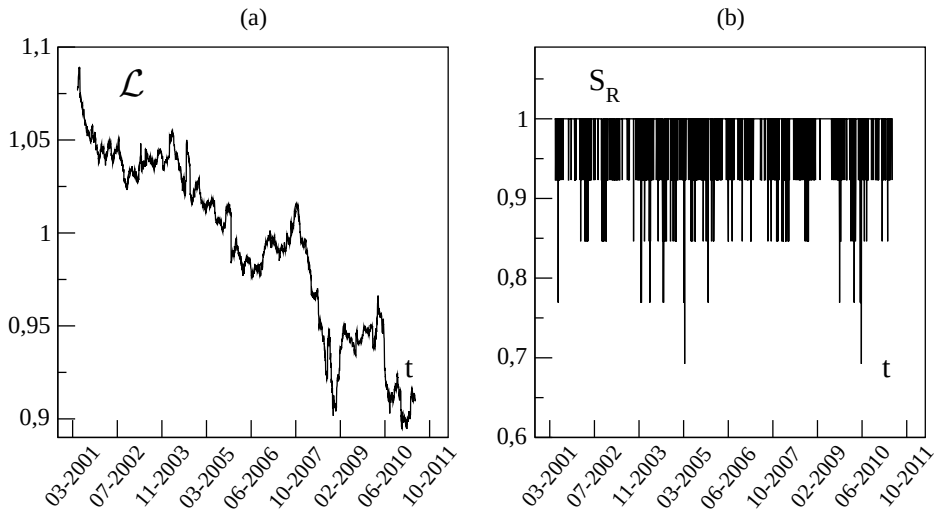


Figure 7: Spatial dimension $2001-2011$. (a): Normalized tree's length; (b): Survival ratios.

Fall 2010. Last but not least, most of the products exhibit a strong increase, except for natural gas and interest rate contracts. Thus, whereas the core of the graph becomes more and more integrated, the peripheral assets do not follow this movement.

5.2.3 How does the length of the Minimum Spanning Trees behave?

Let us now examine some of the properties of the filtered information. The normalized tree's length can be defined as the sum of the lengths of the edges belonging to the MST:

$$\mathcal{L}(t) = \frac{1}{N-1} \sum_{(i,j)\in MST} d_{ij}, \tag{8}$$

where t denotes the date of the construction of the tree and $N-1$ is the number of edges. The length of a tree is longer as the distances increase, and consequently when correlations are low. Thus, the more the length shortens, the more integrated the system is. On the contrary, in the case of random co-movements, the length of the tree is equal to $\sqrt{2}$.

Figure (7)-a represents the dynamic behavior of the normalized length of the MSTs in their spatial dimension. The general pattern is that the length decreases, which reflects the integration of the system. This information confirms what was observed in the fully connected graph on the basis of the nodes strength. In adition, it shows that the most efficient transmission path for price fluctuations becomes shorter as times goes on. A more in-depth examination of the figure also shows a very important decrease between October 2006 and October 2008, as well as significant fluctuations in September and October 2008. We leave the analysis of such events for future studies. In the maturity dimension, as integration increases, the normalized tree's length also diminishes. Figures (8)-a and -b illustrate this phenomenon by representing the evolutions recorded for the Eurodollar contract and Brent crude. As far as the interest rate contract is concerned, the tree's length first increases, then in mid-2001 it drops sharply and remains fairly stable after that date. For crude oil, the decrease is constant and steady, except for a few surges.

5.2.4 Survival ratios and the stability of the prices system

The robustness of the MSTs over time is examined by computing the single-step survival ratio of the links, S_R. This quantity refers to the fraction of edges in the MST, that survives between two consecutive trading days (Onnela (2003)):

$$S_R(t) = \frac{1}{N-1} \left| E(t) \cap E(t-1) \right|. \tag{9}$$

In this equation, $E(t)$ refers to the set of the tree's edges at date t, $\cap$ is the intersection operator, and $|.|$ gives the number of elements contained in the set. The survival ratios are very important for our study. Under normal circumstances, the topology of the

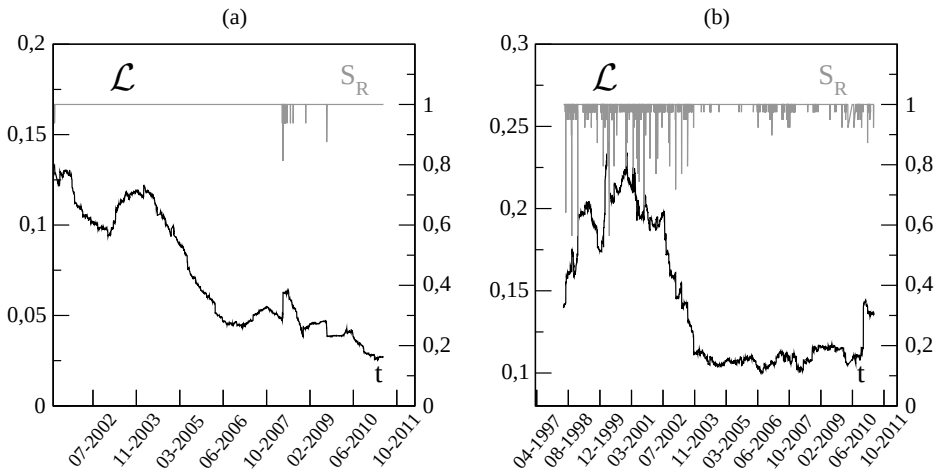


Figure 8: Maturity dimension, normalized tree's length (left axis) and survival ratios (right axis) for the Brent (a) and the Eurodollar (b).

trees should be very stable and the value of the survival ratio around one.

Figure (7)-b represents the evolution of the survival ratios in the spatial dimension. Most of the time, they remain constant, with a value greater than 0.9 in more than 96% of the cases. Thus, the topology of the trees is very stable: the shape of the most efficient path for the transmission of price shocks does not change much over time. However, it is possible to identify two events where 30% of the edges has been shuffled. Such a result could be further investigated in a specific study devoted to price shocks. In the maturity dimension, Figures (8)-a and -b represent the survival ratios and the length of the MSTs for two representative contracts, that is, the Brent crude and the Eurodollar contracts. As far as the former is concerned, the organization of the MST is very stable: the survival ratio is equal to one most of the time, with few exceptions since the end of 2008. Meanwhile, the tree shrinks, in the metric sense: the market becomes more and more integrated. Again, what happens with the Eurodollar is very different. In $2000 - 2001$, around the period of the internet crisis, when the length of the tree increases, the tree also becomes more spaced out. This sparseness

comes with an important amount of reorganizations, and fluctuations in the survival ratios are greater as the length increases. Lastly, as far as the 3-D trees are concerned, the survival ratios do not give any further information. However, we propose a more specific analysis of these trees, based on a pruning method.

5.2.5 Interconnections between markets in 3-D: pruning the trees

Concerning the stability of the trees, especially in 3-D, when focusing on the whole system, it is interesting to distinguish between reorganizations occurring in a specific market (i.e., between different delivery dates of the same contract) and reorganizations that change the nature of the links between two markets or even between two sectors. However, Equation (9) gives the same weight to every kind of reorganization, whatever its nature. The trouble is, a change in intra-maturity links does not have the same meaning, from an economic point of view, as a movement affecting the relation between two markets or sectors. Because we are interested in the strong events that affect the markets, inter-market and inter-sector reorganizations are more relevant. Thus, in order to distinguish between these categories of displacements, we "prune" the 3-D trees, that is, we only consider the links between markets, whatever the maturity considered. This pruning does not mean that maturity is removed from the analysis, but that the information on the specific maturity that is responsible for the connection between markets is no longer identified.

Pruned trees enable us to compute the length and the survival ratios on the sole basis of market links. The comparison between Figures (7) and (9) shows that the level of integration is higher in the pruned tree than in the spatial dimension. More precisely, near the Lehmann crisis, the network is more contracted when only intra-market linkages are considered, than in the spatial case.

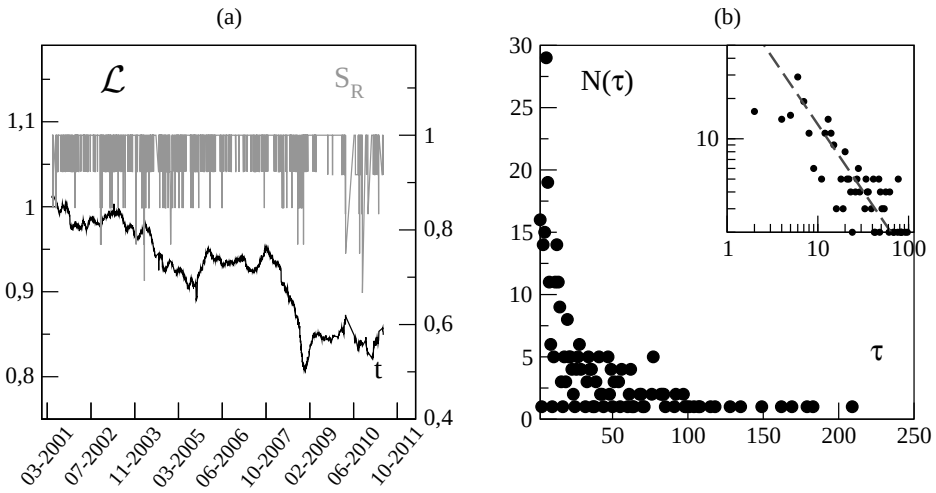


Figure 9: Properties of the pruned trees. (a): Survival ratios and pruned tree length; (b): Number of occurrences of stable periods of length τ. Inset: same as in (b), but in log-log scale. The dashed line corresponds to τ^{-1}

Another interesting characteristic of the pruned survival ratios is that they provide information on the lifetime of a configuration of such trees. In what follows, we measure the length of time τ between two different consecutive configurations and compute the occurrences $N(\tau)$ of these periods. Figure (9)-b displays our results. It shows that $N(\tau)$ decreases quickly with τ. The dashed line in the inset (in log-log scale) indicates that $N(\tau)$ is roughly proportional to τ^{-1}. An economical interpretation of this result is that there is not a typical lifetime for a new configuration of the MST.

Further, another result, as far as robustness is concerned, lies in the analysis of those links that are the most frequently responsible for the reorganization of the pruned trees. With 14 markets, there are 91 inter-market links in our system. Some of them - 26 - never appear. Among the remaining 65 links, some appear very frequently and, on the contrary, others display very few occurrences. Figure (10) reproduces these two categories of links and the frequency in which they appear in the MSTs. The most robust links have a frequency equal to one, which means that the links are always

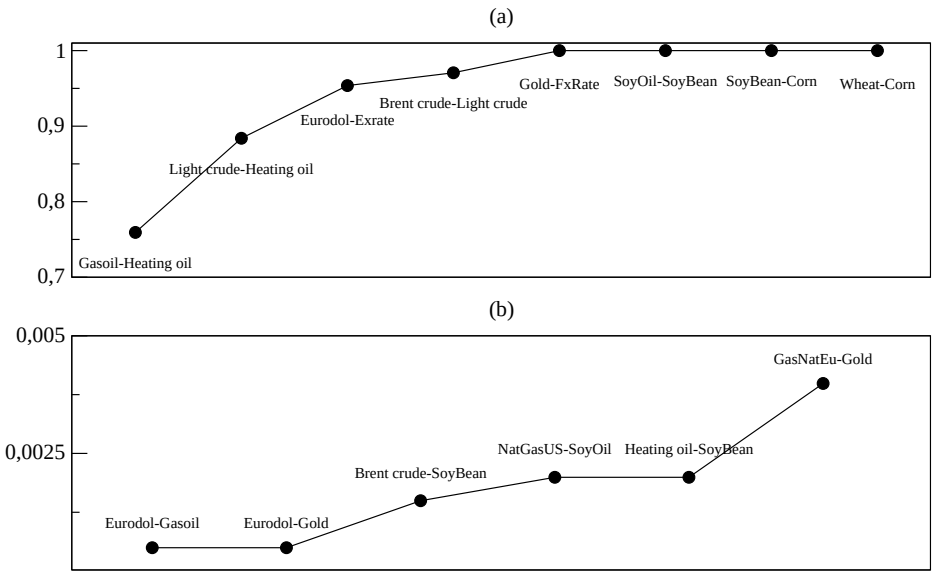


Figure 10: Frequency of the links in the pruned MST. Figure (a): frequency greater than 0,75. Figure (b): frequency lower than 0,005

present. They mainly correspond to the agricultural sector, with the following pairs: wheat and corn; soy beans and corn; soy oil and soy beans. The link between gold and the USD/EUR exchange rate is also always present. As expected, the relation between the two crude oils is very stable, with a frequency greater than 0.9. The same is true for the links between the interest rates and the exchange rate. This is also reasonable from an economic point of view, as interest rates are embedded in forward exchange rates.

6 Conclusions and policy implications

In this article, we study systemic risk in energy derivative markets based on two choices. First, we focus on market integration, which is a favorable condition for the propagation of a price shock. Second, based on the fact that previous studies mainly focus on the spatio-temporal dimension of integration, we introduce the maturity di-

mension and perform a three-dimensional analysis.

In the context of an empirical study that aims to understand the organization and the dynamic behavior of a highly dimensional price system, our methodology, based on graph-theory, has proven very useful. Needless to say, it could also be used for higher systems, or for smaller ones. Moreover, Minimum Spanning Trees are particularly interesting in our framework, as they are filtered networks enabling us to identify the most probable and the shortest path for the transmission of a price shock.

We show that the topology of the MSTs tends towards a star-like organization in the spatial dimension, whereas chain-like trees characterize the maturity dimension. These two topologies merge in the 3-D analysis, and all of them are very stable. The star-like organization reproduces the three different sectors studied (energy, agriculture, and finance), and the chain-like structure reflects the presence of a Samuelson effect. The reasoning behind these findings is very important: the robustness of our methodology is embedded in these topologies.

Another contribution is to show that the American and European crude oils are both at the center of our large scale system; furthermore, they provide the links with the subsets of agricultural products and financial assets. Thus, crude oil is the best candidate for the transmission of price shocks. If such a shock appears at the periphery of the graph, it will necessarily pass through crude oil before spreading to other energy products and sectors. Moreover, a shock will have an impact on the whole system that will be all the greater the closer it is to the heart of the system.

Another important conclusion is that integration increases significantly on both the spatial and maturity dimensions. Such an increase can be observed in the whole price system. It is even more evident in the energy sector (with the exception of the natural gas markets), which becomes highly integrated at the end of our period. Thus, as time

goes on, the heart of the price system becomes stronger whereas the peripheral assets do not change significantly. Moreover, the level of integration is higher in the maturity than in the spatial dimension: arbitrage operations are easier with standardized futures contracts written on the same underlying asset.

These results have very important consequences for regulatory as well as for diversification and hedging purposes.

Whereas the move towards integration started some time ago (and there is probably no way to refrain it), knowledge of its characteristics remains poor, especially from a holistic perspective. On the basis of this study, regulation authorities can see that their actions against systemic risk will not have the same impact depending on the market they are addressing. They should pay particular attention to the heart of the system: this is the place where price shocks spread more easily to other markets.

As far as diversification is concerned, portfolio managers must concentrate their positions on the most stable parts of the graph. More precisely, the benefits associated with diversification that rely on sub-indexes and focus on specific sectors of activity (agricultural products for example) should be more recurrent than those associated with large scale indexes.

Lastly, one important concern for hedging is the information conveyed by futures prices and its meaning. The increasing integration of derivative markets is probably not a problem for hedging purposes, until a price shock appears somewhere in the system. In such a case, the information related to the transmission path of the shock is important, as prices might temporarily become irrelevant.

These results call for further work. First, survival ratios make it possible to identify a few events leading to important reconfigurations of the trees. A thorough analysis of such phenomenon can provide the regulating authorities with a battery of stylized facts

about the different possible manifestations of prices shocks and the signs announcing a future shock. Second, now that we have defined the paths for shock transmission, it is important to obtain directed graphs to determine the direction of the propagation of price movements. Third, a focus on the gas market, which exhibits a striking pattern of cross-maturity connections, can be of interest for energy specialists.

References

Albert, R., H. Jeong, and A.-L. Barabàsi (2000). Error and attack tolerance of complex networks. *Nature 406*.

Banavar, J. (1999). Size and form in efficient transportation networks. *Nature 399*, 130.

Bech, M., J. Chapman, and R. Garratt (2010). Which bank is the "central" bank? *Journal of monetary Economics 57*, 352–363.

Bradley, M. and S. Lumpkin (1992). The treasury yield curve as a cointegrated system. *Journal of Financial and Quantitative Analysis 27*, 449–63.

Buldyrev, S. (2010). Catastrophic cascade of failures in interdependent networks. *Nature 464*, 1025–1028.

Buyukşahin, B., M. Haigh, and M. Robe (2010). Commodities and equities: Ever a "market of one"? *Journal of Alternative Investment*, 76–95.

Buyuksahin, B. and M. Robe (2010). Speculators, commodities and cross-market linkages. *Working Paper, American University, November*.

Buyuksahin, B. and M. Robe (2011). Does "paper oil matter?" energy markets' financialization and equity-commodity co-movements. *Working Paper, American University, (Revised) July*.

Chong, J. and J. Miffre (2010). Conditional returns correlations between commodity futures and traditional assets. *Journal of Alternative Investments 12*, 61–75.

Cohen-Cole, E., A. Kirilenko, and E. Patacchini (2011). Are networks priced? network topology and systemic risk in high liquidity markets. *Working Paper, University of Maryland, College Park*.

Gower, J. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika 53*(3/4).

Haigh, M. and D. Bessler (2004). Causality and price discovery: An application of directed acyclic graphs. *Journal of Business 77*, 1099–1121.

Jumah, A. and S. Karbuz (1999). Interest rate differentials, market integration, and the efficiency of commodity futures markets. *Applied Financial Economics 9*.

Korniotis, G. (2009). Does speculation affect spot prices levels? the case of metals with and without futures markets. *Finance and Economics Discussion Series Working Paper 2009-29, Board of Governors of the Federal Reserve*.

Lautier, D. (2005). Segmentation in the crude oil term structure. *Quarterly Journal of Finance IX*(4), 1003–1020.

Mantegna, R. (1999). Hierarchical structure in financial markets. *Eur. Phys. J. B 11*.

Miceli, M. and G. Susinno (2003). Using tree to grow money. *Risk 16*, 11.

Modigliani, F. and R. Sutch (1966). Innovation in interest rate policy. *American Economic Review 56*, 178–197.

Onnela, J. (2003). Dynamic asset trees and black monday. *Physica A 324*.

Onnela, J.-P., A. Chakraborti, K. Kaski, J. Kertész, and A. Kanto (2003). Dynamics of market correlations: taxonomy and portfolio analysis. *Phys. Rev. E 68*(5), 056110.

Pindyck, R. and J. Rotenberg (1990). The excess co-movement of commodity prices. *Economic Journal 100*.

Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 923.

Sieczka, P. and J. A. Holyst (2009). Correlations in commodity markets. *Physica A 388*.

Stoll, H. and R. Whaley (2010). Commodity index investing and commodity futures prices. *Journal of Applied Finance*.

Tang, K. and W. Xiong (2011). Index investing and the financialization of commodities. *NBER, Working Paper No. 16325*.