



**HAL**  
open science

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

Delphine Lautier, Franck Raynaud

► **To cite this version:**

Delphine Lautier, Franck Raynaud. Systemic risk in energy derivative markets: a graph theory analysis. *Energy Journal*, 2012, 33 (6), pp.215-239. halshs-00738201

**HAL Id: halshs-00738201**

**<https://shs.hal.science/halshs-00738201>**

Submitted on 3 Oct 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Delphine Lautier <sup>\*1,2</sup> and Franck Raynaud <sup>†1</sup>

<sup>1</sup>*University Paris-Dauphine*

<sup>2</sup>*DRM-Finance, UMR CNRS 7088*

## Abstract

Considering it as a necessary condition for systemic risk to appear, this article focuses on integration in energy derivative markets, through a three-dimensional approach: observation time, space and the maturity of futures contracts. Such a method indeed makes it possible to investigate prices shocks in the physical as well as in the paper markets. In order to understand the underlying principles and the dynamic behavior of our prices system, we select specific tools of the graph-theory. More precisely, we use minimum spanning trees as a way to identify the most probable path for the transmission of prices shocks. We study the organization of these trees and their dynamic behavior. Examining

---

<sup>\*</sup>Professor Delphine Lautier is also member of the Fime-FDD laboratory and associate research fellow at Mines ParisTech

<sup>†</sup>Dr Franck Raynaud is currently holding a postdoc position at University Paris-Dauphine

three categories of underlying assets (energy and agricultural products, as well as financial assets), we find that crude oil stands at the heart of the system, and that energy markets are becoming more and more integrated.

*Financial support of the French Energy Council is gratefully acknowledged*

## 1 INTRODUCTION

Considering it as a necessary condition in order for systemic risk to appear, this article examines integration in energy derivative markets. Concerns about such phenomenon have recently grown in derivative markets, notably on energy commodities. The latter are supposed to be more and more integrated, both as regards each other and as regards other markets. For some months now, fluctuations in the prices of energy products have often been invoked to explain those of soft commodities like soy, corn or wheat. Moreover, since commodities are nowadays considered as a new class of assets, they are used by portfolio managers for diversification purposes. Part of the price fluctuations recorded in commodity markets might therefore be explained by external events like the fall in stock prices or in interest rates.

In order to fully understand it, integration is comprehended through a three-dimensional approach: observation time, spatial relationships and the maturity of futures contracts. Such an analysis is crucial as it makes it possible to take into account the eventuality that a prices shock occurring on a specific asset's physical market can spread, not only through its own futures market, but also onto other physical and / or paper markets, and vice versa. To the best of our knowledge, this is the first time that such an approach is performed.

Taking into account three dimensions requires collecting a huge volume of data and

understanding the behavior of complex evolving systems. This explains why we use recent methods originated from statistical physics. Many theoretical and numerical tools indeed have been developed recently in order to investigate the behavior of dynamic complex systems. Among these tools, complex networks have been used to study the interconnections between a wide variety of entities, such as social networks or electricity networks (Albert and Barabási (2002)). For example, Albert, Jeong, and Barabási (2000) examines the tolerance of complex networks to errors and attacks. More recently Buldyrev (2010) study catastrophic cascade of failures in interdependent networks. Since the pioneer work of Mantegna (1999) physicists also started to apply the graph-theory to financial markets. Until now however, systemic risk in derivative markets has never been investigated on this basis.

We thus choose, in the field of statistical physics, several measures which we found relevant for our study of market integration. Firstly we use minimum spanning trees (MST) as a way to filter the information given by the correlation matrix of price returns. In our case, these specific networks are especially interesting: i) correlations are an important dimension of integration; ii) while constructing a MST, we dramatically compress the amount of information, which is appreciable when working with a huge amount of data. Secondly, we study the topology of the filtered networks. The organization between the different nodes of the graph produces interesting information. Thirdly, given the time dependency of the MST, we study their evolution over time. Our first main results lie in the economic meaningfulness of the emerging taxonomy. In the spatial as well as in the 3-D analyses, the trees are organized into sub-trees which correspond to the different sectors of activity under examination (energy commodities, agricultural products, and financial assets). More importantly, especially for regulatory purposes, the connection between the different sectors is always insured by

energy products. Meanwhile, the analysis of the maturity dimension provides evidence of a chain-like organization of the trees and, more precisely, a hierarchical ordering of the futures contracts based on their delivery dates. Such results are very important, as they are a key justification for the use of our methodology.

A second set of important results lies in the structure of integration that is observed. The identification of the central node of the tree is crucial in such a study, as if a shock emerges at this central node, it will have a more important impact than anywhere else. Moreover, if integration is a prerequisite for systematic risk to occur, then the central node is at the heart of all concerns. Our empirical study shows that, in the spatial as well as in the 3-D analysis, between 2000 and 2009, among the 14 different underlying assets under examination, crude oil stands at the center.

A third category of results concerns the evolution of integration over time. Several measures lead us to the conclusion that both spatial and maturity dimensions tend to be more integrated. Moreover, the analysis of the robustness of the trees' topology over time shows that during a crisis, the trees shrink topologically, on the spatial as well as on the maturity dimensions, thus reflecting an intensification of the system's integration. At the same time, the trees become less stable: their topology temporarily witnesses abrupt and significant changes.

In section 2 of this paper, we briefly review the previous literature on market integration and complex networks in finance and physics. Section 3 presents the data. Section 4 focuses on the methodology adopted for the study of market integration and prices shocks. In section 5, we present the empirical results. Section 6 is devoted to conclusions and policy implications.

## 2 PREVIOUS STUDIES ON INTEGRATION AND COMPLEX NETWORKS

In this section we give a brief overview, firstly on the literature on integration, and secondly on the use of statistical physics in finance.

The economic literature has investigated the question of integration through different ways. Pindyck and Rotemberg (1990) began to study the herding behavior of investors on commodity derivative markets. Their seminal work shows 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 macro-economic variables. This article has inspired several other researches on co-movement. Yet, in this kind of work the identification of the relevant economic variables is tricky. This could explain why empirical tests do not really succeed in concluding that there is herding behavior in commodity markets.

Focusing on spatial integration, Jumah and Karbuz (1999) initiated another approach to the systemic risk in commodity markets. Such a study is centered on the relationships between the prices of raw materials negotiated in different places. The authors initiated several works on spatial integration, based on the methodology of the co-integration. The empirical tests show that commodity markets are more and more spatially integrated. Among these articles, a special mention must be made of Haigh and Bessler (2004) as they use the graph theory in order to investigate spatial relationships. Their method (Directed Acyclic Graphs) is very interesting because it allows causality analysis. However, it becomes very difficult to undertake for large scale studies. In the same vein, Büyüksahin, Haigh, and Robe (2010) examined the links between equity and commodity markets.

Integration has also a temporal dimension, in the sense of the preferred habitat theory (Modigliani and Sutch (1966)). In Lautier (2005), the author studied the segmentation of the term structure of commodity prices and examined the propagation of shocks along the prices curve, on the crude oil petroleum market. She showed that temporal integration progresses through time.

Thus, while these studies confirm that it is highly likely that integration is progressing, none of them tried to simultaneously study spatial and maturity evolutions.

In statistical physics, the minimum spanning tree is heavily used in order to understand the evolution of complex systems, especially financial assets.

In his pioneer work, relying on MST, Mantegna (1999) investigates cross correlations of asset returns and identifies a clustering of the companies under investigation. In Bonnano (2004), the authors use this correlation based method in order to examine stocks portfolios and financial indexes at different time horizons. They also apply this method in order to falsify widespread markets models, on the basis of a comparison between the topological properties of networks associated with real and artificial markets. The filtering approach of the MST can also be used in order to construct a correlation based classification of relevant economic entities such as banks or hedge funds, as in Miceli and Susinno (2003). As far as commodities are concerned, Siczka and Holyst (2009) recently proposed a study of commodities clustering based on MST. They found evidence of a market synchronization. However, their database contains commodities characterized by low transaction volumes, which can introduce noise in the correlation matrix. Moreover, they do not investigate the maturity dimension. Lastly, the robustness over time of the minimum spanning tree's characteristics has also been examined in a series of studies (see for example Kullmann, Kertész, and Kaski (2002) and Onnela (2003a)). All these works constitute an interesting basis for the study of

systemic risk in three dimensions.

### 3 PRESENTATION OF THE DATABASE

For our empirical study, we selected futures markets corresponding to three sectors, namely energy, agriculture and financial assets. On the basis of the Futures Industry Association's monthly volume reports, we retained those contracts characterized by the largest transaction volumes, over a long time period. Moreover, in the absence of reliable spot data for most commodity markets we always approximate, in this study, the spot prices with the nearest futures prices. Such an approximation is very common in finance.

We used Datastream in order to collect settlement prices on a daily basis. Moreover, we rearranged the futures prices in order to reconstitute daily term structures, *i.e.* the relationship linking, at a specific date, several futures contracts with different delivery dates. Table (1) summarizes the main characteristics of our database.

With such a database, one of the difficulties comes from the fact that prices curves are shorter at the beginning of the period. Indeed, over time, the maturities of contracts usually rise on a derivative market. The growth in the transaction volumes of existing contracts results in the introduction of new delivery dates. Thus, in order to have continuous time series, we had to remove some maturities from the database. Moreover, when performing spatial and 3-D analyses, we had to retain the longest common time period for all underlying assets, between 2000 and 2009. Once these selections have been carried out, our database still contained more than 655,000 prices.



Underlying asset	Exchange-Zone	Period	Maturities	Records
Light crude oil	CME-US	1998-2009	up to 84	2965
Brent crude	ICE-EU	2000-2009	up to 18	2523
Heating oil	CME-US	1998-2009	up to 18	2835
Gasoil	ICE-EU	2000-2009	up to 12	2546
Nat. gas (US)	CME-US	1998-2009	up to 36	3140
Nat. gas (Eu)	ICE-EU	1997-2009	up to 9	3055
Wheat	CME-US	1998-2009	up to 15	3026
Soy bean	CME-US	1998-2009	up to 14	2977
Soy oil	CME-US	1998-2009	up to 15	3056
Corn	CME-US	1998-2009	up to 25	2569
Eurodollar	CME-US	1997-2009	up to 120	3056
Gold	CME-US	1998-2009	up to 60	2877
Exchange rate €/€	CME-US	1999-2009	up to 12	2864
Mini SP500	CME-US	1997-2009	up to 6	3011

Table 1: Main characteristics of the collected data: Nature of the assets, trading location (ie United States - US - or Europe - EU), time period, longest maturity (in months). CME stands for Chicago Mercantile Exchange, ICE for Inter Continental Exchange, NYSE LIFFE for New York Stock Exchange - London International Financial and Futures Exchange.

## 4 METHODOLOGY: MINIMUM SPANNING TREES AND THEIR USE FOR INTEGRATION ANALYSIS

In order to study the integration of derivative markets, we rely on the graph-theory. Among the different tools provided by this method, we select those allowing us to analyze market integration using a three-dimensional approach. We first decided to focus on the synchronous correlation of price returns. Having transformed these correlations into distances, we were able to draw a fully connected graph of the prices system, where the nodes of the graph represent the time series of futures prices. In order to filter the information contained in the graph, we then rely on specific graphs:

Minimum Spanning Trees (MST) Mantegna (1999). Such a tree can be defined as the one providing the best arrangement of the network's different points.

#### 4.1 Synchronous correlation coefficients of prices returns

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

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

In the spatial dimension,  $i$  and  $j$  stand for the nearby futures prices of pairs of assets (like crude oil or corn), whereas they stand for pairs of delivery dates in the maturity dimension. They are a mix of the two in the three-dimensional analysis. 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  $t$ .  $\Delta t$  is the time window, and  $\langle . \rangle$  denotes the statistical average performed other time, on the trading days of the study period.

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

#### 4.2 From correlations to distances

In order to use the graph-theory, we needed to introduce a metric. The correlation coefficient  $\rho_{ij}$  indeed 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)):

- $d_{ij} = 0$  if and only if  $i = j$ ,
- $d_{ij} = d_{ji}$
- $d_{ij} \leq d_{ik} + d_{kj}$

A metric  $d_{ij}$  can however be extracted from the correlation coefficients through a non linear transformation. Such a metric is defined as follows:

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

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

### 4.3 From full connected graphs to Minimum Spanning Trees

A graph gives a representation of pairwise relationships within a collection of discrete entities. A simple connected graph represents all the possible connections between  $N$  points, with  $N - 1$  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. Such weights can represent the distances between the nodes. In order to understand the organizing principles of a system through its representation as a graph, the latter needs to be spanned, *i.e.* all its nodes need to be traversed. However, there are a lot of paths spanning a graph. For a weighted graph, the minimum

spanning tree (MST) is the one spanning all the nodes of the graph, without loops. It has less weight than any other tree.

Through a filtering procedure (the information space is reduced from  $N(N - 1)/2$  to  $N - 1$ ), the MST thus reveals the most relevant connections of each element of the system. In our study, it provides for the shortest path linking all nodes. Thus, it can be seen as a way of revealing the underlying mechanisms of systemic risk: the minimal spanning tree indeed is the easiest path for the transmission of a prices shock.

## 5 EMPIRICAL RESULTS

The first information given by a minimum spanning tree is the kind of arrangement found between the vertices: its topology. In the first part of our study, we thus focus on the topology of the MST and its consequences for systemic risk. The second part is devoted to the dynamic behavior of the prices system.

### 5.1 Topologies of Minimum Spanning Trees and their consequences for systemic risk

The first information given by a minimum spanning tree is the kind of arrangement found between the vertices. Therefore, the first step in studying MST lies in their visualization. We then use the allometric coefficients method in order to determine whether a MST is totally organized, totally random, or is 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 thus perform a static analysis.

### 5.1.1 The emerging taxonomy in the three dimensions

The visualization of the trees is a very important step, as it addresses the meaningfulness of the taxonomy that emerges from the system. Before going further, let us make two remarks: first, we are considering links between markets and/or delivery dates belonging to the MST. Thus, if a relationship between two markets or maturities does not appear in the tree, this does not mean that this relation does not exist. It just does not correspond to a minimal distance. Second, our results naturally depend on the nature and number of markets chosen for the study.

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

As far as the spatial dimension is concerned. In Figure (1)-a three sectors can be identified: energy is at the bottom. It gathers American as well as European markets and is situated between agriculture (on the left) and financial assets (mainly on the right). Moreover, the most connected node in the graph is Brent, which makes it the best candidate for the transmission of price fluctuations in the tree (actually, the same could have been said for Crude, as the distance between these products is very short). Last but not least, the energy sector seems the most integrated, as the distances between the nodes are short. The link between the energy and agricultural products passes through soy oil. This is interesting, as the latter can be used for fuel. The link between commodities and financial assets passes through gold, which is also meaningful, as gold can be seen as a commodity but also as a reserve of value. The only surprise 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 prices move appearing in the energy markets, situated at the heart of the price system,

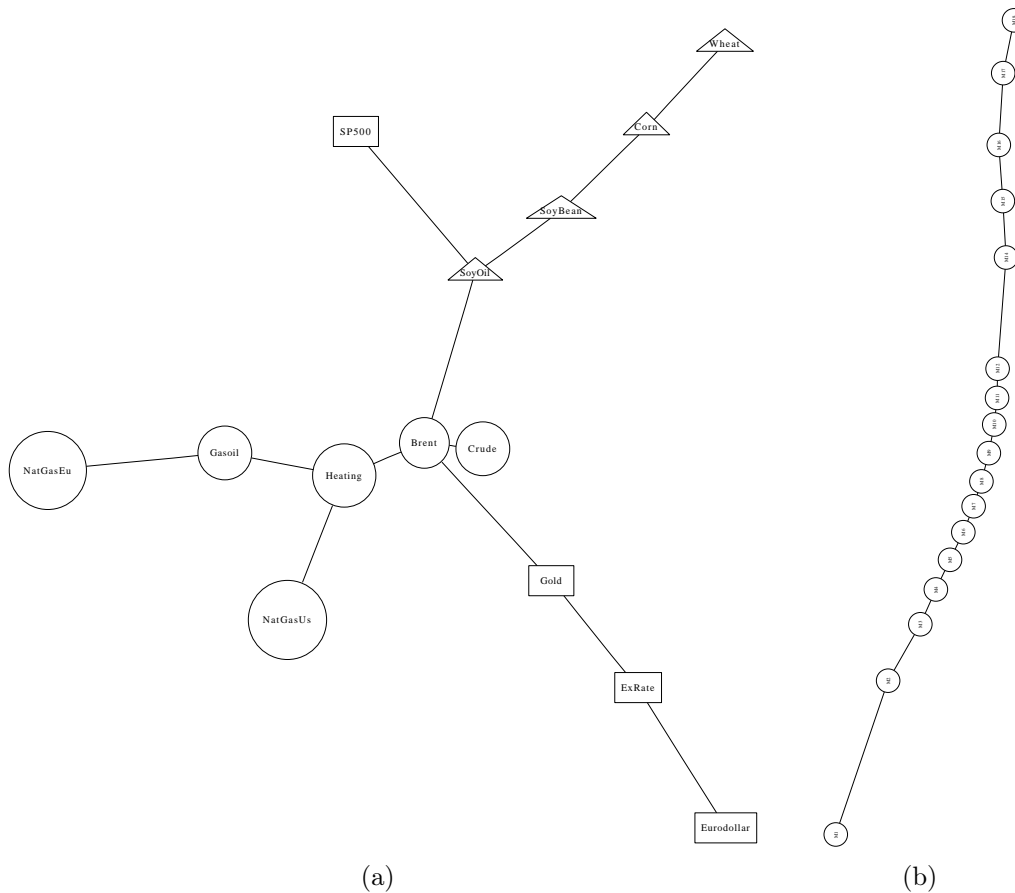


Figure 1: Static minimum spanning trees. Left panel: MST for the spatial dimension, built from the correlation coefficients of prices returns, 30/04/01-01/08/09. Right panel: MST on the maturity dimension, built from the correlation coefficients of the Brent crude oil, 01/04/2000-06/11/09. Maturities increase from bottom to top.

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 case of the Brent crude, depicted by Figure (1)-b (as it was not possible to give an illustration for each of the 14 contracts under examination, we retained a representative one). Here, the MST are linear and the maturities, for all contracts, are regularly

ordered from the first to the last delivery dates.

The results obtained on the maturity dimension give rise to three remarks. Firstly, this linear topology 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 results in a decreasing pattern of volatilities along the prices curve. Secondly, this type of organization impacts the possible transmission of prices shocks. The most likely path for a shock is indeed unique and passes through each maturity, one after the other. Thirdly, the short part of the curves are found to be less correlated with the other parts. This phenomenon can result from prices 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.

Let us now turn to the three-dimensional analysis. Figure (2) represents the 3-D static MST. Its shape brings to mind that observed in the spatial dimension. However, it is enhanced by the presence of the different maturities available for each market. The latter are clearly linearly organized. As previously, the tree shows a clear separation between the sectors. Three energy contracts, American crude oil, European crude oil and American heating oil, are found at the center of the graph. They are the three closest nodes of the graph.

It would have been interesting to know which maturities connect two markets or sectors. Economic intuition suggests two kinds of connections: on the shortest and / or on the longest part of the curves. In the first case, the price's 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 of expected organization. Moreover, the analysis of the tree at different periods does not lead to the conclusion that there is something like a pattern in the

way connections occur. Further investigations are thus necessary in order to study the links between markets and sectors more precisely. We offer an response to this problem at the end of this section.

### 5.1.2 Were 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, which are symptomatic of a random organization, and chain-like trees, which reveal a strong ordering in the underlying structure.

The first model of the allometric scaling on a spanning tree was developed by Banavar (1999). The first step of the procedure consists in initializing each node of the tree with the value 1. Then the root or central vertex of the tree must be identified. In what follows, the root is defined as the node having the highest number of links attached to it. Starting from this root, the method consists in assigning two coefficients  $A_i$  and  $B_i$  to each node  $i$  of the tree, where:

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

$j$  stands for all the nodes connected to  $i$  in the MST. The allometric scaling relation is defined as the relationship between  $A_i$  and  $B_i$ :

$$B \sim A^\eta, \quad (4)$$

$\eta$  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.



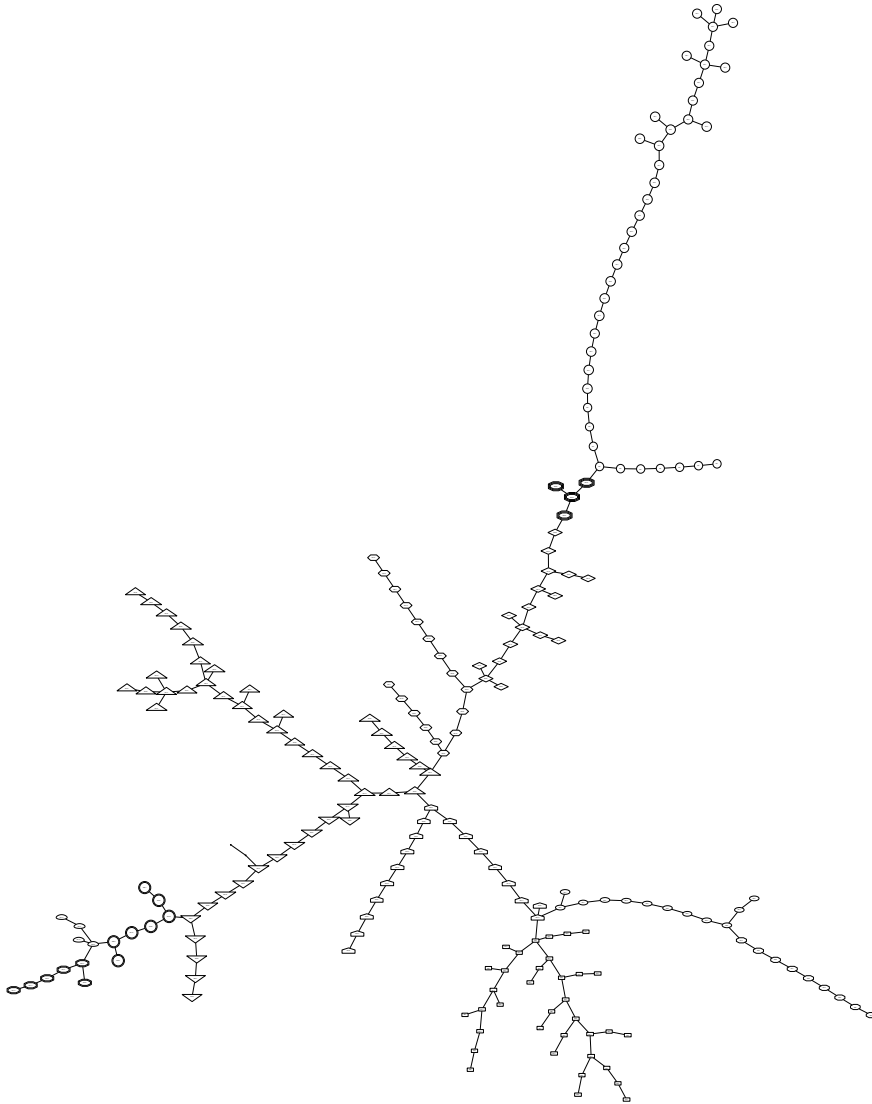


Figure 2: MST in three dimensions, 2000-2009. The different futures contracts are represented by the following symbols: empty circle: Eurodollar, point: Mini SP500, octagon: Natural Gas (Eu), ellipse: Gasoil, box: Natural Gas (Us), hexagon: Brent, triangle: Crude, house: Heating Oil, diamond: Gold, inverted triangle: Soy Oil, triple octagon: Exchange Rate, double circle: Soy Bean, double octagon: Wheat, egg: Corn. For a given futures contract, all maturities are represented with the same symbol. The distance between the nodes is set to unity.

Table (2) summarizes the allometric properties of the MST for each dimension. The left panel reproduces the different exponents and gives the error resulting from a

non-linear regression. Figure (3) gives an illustration of the allometric coefficients in 3-D. The dashed line corresponds to the best fit with an exponent equal to 1.85. The figure shows that the coefficients are well described by the power law with an exponent.

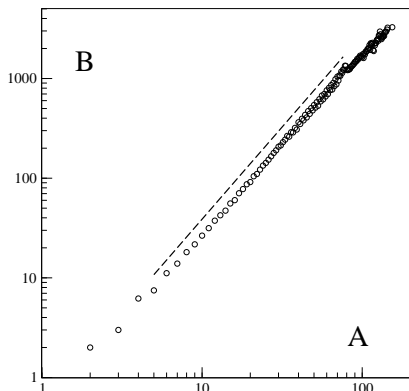


Figure 3: Dynamical allometric scaling of the 3-D in log-log scale. The dashed line corresponds to the exponent 1,85.

	Static	Dynamic
Crude	1,994 ±0,045	1,906 ±0,013
Brent	1,889 ±0,003	1,904 ±0,095
Heating	1,899 ±0,004	1,917 ±0,001
Gasoil	1,88 ±0,003	1,943 ±0,02
Nat. gas (Us)	1,75 ±0,037	1,74 ±0,018
Nat. gas (Eu)	1,874 ±0,002	1,886 ±0,059
Wheat	1,864 ±0,13	1,761 ±0,125
Soy bean	1,848 ±0,095	1,822 ±0,095
Soy oil	1,889 ±0,003	1,886 ±0,032
Corn	1,88 ±0,003	1,834 ±0,024
Eurodollar	1,927 ±0,056	1,913 ±0,011
Gold	1,732 ±0,092	1,908 ±0,013
Spatial	1,493 ±0,056	1,621 ±0,024
3D	1,757 ±0,023	1,85 ±0,09

Table 2: Allometric properties of the trees. Static and dynamical exponents for each futures contract (maturity dimension), as well as spatial and 3D.

As far as the spatial dimension is concerned, the exponents indicate that even if Figure (1) seems to exhibit 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.

Within the maturity dimension, the coefficients tend towards their asymptotic value  $\eta = 2^-$ . They are however 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 maturity dates.

Finally, even if the topologies of the spatial and 3-D trees seem similar, they are quantitatively different. The allometric exponent for the three-dimensional tree is higher: the best fit from our data gives an exponent close to 1.757, which must be compared to the value of 1.493 for the spatial case. Thus, the topology of our system, in 3-D, is rather complex. It is the result of two driving forces: the star-like organization 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 markets integration is intrinsically time dependent. On the basis of the entire graph, we firstly examine the dynamical properties of the correlation coefficients, as well as the node's strength, which provides information on how close a given node is from the other nodes. We then turn to the MST. In order to study the robustness of their topology, we compute their length, which reveals the state of the system at a specific time. Survival ratios also indicate how the topology of the trees evolves over time. Finally, this dynamical study gives us the possibility to propose a deeper investigation of the connections between markets in the three-dimensional analysis.

In what follows, we retain a rolling time window with a size 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 investigated the mean correlations of the returns and their variances (Sieczka and Holyst (2009)). The mean

correlation  $C^T(t)$  for the correlation coefficient  $\rho_{ij}^T$  in a time window  $[t - \Delta T, t]$  can be defined as follows:

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

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

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

Figure (4) represents the mean correlation and its variance on the spatial dimension. It shows that the mean correlation increases over time, especially after 2007. The variance exhibits a similar trend. Moreover, it reaches its maximum on the 09/19/2008, four days after the Lehman Brothers' bankruptcy.

We then examine the maturity dimension. Firstly, we focus on the statistical properties of the correlation coefficients of two futures contracts, represented by Figure (5). They are very different for these contracts. The maturities of Brent crude oil are more and more integrated over time: at the end of the period, the mean correlation is close to 1. 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 was totally different in the spatial case: the mean correlation and its variance were correlated (Onnela (2003b)) also observe such a positive correlation during prices growth and financial crises).

Table (3) summarizes the statistical properties of the mean correlations and vari-

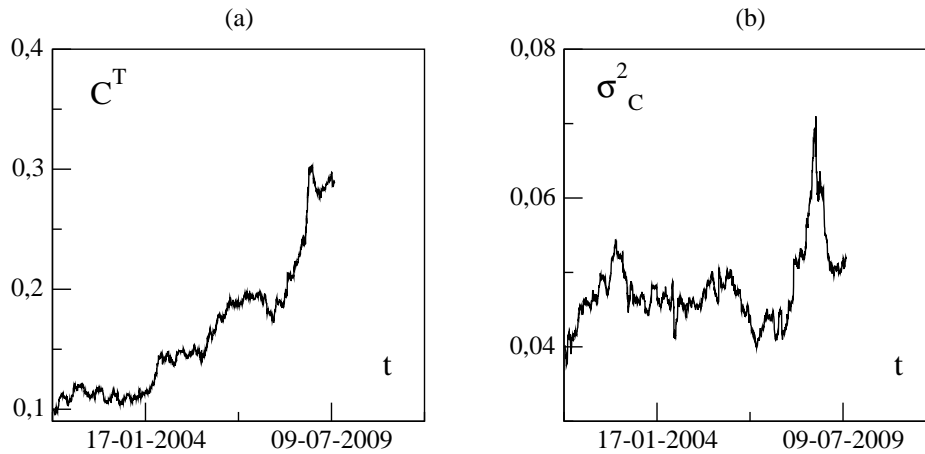


Figure 4: Correlation coefficients in the spatial dimension. Figure (a): Mean of the correlation coefficients; Figure (b): Variance of the correlation coefficients.

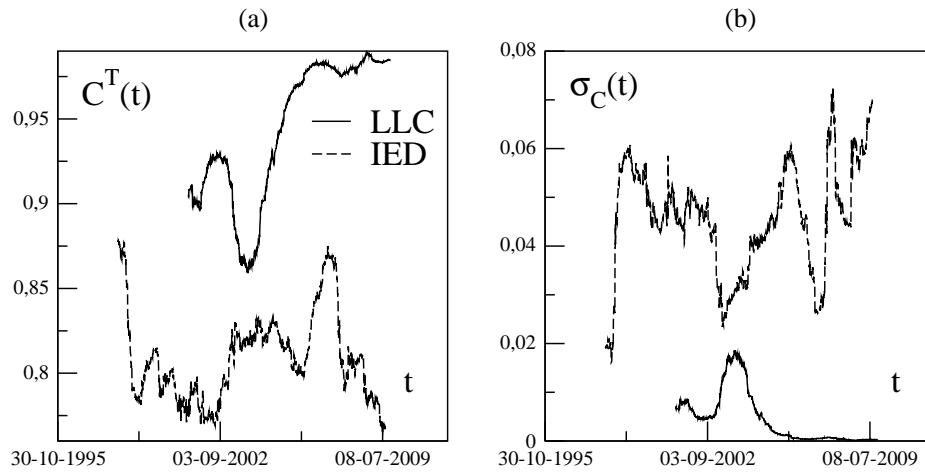


Figure 5: Correlation coefficients in the maturity dimension for the Eurodollar (dashed lines) and the Brent crude oil (black lines). Figure (a): Mean of the correlation coefficients; Figure (b): Variance of the correlation coefficients.

ances for the 14 markets, on the maturity dimension. It confirms that, for almost every contract, the mean correlation is very high and anti correlated with the mean variance. The two natural gases however exhibit more specific figures. Their correlation level is quite low, when compared with other markets, especially for the London Natural Gas. Meanwhile, their mean variance is high.

Table 3: Correlation coefficients in the maturity dimension. Mean correlation coefficients and their variance, correlation between mean and variance, min and max of the correlation coefficients.

Underlying asset	Mean	Variance	Correlations	Min	Max
Light crude oil	0,936	$0,984 \cdot 10^{-3}$	-0,981	0,863	0,973
Brent	0,945	$0,167 \cdot 10^{-2}$	-0,966	0,859	0,99
Heating oil	0,936	$0,266 \cdot 10^{-2}$	-0,978	0,875	0,991
Gasoil	0,95	$0,553 \cdot 10^{-3}$	-0,921	0,891	0,981
Nat. gas (US)	0,617	0,0226	-0,962	0,393	0,855
Nat. gas (Eu)	0,254	0,0275	-0,887	$-0,169 \cdot 10^{-3}$	0,601
Wheat	0,918	$0,232 \cdot 10^{-2}$	-0,943	0,814	0,993
Soy bean	0,912	$0,279 \cdot 10^{-2}$	-0,797	0,769	0,974
Soy oil	0,940	$0,199 \cdot 10^{-2}$	-0,972	0,826	0,997
Corn	0,84	$0,33 \cdot 10^{-2}$	-0,962	0,709	0,902
Eurodollar	0,809	$0,690 \cdot 10^{-2}$	-0,709	0,765	0,878
Gold	0,983	$0,140 \cdot 10^{-3}$	-0,882	0,939	0,996

Merging space and maturity, in three dimensions, we also observe an important rise in the mean correlation and variance, as shown in Figures (6)-a and (6)-b. Moreover, these values are correlated. So our prices system becomes more and more integrated;

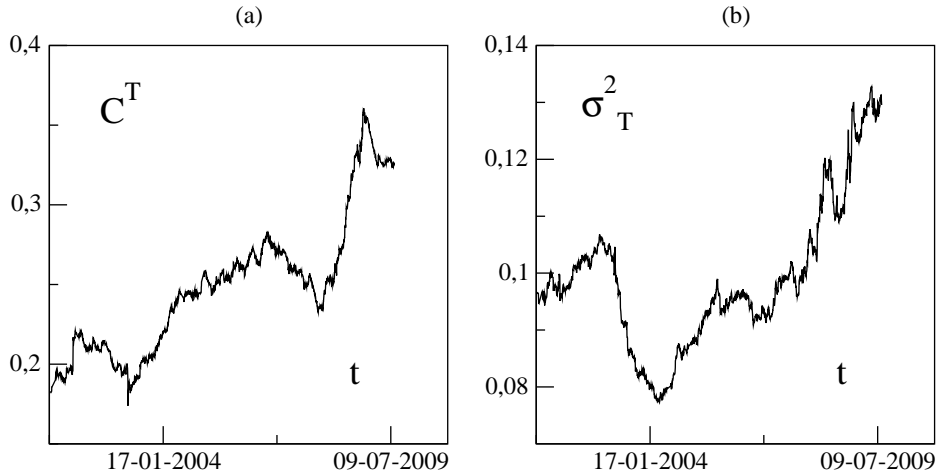


Figure 6: Correlation coefficients in three dimensions. Figure (a): Mean of the correlation coefficients. Figure (b): Variance of the correlation coefficients.

meanwhile, it is less stable.

### 5.2.2 How does markets closeness evolve?

The node's strength, calculated for each node  $i$ , indicates the closeness of one node  $i$  to the others. It is defined as follows:

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

In our case, the node's 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 the others.

Figure (7) represents the time evolution of the node's strength for each node within the fully connected graph, 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 products on the right, the agricultural sector is at the bottom left and financial assets are at the bottom right.

Figure (7) prompts the following remarks: at the end of the period, out of all the assets studied, the two crude oils and American heating oil show the greatest node's strength. These are followed by soy oil, other agricultural assets, the S&P500 contract, gold, the euro dollar exchange rate and European gas oil. The more distant nodes are those representing the eurodollar and natural gases.

When the time evolution of this measure is concerned, the sector shows different patterns: the integration movement, characterized by an increase in the node's strength, emerges earlier for the energy sector than for the agricultural sector. However, it

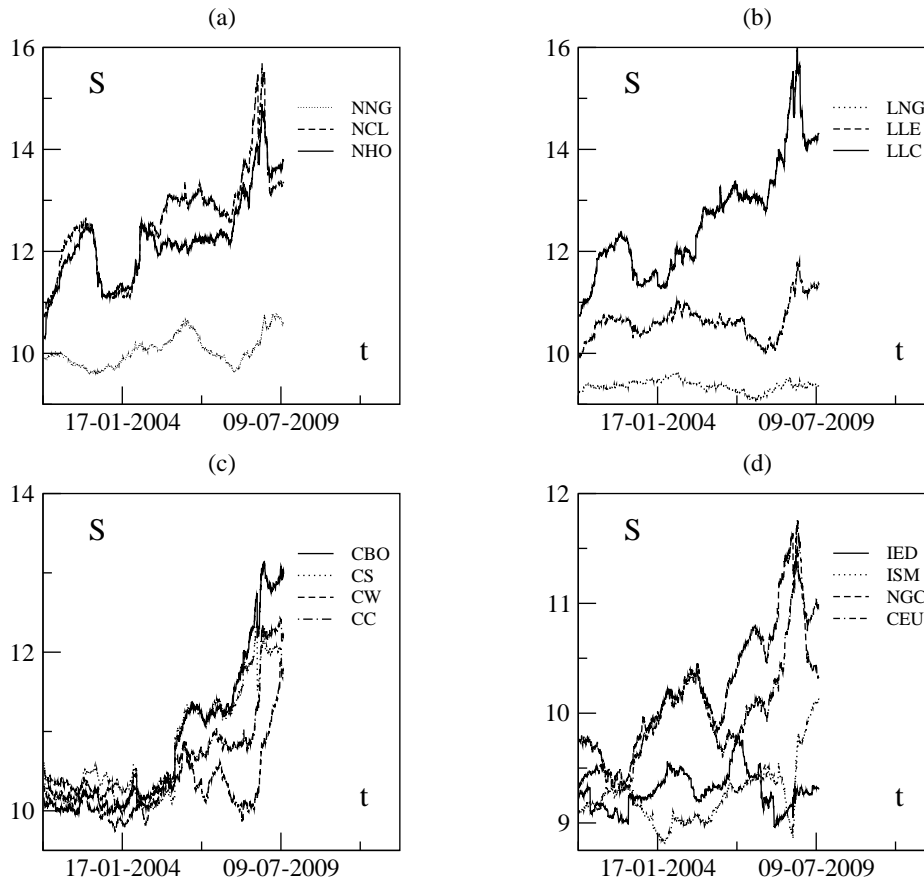


Figure 7: Nodes strength of the markets in the spatial dimension. Figure (a): American energy products. Figure (b): European energy products. Figure (c): Agricultural products. Figure (d): Financial assets.

decreases for energy at the end of the period, which is not the case for agricultural products. Last but not least, most of the products exhibit a strong increase, except for natural gases and interest rate contracts. Thus, whereas the core of the tree becomes more and more integrated, the peripheral assets do not follow this movement.

As far as the maturity dimension is concerned, it was not possible to represent the node's strength for all futures contracts. Therefore, we retained the Brent crude oil and the eurodollar contract examples. We then chose three delivery dates for these



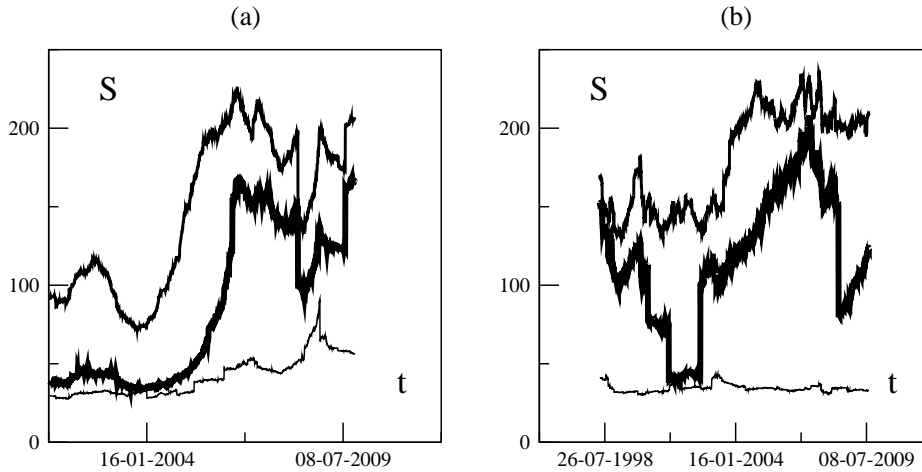


Figure 8: Node strength in the maturity dimension, for three maturities. Figure (a): Eurodollar. Figure (b): Brent .

contracts, as shown in Figures (8)-a and (8)-b. The first maturity is drawn with a fine line, the last maturity with a wide line and the intermediary maturity with a medium width line. All the observed nodes' strength grow over time, except for the first eurodollar maturity. Moreover, in each case, the strongest node is the one which corresponds to the intermediary maturity, whereas the weakest one represents the first maturity.

### 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:

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

where  $t$  denotes the date of the construction of the tree and  $N - 1$  is the number of edges in the MST. The length of a tree is longer as the distances increase, and

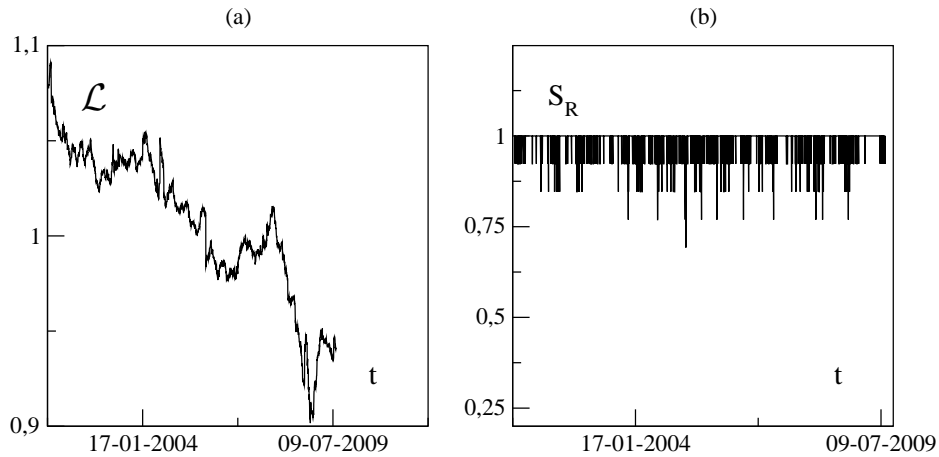


Figure 9: Spatial dimension. Figure (a): Normalized tree's length. Figure (b): Survival ratios.

consequently when correlations are low. Thus, the more the length shortens, the more integrated the system is.

Figure (9)-a represents the dynamic behavior of the normalized length of the MST in its spatial dimension. The general pattern is that the length decreases, which reflects the integration of the system. This information confirms what was observed on the basis of the node's strengths. However we must remember that we are now analyzing a filtered network. Thus, what we see on Figure (9)-a is that the most efficient transmission path for price fluctuations becomes shorter as times goes on. A more indepth examination of the graph 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. This phenomenon is illustrated by Figures (10)-a and -b, which represent the evolutions recorded for the eurodollar contract and for 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 MST 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 (2003a)):

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

In this equation,  $E(t)$  refers to the set of the tree's edges at date  $t$ ,  $\cap$  is the intersection operator, and  $|\cdot|$  gives the number of elements contained in the set. Under normal circumstances, the topology of the trees, between two dates, should be very stable, at least when the window lengths parameter  $\Delta T$  presents small values. While some fluctuations of the survival ratios might be due to real changes in the behavior of the system, it is worth noting that others may simply be due to noise. In this study, we mostly examine the presence of trends in the way these ratios evolve.

Figure (9)-b represents their evolution in the spatial dimension. Most of the time, this measure remains constant, with a value greater than 0.9. Thus, the topology of the tree, in the spatial dimension, is very stable. The shape of the most efficient path for the transmission of prices shocks does not change much over time. However, it is possible to identify four events where 1/4 of the edges has been shuffled. Such a result also calls for further investigation, as a reorganization of the system can be interpreted as the result of a prices shock.

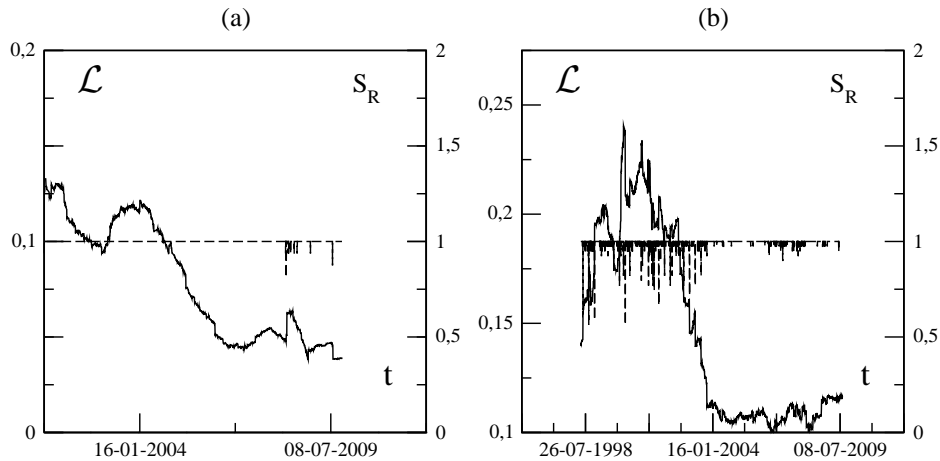


Figure 10: Maturity dimension, normalized tree's length and survival ratios for the Eurodollar (a) and the Brent (b)

In the maturity dimension, Figures (10)-a and -b exhibit different patterns for Brent and interest rates. As far as Brent is concerned, while the trees shrink in the metric sense, the organization of the MST is very stable. Few events seem to destabilize the edges of the trees, except for the very end of the period, *i.e.* from the end of 2008. Again, what happens on the eurodollar is totally different. In mid-2001, around the time 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 ratio 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 than in the spatial and maturity dimension. However, a more specific analysis of these trees, based on a pruning method, provides some interesting results.

### 5.2.5 Interconnections between markets in three dimensions: pruning the trees

As far as the stability of the trees is concerned, especially in 3-D, when focusing on the whole system, it is interesting to distinguish between reorganizations occurring in a specific market, between different delivery dates of the same contract, and reorganization that changes the nature of the links between two markets or even between two sectors. Equation (9) however 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 relationship between two markets or sectors. As we are interested in strong events affecting the markets, inter markets and inter sectors reorganizations seem more relevant. Thus, in order to distinguish between these categories of displacements, we decided to prune the 3-D trees, *i.e.* to only consider the links between markets, whatever the maturity considered. This does not mean, however, that maturity is removed from the analysis. It signifies that with pruned trees, the information on the specific maturity that is responsible for the connexion between markets is no longer identified. Such trees enable us to compute the survival ratios on the sole basis of market links.

Figure (12)-a displays the survival ratio of the reduced trees. As observed previously, the ratio is fairly stable. However, several events cause a significant rearrangement of the tree. This is the case, for example, for two specific dates, namely 02/09/04 and 09/16/08. As illustrated by Figure (11), a focus on these two dates shows that the trees are totally rearranged. In 2004, the MST becomes highly linear, the financial assets sector is at the center of the graph, and commodities appear mainly at the



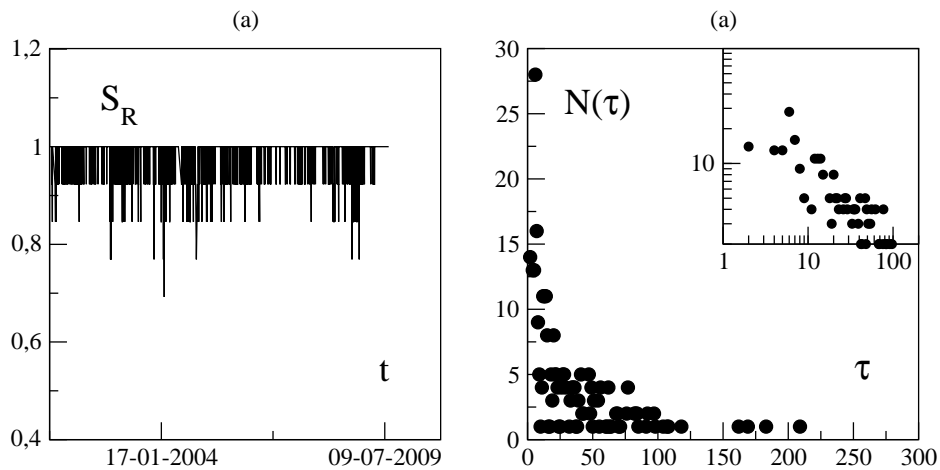


Figure 12: Properties of pruned trees. Figure (a): Survival ratio. Figure (b): Number of occurrences of stable periods of length  $\tau$ . Inset: same as Figure (b), but in log-log scale.

last a long time, and much more stable periods that last a short time. We need to refine the former result, but if such a power law is confirmed, it will mean that the markets can have stable periods of any length.

Finally, another interesting result lies in the analysis of those links which are most frequently responsible for the reorganization of the trees. With fourteen markets, there are ninety one links in our system. Some of them - twenty six - never appear. Among the remaining sixty-five trees, some appear very frequently and, on the contrary, others display very few occurrences. Figure (13) reproduces these two categories of links and the frequency in which they appear in the MST. The most robust links have a frequency equal to one, which means that the links are always 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 euro-dollar exchange rate is also always present. As expected, the relationships between the two crude oils are very stable, with a frequency greater than 09. The same is true for the links between

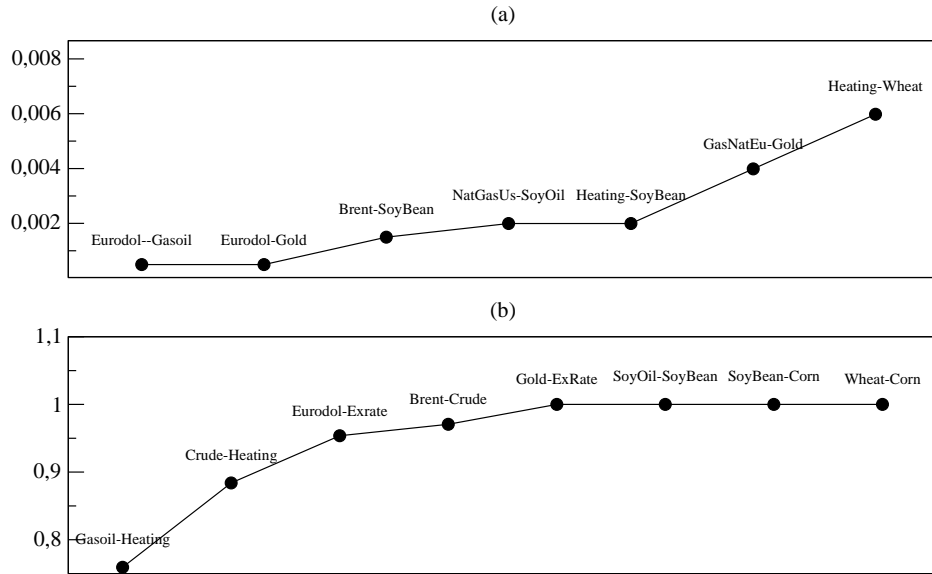


Figure 13: Frequency of links apparition in the pruned minimum spanning trees. Figure (a): frequency lower than 0,008. Figure (b): frequency greater than 0,75.

the interest rates and the exchange rate. This is also intuitive from an economic point of view, as interest rates are embedded in forward exchange rates. The other tail of the curve contains ten links characterized by a frequency lower than 0.01. The lowest values correspond to the association of interest rates and gas oil and that of interest rates and gold.

## 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, as it can be seen as a necessary condition for the propagation of a prices shock. Secondly, based on the fact that previous studies mainly focused solely on the spatio-temporal dimension of integration, we introduce a maturity dimension analysis and perform a three-dimensional analysis.

In the context of an empirical analysis which aims to understand the organization and



the dynamic behavior of a high dimensional price system, the graph-theory has proven very useful. More precisely, Minimum Spanning Trees are particularly interesting in our context, as they are filtered networks enabling us to identify the most probable and the shortest path for the transmission of a prices shock.

The visualization of the MST first shows a star-like organization of the trees in the spatial dimension, whereas the maturity dimension is characterized by chain-like trees. These two topologies merge in the three-dimensional analysis, but the star-like organization still dominates. 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. These intuitive results are very important, as they are a key justification for the use of our methodology.

The American and European crude oils are both found at the center of the graph and ensure the links with agricultural products and financial assets. Thus the first conclusion of importance that we come to is that crude oil is the best candidate for the transmission of prices 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 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 than with products of different natures such as corn bushels and interest rates. The dynamic analysis shows that integration increases significantly on both the spatial and maturity dimensions. Such an increase can be observed on the whole prices system. It is even more evident in the energy sector (with the exception of the American and European natural gas markets)

as well as in the agricultural sector. The latter become highly integrated at the end of our period. Lastly, as far as the financial sector is concerned, no remarkable trend can be highlighted. Thus, as time goes on, the heart of the price system becomes stronger whereas where the peripheral assets are found does not change significantly.

Last but not least, survival ratio also reveal that the system is fairly stable. This is true, except for specific events leading to important reconfigurations of the trees and requiring a specific analysis. We leave these studies for future analyses.

These results have very important consequences, for regulatory as well as for hedging and diversification purposes. The move towards integration started some time ago and there is probably no way to stop or refrain it. However, knowledge of its characteristics is important, as regulation authorities may act in order to prevent prices shocks from occurring, especially in places where their impact may be important. As far as diversification is concerned, portfolio managers should probably focus on the most stable parts of the graph. 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 prices 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.

## References

- Albert, R. and A.-L. Barabási (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74(1), 47–97.
- Albert, R., H. Jeong, and A.-L. Barabási (2000). Error and attack tolerance of complex

- networks. *Nature* 406.
- Banavar (1999). *Nature* 399, 130.
- Bonnano, G. (2004). Networks of equities in financial markets. *Eur. Phys. J. B* 38.
- 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.
- 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.
- Kullmann, L., J. Kertész, and K. Kaski (2002). Time-dependent cross-correlations between different stock returns: A directed network of influence. *Phys. Rev. E* 66(2), 026125.
- 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). *Risk* 16, 11.

- Modigliani, F. and R. Sutch (1966). Innovation in interest rate policy. *American Economic Review* 56, 178–197.
- Onnela, J. (2003a). Dynamic asset trees and black monday. *Physica A* 324.
- Onnela, J.-P. (2003b). 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.
- Sieczka, P. and J. A. Holyst (2009). Correlations in commodity markets. *Physica A* 388.