Energy Economics 34 (2012) 1380–1391

Contents lists available at SciVerse ScienceDirect

Energy Economics

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

# Correlations between biofuels and related commodities before and during the food crisis: A taxonomy perspective

Ladislav Kristoufek [a,b,*], Karel Janda [a,c,d], David Zilberman [e]

[a] Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Opletalova 26, 110 00 Prague, Czech Republic
[b] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 182 08 Prague, Czech Republic
[c] EU Department of Banking and Insurance, Faculty of Finance and Accounting, University of Economics, Prague, Namesti Winstona Churchilla 4, 130 67 Prague, Czech Republic
[d] CERGE-EI, Joint Workplace of Charles University and the Academy of Sciences of the Czech Republic, Politickych veznu 7, 111 21 Prague, Czech Republic
[e] Department of Agricultural and Resource Economics, University of California, Berkeley, 207 Giannini Hall, Berkeley, CA 94720, USA

ARTICLE INFO

*Article history:*
Received 19 December 2011
Received in revised form 4 June 2012
Accepted 5 June 2012
Available online 19 June 2012

*JEL classification:*
C38
Q16
Q42

*Keywords:*
Biofuels
Networks
Minimal spanning tree
Hierarchical tree

ABSTRACT

In this paper, we analyze the relationships between the prices of biodiesel, ethanol and related fuels and agricultural commodities with a use of minimal spanning trees and hierarchical trees. To distinguish between short-term and medium-term effects, we construct these trees for different frequencies (weekly and monthly). We find that in short-term, both ethanol and biodiesel are very weakly connected with the other commodities. In medium-term, the biofuels network becomes more structured. The system splits into two well separated branches — a fuels part and a food part. Biodiesel tends to the fuels branch and ethanol to the food branch. When the periods before and after the food crisis of 2007/2008 are compared, the connections are much stronger for the post-crisis period. This is the first application of this methodology on the biofuel systems.


## 1. Introduction

In this paper, we utilize a straightforward methodology of taxonomy standardly used in networks and complex systems analysis for clear identification of relationships between components of the system. For the first time here, we apply the methodology on the system of biofuels and related agricultural and fuel commodities. We quantify these relationships over different market phases and time dimensions using a graphical display of price transmission network. In this way, we contribute to important policy discussion about impact of biofuels and energy prices on food prices.

Biofuels became of high interest after the oil crisis of the 1970s as a possible replacement for fossil liquid fuels used in transportation. Increased interest in climate and environmental issues in the last three decades also contributed to the popularity of biofuels as alternative fuels. Global production of biofuels experienced a rapid increase since then, especially during the last decade. The main drivers behind this growth are government policies such as mandates, targets and subsidies which have been justified on the grounds of energy security and climate change considerations. However, the concerns raised by the global food crisis in 2007/2008 and ambiguity with respect to environmental impact of biofuels led many government to reconsider their earlier optimism with respect to biofuels.

Very important factor leading to expansion of ethanol was a phase-out of the gasoline additive methyl tertiary butyl ether (MTBE) which was used as an oxygenate to raise the octane number. MTBE was banned or restricted in multiple US states (California, New York, etc.) since it was found to contaminate ground water where it leaked from tanks and pipelines. Unlike other ingredients contained in gasoline fuel, MTBE dissolves in water during the gasoline spills and moves away from spill sites with water flow. MTBE was classified as a possible carcinogen. The fuel industry therefore substituted ethanol as an alternative source of oxygen for fuel blends.

The economics of biofuels constitutes a very active and growing research area as documented in recent review article by Janda et al., forthcoming. Simulation models of economic impacts of biofuels, which are based on long-run parameters (the leading source being GTAP database of Thomas Hertel et al., for recent references see

* Corresponding author at: Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague, Opletalova 26, 110 00 Prague, Czech Republic.
*E-mail addresses:* kristoufek@ies-prague.org (L. Kristoufek), Karel-Janda@seznam.cz (K. Janda), zilber11@berkeley.edu (D. Zilberman).

0140-9883/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.eneco.2012.06.016

Beckman et al. (2011)) and on partial or general equilibrium economic theory, assume links between prices of food, biofuels and fossil fuels. But empirical evidence for these links is largely inconsistent.

Current empirical research on biofuels and fuels price dynamics varies widely from value-at-risk estimation (Chang et al., 2011) to various cointegration estimations (Peri and Baldi, 2010) to volatility spillovers (Serra, 2011) and wavelet coherence analysis (Vacha and Barunik, 2012) and others. The common feature of this research is growing sophistication of econometric estimation which usually comes at the cost of imposing many structural or distributional assumptions on the processes underlying the interactions between the prices of biofuels and related commodities. In this article, we present different methodological approaches to this problem. We analyze connections between biofuels and related commodities (energy-related and food-related) with a use of minimal spanning trees (MST) and hierarchical trees (HT) to uncover the most important connections in the network of commodities.

MST and HT are methodologically very straightforward approaches using only simple correlations as a starting point with no additional prior assumptions. The MST and HT methods are now being increasingly used for analysis of stocks connections (Bonanno et al., 2004; Tumminello et al., 2007), foreign exchange rates (Jang et al., 2011), import/export networks (Kantar et al., 2011), interest rates systems (Tabak et al., 2009), portfolio selection (Onnela et al., 2002) as well as commodities networks (Lucey et al., 2011; Tabak et al., 2010), yet mainly in the journals of interdisciplinary physics, specifically econophysics.

This paper presents the first MST and HT analysis applied on the network containing biofuels. The advantage of our approach is a natural possibility to include simultaneously different biofuels and many different related commodities into our analysis. This contrasts with previous time-series econometric studies which usually focus only on a small selected group of commodities. Our analysis allows the integration of the principal findings in the literature on price transmission between food, fuels and biofuel markets in a clear and elegant way. The correlation clusters formed as results of our analysis may serve as good starting points for further econometric analysis of the price interactions within these clusters. Indeed, the fact that the MST and HT methodology is very straightforward is not only its advantage but also its limitation as well — we are not able to comment on causality between commodities, the methodology does not take into consideration possible cointegration or lagged values of variables of interest. Further, as the methodology is constructed for the stationary series, we might loose information if the analyzed series need to be first-differenced to attain stationarity, which is the case for all stationarity-assuming approaches.

In this paper, we focus on the most popular biofuels — ethanol and biodiesel. Ethanol is mainly produced from crops rich in sugar and starch like sugarcane and corn. Biochemical technologies for conversion of sugar and starch are the most technologically and commercially mature today. Biodiesel is produced from oilseed crops like soybean, rapeseed, and oil palm. Therefore, we are mainly interested whether a dynamic behavior of ethanol and biodiesel forms clusters with food commodities and/or energy commodities. Moreover, we want to analyze the behavior at different frequencies (weekly and monthly) to see whether the relationships apply in short and/or medium term. Further, the connections between the commodities might vary for different phases of the market depending on binding regulatory or technological constraints and market development. To analyze this possibility, we examine the interconnections for two periods separated by the outburst of the food crisis, which was characterized by joint occurrence of high prices of both energy and food commodities (Rajagopal et al., 2009). Indeed, we find that the connections between commodities differ in the pre-crisis and post-crisis periods, which indicates that the relationship between commodities is dependent on the prices of food commodities and is thus non-linear. Importantly, the links between analyzed commodities are stronger in the crisis period. From the biofuels perspective, ethanol tends to be strongly and stably correlated with corn, wheat and soybeans in the food-crisis and post-food-crisis markets with high food prices while biodiesel is more connected to the other fuels (gasoline, diesel and crude oil).

Our empirical results are consistent with the existence of time, space and commodity separate constrained partial equilibria on biofuel markets. We discover different correlation structures among the elements of biofuels system depending on which equilibrium determining constraints are binding on particular market during particular period. In this way, we contribute to reconciliation of seemingly contradictory results of previous studies concerned with price links between biofuels and related commodities.

The rest of the paper is structured as follows. In Section 2, we present a brief review of a current research dealing with links among biofuels and related commodities. In Section 3, we describe the basic notions of the used methodology. In Section 4, the data choice and description is given. Section 5 presents the results of our analysis. Section 6 concludes.

## 2. The relation to current research

Our research is motivated by empirically observed non-linearities in price transmission in biofuel markets, which may be in the most simple form characterized by a following partial equilibrium framework based on Serra et al. (2010). The simulation models of biofuels usually assume an equilibrium determination of biofuel price like the one given by intersection point $E$ of biofuel demand curve $D(P_B,P_G)$ and biofuel supply curve $S(P_B,P_F)$ in Fig. 1, where $P_B,P_F$, and $P_G$ are the prices of relevant biofuel (usually ethanol or biodiesel), its feedstock (corn, soybean, etc.), and an appropriate fossil fuel (gasoline or diesel), respectively. Such an equilibrium price determination implies that the price of biofuel increases with demand curve shifts caused by increase in the price of fossil fuel, eventually reaching new equilibrium level $E_1$ with its associated higher equilibrium price and quantity of a relevant biofuel. Similarly, the supply curve shift caused by increase of feedstock price may lead to a new equilibrium $E_2$ with higher price and lower quantity of a relevant biofuel. This simple unrestricted equilibrium analysis implies that at least in medium- or long-run, after the adjustment to a new equilibrium, the prices as well as price changes of biofuels, fossil fuels and feedstock are strongly positively correlated.

Since the major determining forces of biofuels development are the regulatory support (mandates, blending obligations, subsidies and other measures promoting the use of biofuels) and the technological

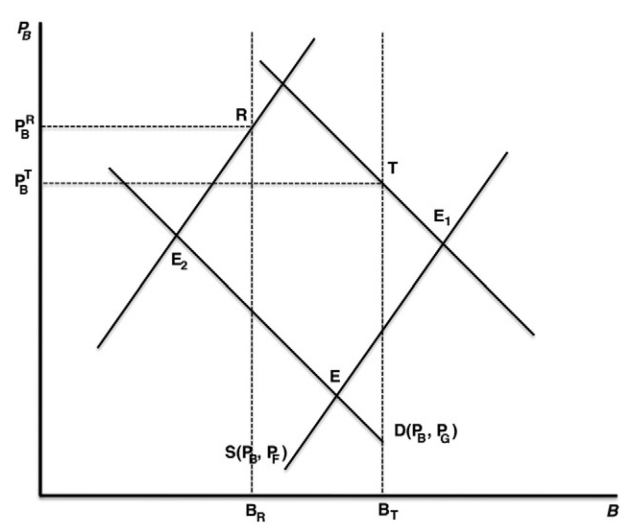


**Fig. 1.** Determination of the price of biofuel.

feasibility (production capacities and technological possibilities of biofuels utilization), the description of supply and demand in Fig. 1 has to include regulation and technological constraints denoted by vertical straight lines through points $B_R$ and $B_T$, respectively. Once we take these constraints, which determine minimal and maximal possible quantities of biofuels on the market, into account, the equilibria denoted as $E_1$ and $E_2$ are no longer feasible. The constrained equilibria will be $T$ or $R$ with associated biofuel prices $P_B^T$ or $P_B^R$, respectively.

The existence of capacity or technological constraints may therefore explain different influences of fossil fuel or feedstock prices on the prices of biofuels, depending on whether any of these two constrains is binding during the analyzed period. The theoretical framework of Fig. 1 assumes that prices of feedstock and fossil fuels are exogenously given while the biofuel prices are determined endogenously by the interaction of supply and demand and by the exogenously given constraints limiting minimal and maximal possible quantities of a given biofuel on the marketplace. This is obviously a very simplifying assumption since there exist feedback effects from prices and quantities of biofuels to prices and quantities of feedstock and maybe even to fossil fuels. Nevertheless, even this most simple framework delivers the message that capacity and technological constraints may prevent high positive correlation among the prices in the wide biofuels related system.

The important feature of biofuel mandates is the determination of quantities, not the determination of prices of biofuels. Therefore it is appropriate to consider market prices of biofuels as function of their quantities produced, which in turn depend not only on mandates but also on prices of feedstock and fossil fuel. While the focus of this paper is on empirical investigation of price-correlation without imposing theory-driven assumptions, a more structured approach may be based on less-simplified competitive market models of biofuel policies and interaction among food and biofuel prices developed by de Gorter and Just (2009a, 2009b); Ciaian and dArtis Kancs (2011a, 2011b), and Drabik (2012) or on models assuming market power with respect to oil markets (Hochman et al., 2010, 2011a, 2011b).

This simple partial equilibrium framework provides theoretical underpinning to our empirical investigation of different correlation patterns of the elements of biofuels system which were reported in the previous studies. In the reminder of this section, we briefly review these most recent time-series studies on links between prices of biofuels and related commodities. More detailed recent reviews are provided by Janda et al. (forthcoming) and Zilberman et al. (forthcoming).

Zhang et al. (2009) focus on volatility of ethanol and commodity prices using cointegration, vector error corrections models (VECM) and multivariate generalized autoregressive conditional heteroskedasticity (mGARCH) models. The authors analyze weekly wholesale price series of the US ethanol, corn, soybean, gasoline and oil from the last week of March 1989 through the first week of December 2007. They find that there are no long-run relations among fuel (ethanol, oil and gasoline) prices and agricultural commodity (corn and soybean) prices in recent years.

The same authors further analyze long- and short-run interactions with a use of cointegration estimation and vector error corrections model with Granger-type causality tests (Zhang et al., 2010). They examine corn, rice, soybeans, sugar, and wheat prices along with prices of energy commodities such as ethanol, gasoline and oil from March 1989 through July 2008. They find no direct long-run price relations between fuel and agricultural commodity prices, and only limited if there are any direct short-run relationships.

Tyner (2010b) finds that since 2006, the ethanol market has established a link between crude oil and corn prices that did not exist historically. He finds that the correlation between crude oil and corn prices was negative (−0.26) from 1988 to 2005; in contrast, it reached a value of 0.80 during the 2006–2008. However, only the price series are analyzed, which raises serious questions about stationarity of the data.

Du et al. (2011) investigate the spillover of crude oil price volatility to agricultural markets (specifically corn and wheat). They apply stochastic volatility models on weekly crude oil, corn and wheat futures prices from November 1998 to January 2009. Their model parameters are estimated using Bayesian Markov Chain Monte Carlo methods. They find that the spillover effects are not statistically significant from zero over the period from November 1998 to October 2006. However, the results indicate significant volatility spillover from the crude oil market to the corn market between October 2006 and January 2009.

In a pair of papers focusing on the cointegration of prices for oil, ethanol and feedstocks, Serra, Zilberman et al. study the US (Serra et al., 2011a, 2011b) and Brazilian (Serra et al., 2011a, 2011b) ethanol markets. In the case of the US, they find the existence of a long-term equilibrium relationship between these prices, with ethanol deviating from this equilibrium in the short term. Further for the US, they find the prices of oil, ethanol and corn to be positively correlated as might be expected. The authors estimate that a 10% perturbation in corn prices boosts ethanol prices by 15%. From the other side, they find that a 10% rise in the price of oil leads to a 10% rise in ethanol. In terms of temporal response time, they find that the response to corn prices is much quicker (1.25 months to full impact) than for an oil price shock (4.25 months). For Brazil, the relevant feedstock is sugarcane. The authors find that sugar and oil prices are exogenously determined and focus their attention on the response of ethanol prices to changes in these two exogenous drivers. The authors conclude that ethanol prices respond relatively quickly to sugar price changes, but more slowly to oil prices. A shift in either of these prices has a very short-run impact on ethanol price volatility as well. These commodity markets are not as quick to achieve long-run equilibrium again as those in the US according to these two studies.

Rajcaniova and Pokrivcak (2011) analyze the relationship between fuel prices (oil, gasoline, and ethanol) and prices of food (corn, wheat, and sugar) serving as ethanol feedstock. They do not find any cointegration in the period January 2005–July 2008, while they find cointegration among majority of their price time series for more recent time period of August 2008–August 2010. Pokrivcak and Rajcaniova (2011) investigate the relationship among the prices of ethanol, gasoline and crude oil in a vector autoregression and impulse–response framework. Their results confirm the usual finding in the literature that the impact of oil price shock on transport fuels is considerable larger than vice versa.

The interaction between monthly prices of crude oil, the US gasoline and the US ethanol between 1994 and 2010 is investigated in a joint structural vector auto regression (SVAR) model by McPhail (2011). His structural VAR model allows to decompose price and quantity data into demand and supply shocks. Since the US ethanol demand is driven mainly by government support through blending mandates and tax credits, he assumes that ethanol demand reflects primarily changes in government policy. As opposed to policy driven demand, ethanol supply shocks are determined by changes in feedstock prices. The author shows that policy-driven ethanol demand expansion leads to statistically significant decrease in real crude oil prices and the US gasoline prices. He also shows that ethanol supply expansion does not have a statistically significant influence on real oil prices.

Ziegelback and Kastner (2011) investigate the relationship between the future prices of European rapeseed and heating oil. They use 2005–2010 daily data to show the asymmetry in price movements. The results of their three-regime threshold cointegration model are similar to the results of Peri and Baldi (2010). Related paper by Busse et al. (2010) deals with the connections between prices of rapeseed oil, soy oil, biodiesel and crude oil during the rapid growth of German biodiesel demand from 2002 until its decline in 2009. They found an evidence for a strong impact of crude oil price on German biodiesel prices, and of biodiesel prices on rapeseed oil prices. However, in both cases, the price adjustment behavior was found to be regime-dependent.

Different results with respect to mutual interactions between the prices of biofuels and related commodities may be due to a number of factors. In our research, we focus on the differences in investment horizon (comparing different frequencies), on the role of technological and regulatory constraints and also on geographic factors of the US and European biofuel markets.

## 3. Methodology

In this section, we describe the basics of construction of minimal spanning trees and hierarchical trees. As this methodology is not well known in the economics literature, we present quite careful description of the methods. For the first application of minimal spanning trees and hierarchical trees to the financial time series and a more detailed description, see Mantegna (1999).

### 3.1. Distance measure

The interconnections in a group of assets are standardly measured by sample correlation coefficients. For a pair of assets $i$ and $j$ with values $X_{it}$ and $X_{jt}$ and $t=1,\dots,T$, the sample correlation coefficient $\hat{\rho_{ij}}$ is calculated as

$$\hat{\rho_{ij}} = \frac{\sum_{t=1}^{T}\left(X_{it}-\overline{X_i}\right)\left(X_{jt}-\overline{X_j}\right)}{\sqrt{\sum_{i=1}^{T}\left(X_{it}-\overline{X_i}\right)^2\sum_{i=1}^{T}\left(X_{jt}-\overline{X_j}\right)^2}}, \quad (1)$$

where $\overline{X_i} = \frac{\sum_{t=1}^{T} X_{it}}{T}$ and $\overline{X_j} = \frac{\sum_{t=1}^{T} X_{jt}}{T}$ are respective time series averages. Linear correlation $\rho_{ij}$ ranges between $-1$ (perfectly anti-correlated) and 1 (perfectly correlated) with $\rho_{ij}=0$ meaning that the pair is uncorrelated. Note that it only makes sense to estimate correlations for the series with well defined means and variances, i.e. weak stationarity of the series is needed.

For a portfolio of $N$ assets, we obtain $N(N-1)/2$ pairs of correlations. Mantegna (1999) showed that the correlation coefficients can be transformed into distance measures, which can in turn be used to describe hierarchical organization of the group of analyzed assets. Distance measure

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

is constructed so that it fulfills three axioms of a metric distance:

- $d_{ij}=0Z$ if and only if $i=j$;
- $d_{ij}=d_{ji}Z$;
- $d_{ij}\leq d_{ik}+d_{kj}$ for all $k$

From the definition of the correlation coefficient, the distance ranges between 0and 2, while $d_{ij}\rightarrow 0$ means that the pair is strongly correlated, $d_{ij}\rightarrow 2$ implies strongly anti-correlated pair and $d_{ij}=\sqrt{2}$ characterizes an uncorrelated pair.

### 3.2. Minimal spanning tree and hierarchical tree

Minimal spanning tree (MST) is used to extract the most important connections in the whole network. For our purposes, the connections are characterized by correlation coefficients between pairs of assets. The basic idea behind MST is to reduce the number of $N(N-1)/2$ pairs to only the $N-1$ most important connections while the whole system remains connected. The procedure is very straightforward and described in detail in Mantegna (1999). In short, we transform the correlation matrix $\mathbb{C}$ into a distance matrix $\mathbb{D}$, discarding the diagonal elements (containing zero distances). We then find the closest pair of assets, which creates the first two nodes in the network connected by the first link (with a weight equal to the distance $d_{ij}$). Each node now has a single edge (the link connected to the node). We proceed to the second closest pair which creates the second pair of nodes. At this point, if a node from the second pair is already present in the network, the new node is simply connected to the existing pair. The steps are repeated until $N-1$ links are reached, while the network must not be closed or create closed loops. If the link would create a loop, it is not added into the network. We use Kruskal's algorithm in our application (Kruskal, 1956).

MST helps us to construct hierarchical trees (HT) which are important for the analysis of clusters. With a use of HT, it has been shown that stocks form clusters based on the industrial branches (Mantegna, 1999; Tabak et al., 2010) and that foreign exchange rates create clusters with respect to the geographical location (Jang et al., 2011; Keskin et al., 2011; Mizuno et al., 2006). In order to construct HT with a use of MST and $\mathbb{D}$, we first need to determine the subdominant ultra metric distance matrix $\mathbb{D}^*$. The elements of the matrix $\mathbb{D}^*$ are defined as the subdominant ultra metric distances $d^*_{ij}$. Such a distance is equal to the maximal weight of the link which needs to be taken to move from node $i$ to node $j$ in the MST. More formally, $d^*_{ij}=max(d_{kl})$, where $k$ and $l$ stand for all nodes connecting $i$ and $j$ (including $i$ and $j$) in the corresponding MST. In matrix $\mathbb{D}^*$, we find the minimal distance $d^*_{ij}$ and create the first pair of assets. We follow in connecting the assets and if we find more assets with same $d^*_{ij}$, we connect the clusters together. In the end, we obtain the whole HT which clearly separates clusters of the analyzed variables (Mantegna, 1999). For illustration, consider three commodities $a$, $b$ and $c$, which form MST such that $a$–$b$–$c$ with $d_{ab}=0.4$ and $d_{bc}=0.7$. Since the lowest distance is $d_{ab}$, then the ultra metric distance is $d^*_{ab}=0.4$. The second lowest distance is $d_{bc}$ which implies $d^*_{bc}=0.7$. Now, we need to find $d^*_{ac}$. To get from $c$ to $a$ in this simple MST, we need to cross $b$. $d^*_{ac}$ is then a maximum of distances between $a$–$b$ and $b$–$c$, i.e. $d^*_{ac}=max(d_{ab},d_{bc})$. We arrive at $d^*_{ab}=0.4$ and $d^*_{ac}=d^*_{bc}=0.7$, which means that $a$ and $b$ are connected and form a pair while $c$ is separated from this simple cluster as it has the same ultra metric distance from both $a$ and $b$, and we are able to construct the hierarchical tree. The procedure will be better illustrated on the analyzed dataset arriving at more complicated hierarchical structures in the following sections.

Depending on the structure of HT, we can discuss interconnections between specific clusters or separate assets and commodities. In general, HT translates relatively unstructured MST and creates a unique hierarchical structure. From the point of view of our research and focus on clusters in biofuels and related commodities, HT gives a more informative picture of existing clusters. Without HT, MST would give only limited information.

### 3.3. Stability of links

The major weakness of the described methodology lies in the fact that the calculated MST and HT might be unstable. Moreover, without further statistical analysis, we cannot be sure whether the links present in the MST are actually the important links in the network or are rather a statistical anomaly, i.e. whether the results are sensitive to the sampling. To deal with the problem, we use a bootstrapping technique proposed by Tumminello et al. (2007) specifically for MST and HT analysis.

In the procedure, we first construct the original MST and HT. Then, we construct a bootstrapped time series from the original while keeping the time series length fixed (i.e. the observations may repeat in the bootstrapped sample). MST and HT are then constructed for the bootstrapped time series and links are recorded. It is then checked whether the connections in the original MST are also present in the new MST based on bootstrapped time series. We repeat such procedure 1000 times so that we can distinguish whether the connections in the original MST and HT are the strong ones or statistical anomalies (Keskin et al., 2011). The share of the bootstrapped cases where the link appears between nodes $i$ and $j$ will be labeled as $b_{ij}$ with an obvious range of $0 \le b_{ij} \le 1$.

## 4. Data

We analyze weekly and monthly prices of Brent crude oil (CO), ethanol (E), corn (C), wheat (W), sugar cane (SC), soybeans (S), sugar beets (SB), consumer biodiesel (BD), German diesel and gasoline (GD and GG), and the US diesel and gasoline (UD and UG) from 24.11.2003 to 28.2.2011. The Bloomberg tickers and contracts specification are summarized in Table 1. Gasoline and diesel prices were obtained from the US Energy Information Administration and they present average prices of the countries. We use both the US and the German prices to uncover potential connection to ethanol and biodiesel as biodiesel production used to be rather a European activity while ethanol production is more an American activity. Ethanol price is the New York Harbor price for ethanol according to ASTM D4806 specification. This is a denaturated anhydrous fuel ethanol for blending with gasoline. Crude oil price refers to current pipeline export quality Brent blend as supplied at Sullom Voe. Corn price is for Corn No. 2 Yellow. Wheat price is for various types of wheat (No. 2 Soft Red Winter Wheat, No. 2 Hard Red Winter Wheat, No. 2 Dark Northern Spring Wheat, and No. 2 Northern Spring Wheat at par (contract price); and No. 1 Soft Red Winter Wheat, No. 1 Hard Red Winter Wheat, No. 1 Dark Northern Spring Wheat and No. 1 Northern Spring Wheat at 3 cents per bushel over contract price.) Sugar price is for raw centrifugal cane sugar based on 96° average polarization. Soybeans price is for Soybeans No. 2 Yellow. Sugar beets price is for white beet or cane crystal sugar or any other refined sugar. Biodiesel price is for commodity type consumer biodiesel, as reported by F.O. Licht. Daily data are not used in our analysis as the spot markets (ethanol and biodiesel) are not liquid enough and the analysis would not be meaningful.

The weekly (Monday) logarithmic prices are shown in Fig. 2 (all prices are in the US dollars) and basic descriptive statistics for logarithmic returns are presented in Table 2. In the Fig. 2a–d, we separate the series according to their co-movements. Retail fuels dynamics practically overlaps for the whole examined period, both analyzed sugar commodities follow very similar price-path. Similar yet weaker connections are visible for the triple of corn, wheat and soybeans. Both analyzed biofuels follow very different path. Period of the food crisis is clearly visible in Fig. 2c, where we observe that prices of corn, wheat and soybeans rapidly increased between 2007 and 2008 but also remained higher afterwards compared to the period before 2007. The period of high fodder prices is accompanied by high prices of crude oil and retail fuels between 2007 and 2008. However, there is no such obvious change in prices of sugar commodities.

Taking $X_t$ as a Monday closing prices, we analyze returns $r_t = log(X_t - X_{t-1})$. From Table 2, we can see that all the retail fuels are positively skewed while all the others are negatively skewed (except for positively skewed wheat and symmetric ethanol). Returns of the US fuels are more leptokurtic than the German fuels but are also less volatile (approximately twice). Note that leptokurtosis (the fat tails) can cause problems to simple linear correlations since the correlation of extreme events, i.e. the correlation in the tails of a multivariate distribution, can differ from the correlation of observations near to the mean of the distributions (Patton, 2006). However, checking dependence in tails for weekly and monthly data would be difficult and statistically implausible so that we leave this possible complication aside.

**Table 1**
Analyzed Bloomberg commodities.

| Commodity | Ticker | Contract type |
|---|---|---|
| Crude oil | CO 1 Comdty | 1st month futures, ICE |
| Ethanol | ETHNNYPR Index | Spot, FOB |
| Corn | C 1 Comdty | 1st month futures, CBOT |
| Wheat | W 1 Comdty | 1st month futures, CBOT |
| Sugar cane | SB 1 Comdty | 1st month futures, ICE |
| Soybeans | S 1 Comdty | 1st month futures, CBOT |
| Sugar beets | QW 1 Comdty | 1st month futures, LIFFE |
| Biodiesel | BIOCEUGE Index | Spot, Germany |

As we analyze the structure of distances, which are simply transformed correlations, between the commodities, stationarity of the series becomes crucial. Table 5 summarizes the results for three stationarity tests — ADF test with a constant, ADF test without a constant and KPSS test. The results are quite straightforward — all the logarithmic returns are stationary, which implies that we can proceed to the estimation of correlation coefficients and distances from the logarithmic returns series without further adjustments. Note that we try to keep the methodology as straightforward as possible. To do so, we present only the results for unadjusted logarithmic returns, which are standardly done in the literature. We also applied the methodology on AR (1)–GARCH (1,1)-filtered series, i.e. the estimated correlations were robust to autocorrelation and heteroskedasticity in the processes. However, the sample correlations differ only a little for the adjusted series and the resulting MSTs and HTs are qualitatively the same as the ones presented in this paper. Again, the methodology can be extended to various frameworks modeling time-dependent correlations (Long et al., 2011) or even time- and frequency-dependent correlations (Vacha and Barunik, 2012).

Apart from analyzing the network between commodities for the whole sample period, we also examine connections between series for different phases of the market. We use a coincidental fact that the central point in our sample is connected to the beginning of the food crisis of years 2007 and 2008. Therefore, we split the sample into two – the “pre-crisis” (24.11.2003–9.7.2007) and the “post-crisis” (16.7.2007–28.2.2011) periods – each with 190 weekly observations. The basic descriptive statistics of these two sub-samples are summarized in Table 3. When the two periods are compared, we find some interesting differences — all food commodities have higher volatility during the food crisis, the average growth rate is higher for the food commodities in the crisis period, compared to the lower growth of crude oil and retail fuels. However, these differences are not strong enough to break the stationarity of the series as documented by the stationarity tests for the whole sample.

## 5. Results

In this section, we present and comment on the results of the minimal spanning trees and hierarchical trees for the studied network of commodities. We first focus on the connections found for the whole sample and then focus on two sub periods — before the food crisis and after the food crisis (inclusive).[1]

### 5.1. Whole sample

We start with the first few steps of construction of minimal spanning tree for weekly returns to illustrate the procedure. The pair with the highest correlation coefficient – and thus the closest one – consists of German diesel and German gasoline with $d_{ij} = 0.5330$. Therefore, the first connected nodes of the MST are GD–GG. The second lowest distance is the one between US gasoline and US diesel ($d_{ij} = 0.6563$). We now have two pairs of nodes GD–GG and UD–UG in the MST. The next lowest distance is found for SB–SC pair ($d_{ij} = 0.7671$). The MST now contains three separate pairs of nodes — GD–GG, UD–UG and SB–SC. We proceed to the fourth lowest distance and obtain a next pair created by corn and wheat ($d_{ij} = 0.8848$). Again, neither corn nor wheat is connected to the other nodes already present in the MST which implies that the MST is now made of four separate pairs. In the

[1] All calculations and construction of MST and HT have been conducted and coded in TSP 5.0.

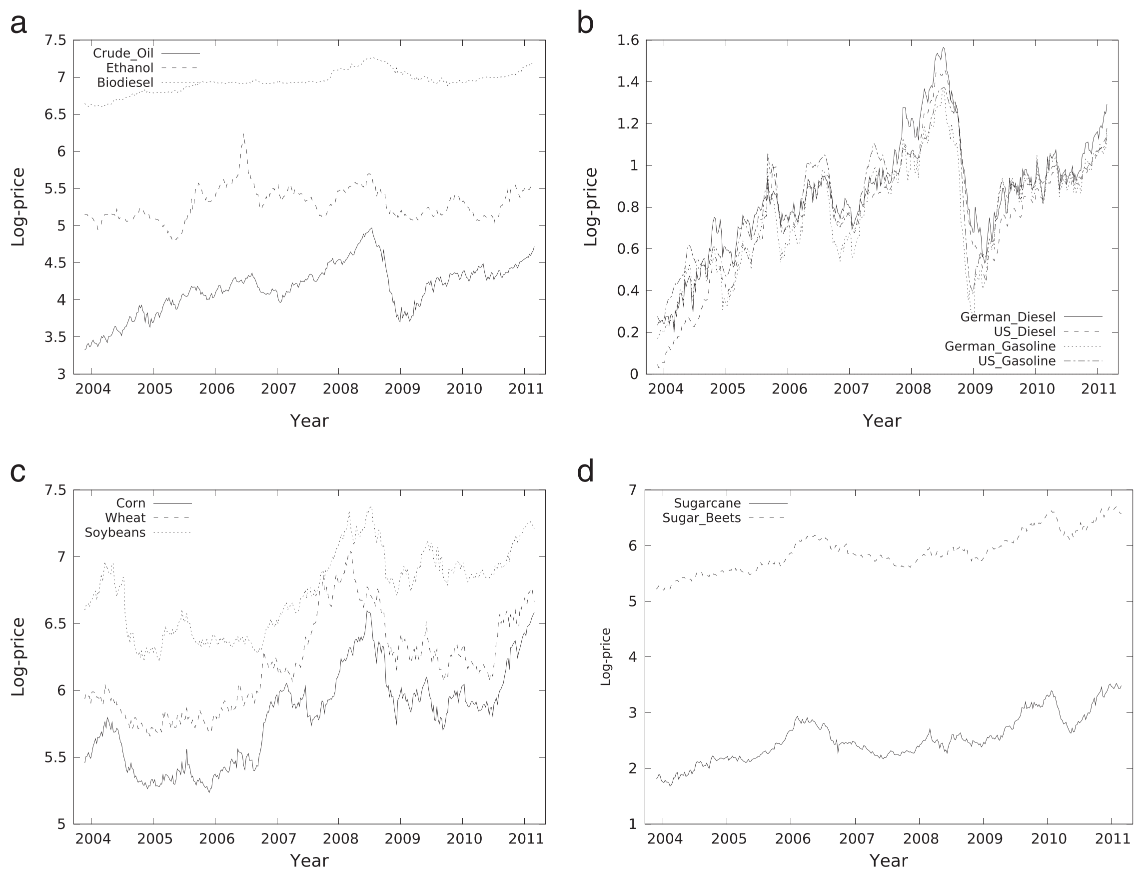


**Fig. 2.** Weekly logarithmic prices of the analyzed commodities.

next step, we find that the fifth lowest distance in the distance matrix $\mathbb{D}$ is for the German and US gasoline ($d_{ij}=0.9181$). Both of the nodes are already present in the MST so that we just connect the nodes GG and UG. The MST is now created by two pairs C–W, SB–SC and one quadruple GD–GG–UG–UD. Next pair is formed by soybeans and corn with $d_{ij}=0.9369$. Corn is already a part of the MST so that soybeans are just connected to the existing couple C–W. The MST is now formed by a pair SB–SC, a triple C–W–S and a quadruple GD–GG–UG–UD. The next closest pair is the one of German gasoline and US diesel. Both nodes are already present in the MST. Moreover, they are both a part of the quadruple GD–GG–UG–UD and are therefore already connected. If we added a new link GG–UD, we would create a loop, which is not desirable. Eventually, no new link is added for this pair. Following these simple rules, we arrive at the final MST presented in Fig. 3a.

In the similar way, we describe the construction of the hierarchical tree for the weekly returns. We start with finding the closest pair in the MST — that is GG–GD pair, which in turn forms the first pair in the HT. Next is the UG–UD pair, which again forms a pair in the HT. In the same way, the C–W and SC–SB pairs are formed. The next lowest distance is between GG–UG link. Now, both nodes are already present in the HT so that we connect the pairs GG–GD and UG–UD but assign

**Table 2**
Descriptive statistics (2003–2011).

| | Mean | Min | Max | SD | Skewness | Ex. kurtosis |
|---|---|---|---|---|---|---|
| Crude oil | 0.0035 | −0.1595 | 0.2019 | 0.0477 | −0.1705 | 1.6535 |
| Ethanol | 0.0011 | −0.2097 | 0.2085 | 0.0497 | 0.01177 | 2.9204 |
| Corn | 0.0029 | −0.1905 | 0.1774 | 0.0468 | −0.1059 | 1.4742 |
| Wheat | 0.0019 | −0.1204 | 0.1621 | 0.0469 | 0.2776 | 0.1755 |
| Sugar cane | 0.0044 | −0.2372 | 0.1493 | 0.0513 | −0.2966 | 1.5316 |
| Soybeans | 0.0016 | −0.2810 | 0.1390 | 0.0452 | −1.0082 | 4.7419 |
| Sugar beets | 0.0036 | −0.1713 | 0.1135 | 0.0408 | −0.4910 | 1.2595 |
| Biodiesel | 0.0014 | −0.2532 | 0.2279 | 0.0283 | −0.2581 | 44.7410 |
| German diesel | 0.0028 | −0.1206 | 0.1420 | 0.0437 | 0.2026 | 0.3521 |
| US diesel | 0.0030 | −0.1099 | 0.1379 | 0.0260 | 0.3980 | 4.4957 |
| German gasoline | 0.0026 | −0.1806 | 0.2256 | 0.0507 | 0.3609 | 1.9525 |
| US gasoline | 0.0024 | −0.1049 | 0.1831 | 0.0276 | 0.1612 | 6.1811 |

**Table 3**
Descriptive statistics — pre- and post-crisis periods.

| | Mean | Min | Max | SD | Skewness | Ex. kurtosis |
|---|---|---|---|---|---|---|
| *Pre-crisis* | | | | | | |
| Crude oil | 0.0053 | −0.1431 | 0.1150 | 0.0411 | −0.3442 | 0.5150 |
| Ethanol | 0.0012 | −0.2097 | 0.2085 | 0.0551 | 0.0800 | 2.8789 |
| Corn | 0.0018 | −0.1509 | 0.1368 | 0.0416 | −0.0616 | 1.2964 |
| Wheat | 0.0023 | −0.0782 | 0.1305 | 0.0399 | 0.6645 | 0.1841 |
| Sugar cane | 0.0024 | −0.2372 | 0.1493 | 0.0468 | −0.5419 | 3.9253 |
| Soybeans | 0.0009 | −0.2810 | 0.1080 | 0.0435 | −1.4926 | 8.5903 |
| Sugar beets | 0.0030 | −0.1713 | 0.0884 | 0.0357 | −0.8138 | 2.9147 |
| Biodiesel | 0.0016 | −0.0369 | 0.0388 | 0.0121 | 0.6104 | 1.1479 |
| German diesel | 0.0036 | −0.1206 | 0.1377 | 0.0423 | 0.2037 | 1.0024 |
| US diesel | 0.0044 | −0.1099 | 0.1379 | 0.0273 | 0.6974 | 6.5131 |
| German gasoline | 0.0040 | −0.1487 | 0.2256 | 0.0482 | 0.4812 | 2.7339 |
| US gasoline | 0.0040 | −0.0601 | 0.1831 | 0.0276 | 1.3953 | 8.4637 |
| *Post-crisis* | | | | | | |
| Crude oil | 0.0020 | −0.1595 | 0.2019 | 0.0536 | −0.0492 | 1.6757 |
| Ethanol | 0.0011 | −0.1610 | 0.1484 | 0.0438 | −0.1267 | 2.0017 |
| Corn | 0.0041 | −0.1905 | 0.1774 | 0.0515 | −0.2140 | 1.3049 |
| Wheat | 0.0015 | −0.1204 | 0.1621 | 0.0530 | 0.1125 | −0.1299 |
| Sugar cane | 0.0064 | −0.1650 | 0.1455 | 0.0563 | −0.1948 | 0.2212 |
| Soybeans | 0.0023 | −0.1711 | 0.1390 | 0.0467 | −0.6265 | 1.7960 |
| Sugar beets | 0.0042 | −0.1344 | 0.1135 | 0.0455 | −0.3361 | 0.3190 |
| Biodiesel | 0.0011 | −0.2532 | 0.2279 | 0.0381 | −0.2135 | 25.9360 |
| German diesel | 0.0020 | −0.0916 | 0.1420 | 0.0452 | 0.2079 | −0.1679 |
| US diesel | 0.0016 | −0.0735 | 0.0775 | 0.0246 | −0.0439 | 1.1339 |
| German gasoline | 0.0012 | −0.1806 | 0.2139 | 0.0533 | 0.2848 | 1.3473 |
| US gasoline | 0.0007 | −0.1049 | 0.0800 | 0.0276 | −1.0606 | 3.4271 |

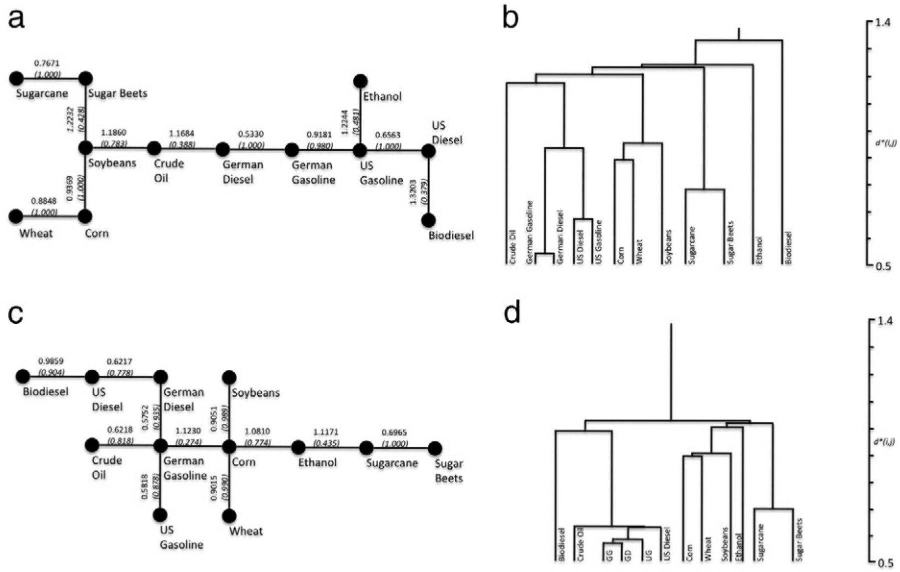


**Fig. 3.** Minimal spanning trees (first column) and hierarchical trees (second column) for network of returns and different frequencies (from the top — one week and one month).

the distance $d^*_{ij} = 0.9181$ to all pairs which might be formed by these four nodes. Therefore, the distance between the pairs is now 0.9181. This is graphically shown in Fig. 3b. The next lowest distance in the MST is present for C–S pair. Corn is already a part of the HT and forms a pair with wheat. We now check what the maximum distance between soybeans and wheat is and we find that it is the distance between corn and soybeans. In turn, we assign $d^*_{ij} = 0.9369$ to both possible pairs formed from the three. Graphically, we connect S to the pair C–W. Again, if we follow these simple rules, we finally arrive at the HT presented in Fig. 3b. In the same way, we constructed the HT for monthly frequency.

Let us first focus on the minimal spanning trees for a higher frequency — a trading week. It is clearly visible that the minimal spanning tree is formed from two parts — a food part (SC, SB, W, C, and S) and a fuels part (CO, GD, GG, UG, UD, E, and BD). In the MST charts, we also show the distances $d_{ij}$ between nodes (regular font) as well as a bootstrapped value $b_{ij}$ (italics in brackets). The bootstrapped value represents the proportion of times when the specific link has been present in the bootstrapped MST. For example, the value of 0.783 for S–CO link means that out of 1000 bootstrapped realization, the S–CO link has been found in 783 final MSTs. Using these values, we can comment on a strength or a stability of a link in the MST. In the food part of the MST, we observe a triple W–C–S and a pair SC–SB which have been found in all bootstrapped realizations. These links are thus very stable. The connection between the triple and the pair is quite weaker ($b_{ij} = 0.428$). We can see similarly strong connections in the fuels part of the MST, mainly for a foursome GD–GG–UG–UD which has been found in almost all the bootstrapped cases. Both biofuels are linked to the US fuels. Relatively low bootstrapped value for CO–GD link ($b_{ij} = 0.388$) is caused mainly by the fact that crude oil is correlated to GG, GD, UD and UG at similar levels so that the links alter between the four in the bootstrapped cases.

Very similar results can be read from the HT. Here, we can see that there are several clusters — a fuels cluster, a sugar cluster and a fodder cluster. The other commodities – crude oil, ethanol and biodiesel – are quite far from these clusters and thus do not interact much in the short term. Importantly, the biofuels are quite remote from the rest of the network, which can be interpreted in a way that in a short term horizon, the behavior of these biofuels is not dependent on the other analyzed commodities.

When we look at the relationships between commodities at the lower (monthly) frequency, both MST and HT are getting more structured. The core of the connections remains the same — we still have the three clusters. However, the behavior of the biofuels changes. Ethanol becomes more connected with the food part and biodiesel with the fuels part. Interestingly, the whole network practically splits into two branches — one branch contains all the retail fuels, crude oil and biodiesel and the other branch includes all the analyzed food and ethanol. However, it has to be noted that a distance between the branches is quite low so that the whole system is well correlated. Moreover, difference in the distances between ethanol and C–W–S cluster, then SC–SB from C–W–S–E cluster and then between the whole food cluster and the fuels cluster is very small (all three ultra metric distances are between 1.08 and 1.12), which means that this separation is very unstable. Nevertheless, the average distance between the analyzed commodities decreases from 0.98 for the weekly frequency to 0.84 for the monthly frequency (Table 4), which implies that the system gets more interconnected with the lower frequency. Apart from the connections of the biofuels to the rest of the network, we observe some other interesting features. First, compared to the weekly frequency, where the GG–GD and UG–UD clusters were well separated, this separation almost disappears for the monthly frequency. This implies that in a short term, behavior of the retail fuels is dominated by geographical features but in medium term, this separation vanishes. Second, crude oil is very well connected to the retail fuels cluster in the medium term, which was not the case for the short term. This implies that it takes several weeks until the effect of the price change of crude oil is reflected in the prices of retail fuels. And last, the feedstock and sugar clusters are well separated for both frequencies.

To summarize the most important findings for ethanol and biodiesel returns with respect to different frequencies, we can say that in the short term, both of these are very weakly connected with the other commodities. Moreover, there is no clear inclination to either of fuels or food parts of the network. In the medium term, biodiesel becomes connected to the fuels section of the system, whereas ethanol gets more connected to the food branch of the system.

**Table 4**
Average tree lengths and average bootstrapped values.

| Period | Frequency | $\overline{d_{ij}}$ | $\widehat{\sigma(d_{ij})}$ | $\overline{b_{ij}}$ | $\widehat{\sigma(b_{ij})}$ |
|---|---|---|---|---|---|
| Whole period | Week | 0.9835 | 0.2481 | 0.7672 | 0.2710 |
| | Month | 0.8373 | 0.2125 | 0.7977 | 0.2253 |
| Pre-crisis | Weekly | 1.0020 | 0.2446 | 0.8202 | 0.2180 |
| | Monthly | 0.9478 | 0.3078 | 0.7046 | 0.2675 |
| Post-crisis | Weekly | 0.8768 | 0.1810 | 0.8219 | 0.2471 |
| | Monthly | 0.7723 | 0.2341 | 0.7661 | 0.2190 |

Unfortunately, the MST and HT analysis is not capable to find the direction of the effects, i.e. whether the effect comes from food to ethanol or the other way around. However our supplementary follow-up analysis of Granger-causality based on the whole sample of data used in this paper shows that prices of corn Granger-cause prices of ethanol in both short and medium term. We found out that this effect is positive, so that increase in price of corn leads to increase in price of ethanol in relatively short time and the effect disappears quite quickly since the aggregate effect is insignificant starting by the 12th week. We did not find statistically significant Granger causality in the other direction (from ethanol to corn). This is in agreement with the findings of Wixson and Katchova (2012) who show on monthly US data from 1995 to 2010 that price of corn Granger-causes price of ethanol and that ethanol does not Granger-causes wheat. Similar results are reported by Saghaian (2010) who shows that corn price Granger-causes price of ethanol with statistical significance on all conventional levels, but the reversed direction of Granger causality is statistically significant only on 10% significance level that disappers quite quickly since the aggregate effect is insignificant starting by the 12th week. We did not find statistically significant Granger causality in the other direction (from ethanol to corn). This is in agreement with the findings of Wixson and Katchova (2012) who show on monthly US data from 1995 to 2010 that price of corn Granger-causes price of ethanol and that ethanol does not Granger-cause wheat. Similar results are reported by Saghaian (2010), who shows that corn price Granger-causes price of ethanol with statistical significance on all conventional levels, but the reversed direction of Granger causality is statistically significant only on 10% significance level. However there also exist studies indicating different causality patterns. For example Zhang et al. (2009) did not find any long-run causality relation between prices of ethanol and corn while in the short-run they found out that prices of ethanol Granger-cause the price of corn. Serra et al. (2011b) show that positive causal relationship from ethanol prices to corn prices does not only prevail in the short-run but also in the longer term. However they also show that a shock to corn price when the ethanol price is far away from its equilibrium level will cause an adjustment in the ethanol price in the same direction.

### 5.2. Food crisis

Here, we present the results for the pre-crisis and the post-crisis periods. Note that the pre-crisis period ranges from 24.11.2003 to 9.7.2007 and the post-crisis period from 16.7.2007 to 28.2.2011.

#### 5.2.1. Pre-crisis period

In the pre-crisis period, we again find strongly linked and stable pairs of GD–GG, UD–UG, SB–SC and a triple S–C–W (Fig. 4a). The connection between the geographically separated retail fuels is also quite strong and very stable (with $b_{ij}=0.987$). Both biofuels are connected to the retail fuels — ethanol to UG (with a distance $d_{ij}=1.1759$ and $b_{ij}=0.762$) and biodiesel to UD (with a distance $d_{ij}=1.2293$ and $b_{ij}=0.731$). Compared to the connections for the whole period, these two are stronger and mainly more stable links. Other than that, the rest of the network is quite separated, which is supported by the hierarchical tree (Fig. 4b). In the figure, we see that the system is separated into three branches — a feedstock branch, a sugars branch and a fuels branch. Biofuels are connected to the fuels branch. This implies that during the low food prices period, the biofuels were correlated with other fuels rather than with their production factors in the short term. Further, compared to the whole period, the links are stronger and more stable.

When the frequency is lowered to one month, we observe a change in a structure of the network (Fig. 4c). Even though the standard clusters of the sugars, feedstock and retail fuels remain, the connections are less stable. What remains is the connection of biodiesel to the US diesel, which is stronger than for the weekly frequency ($d_{ij}=1.1114$) and also quite stable ($b_{ij}=0.781$). On the other hand, ethanol becomes less connected to the network and its connection to German gasoline and sugarcane are quite unstable and weaker than for the weekly frequency. The hierarchical tree again tells a more detailed story (Fig. 4d). There are again the same three separate branches of the network as for the weekly frequency — the fuel, feedstock and sugar branches. However, the separation between these is not as clear as in the previous case. Similarly to the weekly frequency, the S–W–C is quite strongly separated from the rest of the system. This implies that the feedstock commodities did not interact (or only weakly) with the rest of the system before the food crisis, i.e. when the prices of corn, wheat and soybeans were low, the correlations were low. Further in the text, we show that the opposite is true for the periods of high prices (the food crisis). Next, the link between ethanol and the sugars is only moderately stable ($b_{ij}=0.538$) and from the HT, the subdominant ultra metric distance between ethanol and each of the sugars cluster, the fuels cluster and biodiesel is almost the same. Therefore, the connection between the sugars and ethanol is not very convincing and we surely cannot make any strong statements about such a link. Obviously, the same holds for biodiesel — the stability and unambiguity of its links don't tell any strong statements. Lastly, the geographical properties of the retail fuels are less influential compared to the weekly frequency, which is in accord with the results for the whole sample.

#### 5.2.2. Crisis and post-crisis periods

The results differ considerably for the crisis and post-crisis periods. For the weekly frequency, we again observe strong and stable clusters of the retail fuels (geographically separated), the sugars and the feedstock from the MST (Fig. 5a). However, the standard S–C–W cluster is now supplemented by ethanol, which is quite strongly ($d_{ij}=0.9696$) and very stably ($b_{ij}=0.960$) linked to corn. On the other hand, biodiesel is quite modestly ($d_{ij}=1.1756$) and mainly very unsteadily ($b_{ij}=0.187$) connected to the crude oil. The clusters are well illustrated in the HT (Fig. 5b). We observe two strong branches — a retail fuels branch and a feedstock branch. Compared to the HT for the pre-crisis period, the fuels branch does not contain the biofuels. On contrary, ethanol is well linked in the fodder branch and biodiesel is far from the rest of the system. In addition to the interesting results for biofuels, we also find that crude oil is very similarly connected with the whole system with an exception of biodiesel.

In the medium term, the structure of links changes. From the MST (Fig. 5c), we observe that the sugars and feedstock clusters remain strong and stable. The previously standard retail fuels cluster differs. Interestingly, the strongest connection is between crude oil and the US gasoline ($d_{ij}=0.4533$) which is followed by the CO–GG pair ($d_{ij}=0.5050$). Biodiesel is strongly connected to the US diesel ($d_{ij}=0.6900$) with a modestly stable link ($b_{ij}=0.666$). Ethanol is again well and steadily connected to corn ($d_{ij}=0.7093$ with $b_{ij}=0.989$). The clusters are well represented in the hierarchical tree (Fig. 5d). First, we observe that the retail fuels connections are now dominated by the type of fuel rather than by the geographical properties (the same characteristic is present in the pre-crisis period and in the whole period as well). Second, biodiesel is a stable part of the fuels branch of the network, even though it is quite far from the rest of the branch. Third, ethanol creates a strong branch with the corn, wheat and soybeans. And last, the sugars are relatively far from the rest of the network. However, the network is altogether well connected because the longest distance is equal to 1.0076 compared to 1.2527 for the pre-crisis period.

### 5.3. Discussion

An important starting point for the discussion of our results is the comparison of two major biofuel markets covered in our analysis — US and EU. The EU is historically the largest producer, consumer and importer of biodiesel, which is the most important biofuel in

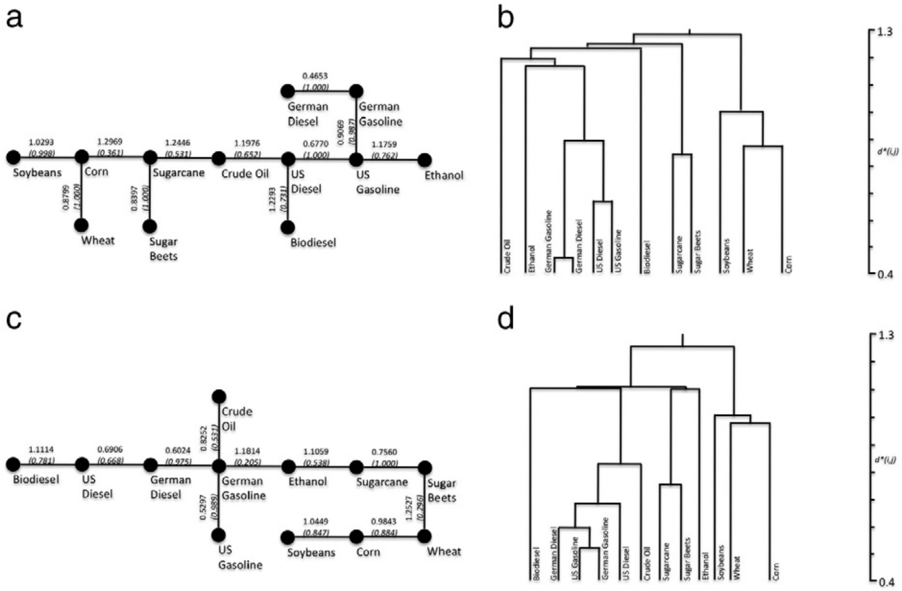


**Fig. 4.** Minimal spanning trees (first column) and hierarchical trees (second column) for network of returns and different frequencies (from the top — one week and one month and one quarter) for the pre-food-crisis period.

EU. According to Flach et al. (2011) on energy basis biodiesel represents about 80% of the total EU biofuels market in the transportation sector. Biodiesel was the first biofuel developed and used in the EU in the transport sector in the 1990s. At the time, the rapid expansion was driven by an increasing crude oil price, the Blair House Agreement of 1992 between US and EU on export subsidy and domestic subsidy reduction and resulting provisions of the EU's set-aside scheme, and generous tax incentives mainly in Germany. The Blair House Agreement allowed the EU to produce oilseeds for non-food use of up to 1 million Mt of soybean equivalent. EU biofuel goals set in directive 2003/30/EC (indicative goals) and in the RED 2009/28/EC (mandatory goals) further pushed the use of biodiesel. In addition, the Fuel Quality Directive gave the industry considerable latitude to market higher blends in the fuel supply. This means that the EU orientation on biodiesel was very much induced by public policies originating in 1990s. On the contrary to the EU situation, the US biofuel markets are dominated by ethanol.

The EU policy of setting a single target for all types of biofuel provides a flexibility for EU fuel markets to select a cost-effective biofuel types and technologies. The US approach of sectoral targets is missing this market flexibility, but it may provide market players a long-term confidence for introducing new investments in a broad range of renewable energy sources. More detailed comparison of the US and EU biofuel markets and policies is provided by Tyner (2010a) and Ziolkowska et al. (2010).

Because of crucial determining role of government policies in biofuel markets development both in the US and the EU, it is important to realize that US biofuels mandate was designed in volumes while the EU targets are in energy units. This means that in the US a liter of ethanol was equivalent to a liter of biodiesel as far as volumetric mandates

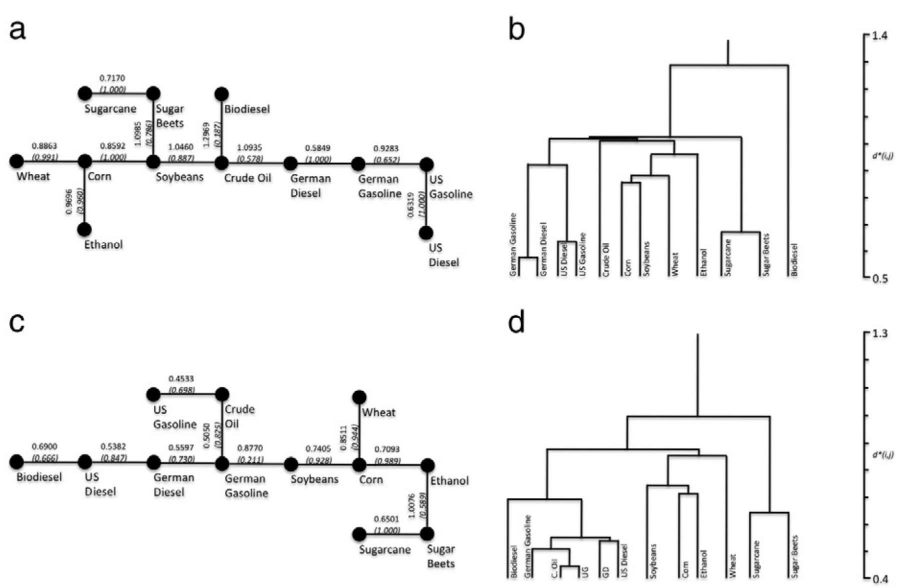


**Fig. 5.** Minimal spanning trees (first column) and hierarchical trees (second column) for network of returns and different frequencies (from the top — one week and one month and one quarter) for the food crisis and post-food-crisis period.

were concerned, while in the EU a kilojoule of ethanol is equivalent to kilojoule of pure biodiesel. According to Tyner (2010a) 1.65 l of ethanol have an energy equivalent of 1 l of biodiesel which means that EU system provides an incentive for private sector to use the biodiesel in order to meet the biofuel mandates while the US policy is biased towards the use of ethanol.

Another important difference among EU and US motor fuel markets is much higher share of diesel-engined car in Europe than in US. This historical difference was again caused by government policies, primarily by taxation of motor fuels. Since the fuel taxes in the US were historically much lower than in Europe, the higher fixed cost of diesel engines as compared to gasoline engines was more important than the variable cost advantage of diesel fuel. In addition, the relative tax differences among diesel and gasoline in Europe and the US meant that over the period covered in our paper, the consumer price of a liter of diesel was higher than that of gasoline in the US and vice versa in the EU.

After taking into account these differences among EU and US biofuel markets which are relevant for our whole analyzed period, we will discuss the differences and similarities in our both sub-periods separated by the global food crisis of 2007/2008. When we want to compare the results for the whole period and the two periods separated by the food crisis, we have to start from the observation that in the short term, both biofuels seem to be practically uncorrelated with either their producing factors or the other fuels. However, this uncorrelatedness is due to the fact that the connections in the whole network differ substantially in the pre-crisis and the post-crisis periods. In effect, these correlations are then nullified from the whole-sample perspective.

If we consider the networks as a whole and consider whether these are more or less connected, we look at the average length of the links in the network. The lower the average length, the closer the nodes in the network are, on average. The system is then more interconnected. The average tree lengths and corresponding standard deviations for both frequencies and three different periods are summarized in Table 4. We observe that the average link length decreases with a decreasing frequency for all three periods, which implies that the system is generally more connected in the medium term than in the short term. Comparing the results for the different periods, we see that the system is more interconnected for the post-crisis period $\left(\overline{d_{ij}}=0.8768\right)$ than both the pre-crisis period $\overline{d_{ij}}=(1.0020)$ and the whole sample $\left(\overline{d_{ij}}=0.9835\right)$ for weekly frequency. The difference is even more profound for the monthly period mainly between the pre-crisis and post-crisis periods with the average link lengths of 0.8373, 0.9478 and 0.7723 for the whole sample, pre-crisis and post-crisis periods, respectively. As a whole, the analyzed commodities are much more correlated in the post-crisis period than the others. Thus, the high prices of food are connected with more visible co-movements between the commodities, i.e. the correlations are not linear as they depend on the price. Taking standard deviations of the average tree lengths into consideration as well, we can see that the system as a whole is more interconnected for the post-crisis period as the standard deviations are markedly lower than for the pre-crisis period. The connections are even more stable for the post-crisis period compared to the pre-crisis period based on the average $b_{ij}$ (Table 4), even though the difference is more profound for the monthly frequency.

From economic point of view, our results show that short-term adjustments, which correspond more to random changes than systematic forces, do not form strong price links in the whole system of biofuels and related commodities. The picture changes by extending the analyzed horizon to one month since the MST and HT constructed with monthly data exhibit considerably more complex structure. Comparison of earlier period up to the food crisis of 2007/2008 reveals that the relaxation of capacity constraints both through production plants capacity building and through improvements in technological capacities of biofuels use in car engines means that especially for ethanol, the resulting equilibrium is driven by the government regulations with respect to minimal use of biofuels and by the prices of agricultural commodities serving as feedstock for biofuels in the more recent period after 2007. This contrasts with the situation on biodiesel market, where determining equilibrium factor are the prices of fossil fuels.

Our finding that the behavior of the biofuels system changed in the period after onset of the food crisis of 2007/2008 provides a supporting evidence to the currently prevailing opinion among agricultural and resource economists (Baffes and Haniotis, 2010) that the increase of commodity prices since 2007 marks a new period in the commodity prices development. While some earlier analyses (Piesse and Thirtle, 2009) treated 2007/2008 food crisis as a bubble, further price development indicates that the increase of commodity prices since 2007 is a long-lasting phenomenon. As documented by Haniotis (2012) the commodity price increase which we observe since about 2007 pertains not only to fuels and foods, but also to all commodities (agricultural, energy, fertilizers, metals, and minerals). The list of possible factors contributing to this increase is much wider than food demand, biofuels, low stocks of agricultural commodities and financialization of commodities which were the focus of policy debate during 2008–2011. It includes both macro factors (economics growth, weak dollar, fiscal expansion, low cost of capital, and financialization of commodities) and agricultural sector-specific factors (energy prices, weather, food demand, biofuels, agricultural policies, agricultural underinvestment, and low stocks of agricultural commodities).

While some earlier evaluations (Mitchell, 2008) pointed to biofuels as a major cause of 2007/2008 food crisis, subsequent research of Hochman et al. (2011a) and other authors shows that biofuels were only one of many contributors of price increase. Majority of this research dealing with the role of biofuels in the 2007/2008 food crisis concentrates on ethanol and main agricultural commodities (corn, soybean, rice, and wheat) and concludes that the role of biofuels in the price increase was noticeably stronger for corn than for soybeans, with soybean prices driven primarily by the increase in demand due to economic growth. This is in line with our results separating soybeans into a "food subgroup" of MST/HT a placing biodiesel into a distinctive "fuels group" as opposed to ethanol with strong connections to food commodities.

An important policy lesson of our analysis is to emphasize that the general statements about biofuels driving up the prices of agricultural commodities miss a critical distinction between different biofuels. We show that ethanol prices and biodiesel prices have clearly different places in a wide system of biofuels-related commodities. Our results confirm that discussion about food and biofuel prices is primarily relevant for ethanol, but not so much for biodiesel. While we present a strong correlation between prices of ethanol and it major feedstock corn and to a lesser extent other feedstocks, we do not obtain such results for biodiesel. The close connection of major biodiesel feedstock – soybeans – with corn and other grains shows that pricing of soybeans is more driven by its competition with corn for land and water resources and as major components of animal feed in livestock production in the US and abroad, especially in China.

In the terms of equilibrium model of price determination in Fig. 1, the biodiesel supply and demand are not as much constrained as the ones for ethanol. While both biodiesel and ethanol processing plants exhibit overcapacity, the use of ethanol is more restricted by a "blending wall" than it is a case for biodiesel. While in the absence of flex-fuel vehicles the "blending wall" for ethanol is on the level of 10 or 15%, for biodiesel the blends up to 20% are commonly accepted and ASTM D6751 regulation in US allows for biodiesel blends containing more than 20% of biodiesel subject to evaluation on case-by-case basis. Therefore the biodiesel demand shift caused by crude oil and fuel price changes may freely reach equilibrium point $E_1$ in Fig. 1 leading to high correlation among the prices of biodiesel and fossil fuels while for ethanol the blending wall restricts the attainment of this equilibrium which disrupt the close correlations between prices of ethanol and fossil fuels.

## 6. Conclusions

We analyzed the relationships between biodiesel, ethanol and related fuels and agricultural commodities with a use of minimal spanning trees and hierarchical trees. To distinguish between short-term and medium-term effects, we constructed the trees for different frequencies (weekly and monthly). Moreover, we were interested in different structure of connections before and during/after the food crisis of 2007/2008.

For the whole examined period, we found that in the short term, both analyzed biofuels are very weakly connected with the other commodities. In the medium term, the network structure becomes more interesting. The system practically splits into two branches — a fuels part and a food part. Biodiesel tends to the fuels branch and ethanol to the food branch. However, the results are much more pronounced for two periods separated by an outburst of the food crisis.

In the pre-crisis period, we mainly find that biofuels are only weakly connected to the whole network, even for the monthly frequency. However, we also uncovered that soybeans, wheat and corn are only weakly correlated with the rest of the network. This implies that when the food prices are low, these are very mildly connected to fuels and biofuels. The situation changes considerably for the crisis and post-crisis period, i.e. the period with higher food prices. Here, we observed that ethanol is well connected to corn, wheat and soybeans even in short term, and more strongly in medium term. On the other hand, biodiesel is very lowly correlated with the rest of the system in short term but becomes strongly and steadily connected to other fuels commodities in medium term. Reversely to the situation before the crisis, corn, wheat and soybeans are well connected with the whole network but the sugars are less correlated. Nevertheless, the correlations are considerably higher in the post-crisis period compared to the pre-crisis period.

Our results contributed to the policy debate about biofuels as possible (major) source of rises in food prices leading to food crises. We confirmed positive correlations among the prices of biofuels and food, but we shoved that the distinction should be made between different biofuels. The policy recommendation of carefully distinguishing between different biofuels is not new to the biofuels and food debate, but so far the distinction was drawn primarily between first generation and second generation biofuels with emphasis on ethanol related feedstock. Our contribution is in highlighting the differences among biodiesel and ethanol with respect to co-movements with food commodity prices and to emphasize time-varying nature of these co-movements. The investigation of time and price varying dynamic causal relations among prices of various biofuels and related commodities is a topic of our further ongoing research in this food-policy relevant area.

Finally, even though the methodology of taxonomy for economic time series is very simple and only transforms the correlations into distances, we were able to find several important results. We identified different biofuel prices network clusters corresponding to different binding constraints for the biofuels price equilibrium formation. The connections among different elements of biofuels network identified in this paper may be used as starting points for more detailed econometric time series investigations (identification of the most important connections in the system, identification of potential co-linearity, or even a basis for an optimal portfolio construction). The simplicity of the minimal spanning trees and hierarchical trees methodology allows to include a large number of prices and we therefore expect future research to expand our analysis both in terms of goods and locations in more detail. This will eventually create a good picture of how the relative food and fuel prices relate over space and time.

Note that the taxonomy methodology has been used for the first time on the biofuel systems in this paper and opens new possibilities for further research. First, a broader range of commodities and assets which might be important in the biofuels discussion – exchange rates, interest rates, commodities futures, stocks, climate conditions, exports and many others – can be included in the MST and HT analysis. A range of possible factors influencing clustering of commodity prices is suggested by Savascin (2011). Second, the proposed methodology can be accompanied by principal component analysis (Pearson, 1901) to give a more complex view on the cluster analysis. Third, conditional (time-varying) correlations can be taken into consideration and incorporated into MST/HT methodology to better describe the evolution in time. However, this would impose a specific model on the data-generating process of the analyzed series, which we wanted to avoid in this paper. Fourth, the time-dependent correlations analysis can be expanded to the frequency domain through wavelets which are able to separate time and frequency characteristics of the series (Vacha and Barunik, 2012). Discrete wavelets and corresponding coherences can be incorporated into the proposed methodology as well while still keeping the framework model-free. And fifth, the biofuels network can be analyzed with a 3D generalization of MST/HT methodology proposed by Song et al. (2011). As a starting point, the proposed methodology and obtained results uncover new frontiers in the biofuel systems research.

## Acknowledgments

The authors acknowledge the support from Energy Biosciences Institute at University of California, Berkeley (EBI). Karel Janda acknowledges research support provided during his long-term visits at University of California, Berkeley and Australian National University (EUOSSIC programme). Our research was supported by the grants P402/11/0948 and 402/09/0965 of the Czech Science Foundation, grant 118310 and project SVV 265 504 of the Grant Agency of the Charles University and by institutional support grant VSE IP100040. The views expressed here are those of the authors and not necessarily those of our institutions. All remaining errors are solely our responsibility.

## Appendix A

**Table 5**
Stationarity tests–Augmented Dickey–Fuller tests with (ADF1) and without (ADF2) a constant term testing the unit-root null hypothesis and KPSS testing the null of stationarity. The unit-root behavior is strongly rejected for all series while stationarity is not rejected for any of the series.

| Commodity | ADF1 | p-Value | ADF2 | p-Value | KPSS | p-Value |
|---|---|---|---|---|---|---|
| *Log-returns* | | | | | | |
| Crude oil | −13.3784 | 0.0000 | −13.2848 | 0.0000 | 0.1157 | >0.1 |
| Ethanol | −11.9235 | 0.0000 | −11.9329 | 0.0000 | 0.0536 | >0.1 |
| Corn | −13.3794 | 0.0000 | −13.3284 | 0.0000 | 0.1249 | >0.1 |
| Wheat | −12.7878 | 0.0000 | −12.7795 | 0.0000 | 0.0916 | >0.1 |
| Sugarcane | −14.3478 | 0.0000 | −14.2195 | 0.0000 | 0.0923 | >0.1 |
| Soybeans | −12.9502 | 0.0000 | −12.9467 | 0.0000 | 0.1046 | >0.1 |
| Sugar beets | −14.7365 | 0.0000 | −14.6001 | 0.0000 | 0.0658 | >0.1 |
| Biodiesel | −14.5312 | 0.0000 | −14.3312 | 0.0000 | 0.1676 | >0.1 |
| German diesel | −13.5802 | 0.0000 | −13.5123 | 0.0000 | 0.1050 | >0.1 |
| US diesel | −9.6001 | 0.0000 | −9.4996 | 0.0000 | 0.2574 | >0.1 |
| German gasoline | −12.6561 | 0.0000 | −12.6276 | 0.0000 | 0.0693 | >0.1 |
| US gasoline | −7.7812 | 0.0000 | −7.7353 | 0.0000 | 0.1472 | >0.1 |

## References

Baffes, J., Haniotis, T., 2010. Placing the 2006/2008 commodity price boom in perspective (July) World Bank Policy Research Working Paper, 5371. World Bank.
Beckman, J., Hertel, T., Tyner, W., 2011. Validating energy-oriented CGE models. Energy Econ. 33 (5), 799–806.
Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N., Mantegna, R., 2004. Networks of equities in financial markets. Eur. Phys. J. B 38, 363–371.

Busse, S., Brummer, B., Ihle, R., 2010. Interdependencies between fossil fuel and renewable energy markets: the German biodiesel market. DARE Discussion Paper, 1010. Georg-August University of Gottingen, Department of Agricultural Economics and Rural Development.
Chang, T.-H., Su, H.-M., Chiu, C.-L., 2011. Value-at-risk estimation with the optimal dynamic biofuel portfolio. Energy Econ. 33 (2), 264–272.
Ciaian, P., dArtis Kancs, 2011a. Interdependencies in the energy–bionergy–food price systems: a cointegration analysis (January) Resour. Energy Econ. 33 (1), 326–348.
Ciaian, P., dArtis Kancs, 2011b. Food, energy and environment: is bioenergy the missing link? (October) Food Policy 36 (5), 571–580.
de Gorter, H., Just, D.R., 2009a. The economics of a blend mandate for biofuels (August) Am. J. Agric. Econ. 91 (3), 738–750.
de Gorter, H., Just, D.R., 2009b. The welfare economics of a biofuel tax credit and the interaction effects with price contingent farm subsidies (May) Am. J. Agric. Econ. 91 (2), 477–488.
Drabik, D., 2012. The theory of biofuel policy and food grain prices (March) Working Paper, 2011–20. Cornell University.
Du, X., Yu, C.L., Hayes, D.J., 2011. Speculation and volatility spillover in the crude oil and agricultural commodity markets: a Bayesian analysis. Energy Econ. 33 (3), 497–503.
Flach, B., Lieberz, S., Bendz, K., Dahlbacka, B., 2011. EU-27 Annual Biofuels Report (June) GAIN Report NL1013. USDA Foreign Agricultural Service.
Haniotis, T., 2012. Does the price talk match the price walk? A review of evidence. Presentation at 123rd EAAE seminar in Dublin.
Hochman, G., Rajagopal, D., Zilberman, D., 2010. Are biofuels the culprit? OPEC, food, and fuel (May) Am. Econ. Rev. 100 (2), 183–187.
Hochman, G., Rajagopal, D., Timilsina, G., Zilberman, D., 2011a. Quantifying the causes of the global food commodity price crisis (August) Policy Research Working Paper, 5744. World Bank.
Hochman, G., Rajagopal, D., Zilberman, D., 2011b. The effects of alternative energy sources on the international oil market (Autumn) Appl. Econ. Perspect. Policy 33 (3), 402–427.
Janda, K., Kristoufek, L. and Zilberman, D., forthcoming. Biofuels: policies and impacts. Agricultural Economics — Czech.
Jang, W., Lee, J., Chang, W., 2011. Currency crises and the evolution of foreign exchange market: evidence from minimum spanning tree. Phys. A 390, 707–718.
Kantar, E., Deviren, B., Keskin, M., 2011. Hierarchical structure of Turkey's foreign trade. Phys. A 390 (20), 3454–3476.
Keskin, M., Deviren, B., Kocakaplan, Y., 2011. Topology of the correlation networks among major currencies using hierarchical structure methods. Phys. A 390, 719–730.
Kruskal, J., 1956. On the shortest spanning subtree of a graph and the traveling salesman problem. Proc. Am. Math. Soc. 7 (1), 48–50.
Long, X., Su, L., Ullah, A., 2011. Estimation and forecasting of dynamic conditional covariance: a semiparametric multivariate model. J. Bus. Econ. Stat. 29, 109–125.
Lucey, B.M., Gilmore, G.G., Boscia, M., 2011. Comovements in commodity markets: a minimum spanning tree analysis. Presentation for Midwest Finance Conference, 2012.
Mantegna, R., 1999. Hierarchical structure in financial markets. Eur. Phys. J. B 11, 193–197.
McPhail, L.L., 2011. Assessing the impact of US ethanol on fossil fuel markets: a structural VAR approach. Energy Econ. 33 (6), 1177–1185.
Mitchell, D., 2008. A note on rising food prices (July) World Bank Policy Research Working Paper, 4682. World Bank — Development Economics Group.
Mizuno, T., Takayasu, H., Takayasu, M., 2006. Correlation networks among currencies. Phys. A 364, 336–342.
Onnela, J.-P., Chakraborti, A., Kaski, K., Kartesz, J., 2002. Dynamic asset trees and portfolio analysis. Eur. Phys. J. B 30, 285–288.
Patton, A., 2006. Modeling asymmetric exchange rate dependence. Int. Econ. Rev. 47, 527–556.
Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. Philos. Mag. 2 (6), 559–572.
Peri, M., Baldi, L., 2010. Vegetable oil market and biofuel policy: an asymmetric cointegration. Energy Econ. 32 (3), 687–693.
Piesse, J., Thirtle, C., 2009. Three bubbles and a panic: an explanatory review of recent food commodity price events (April) Food Policy 34 (2), 119–129.
Pokrivcak, J., Rajcaniova, M., 2011. Crude oil price variability and its impact on ethanol prices. Agric. Econ. — Czech 57 (8), 394–403.
Rajagopal, D., Sexton, S., Hochman, G., Roland-Holst, D., Zilberman, D., 2009. Model estimates food-versus-biofuel trade-off. Calif. Agric. 63 (4), 199–201.
Rajcaniova, M., Pokrivcak, J., 2011. The impact of biofuel policies on food prices in the European Union. J. Econ. (Ekonomicky casopis) 59 (5), 459–471.
Saghaian, S.H., 2010. The impact of the oil sector on commodity prices: correlation or causation? (August) J. Agric. Appl. Econ. 42 (3), 477–485.
Savascin, O., 2011. The Dynamics of Commodity Prices: A Clustering Approach. (November) University of North Carolina, Chapel Hill.
Serra, T., 2011. Volatility spillovers between food and energy markets: a semiparametric approach. Energy Econ. 33 (6), 1155–1164.
Serra, T., Zilberman, D., Gil, J.M., Goodwin, B.K., 2010. Price transmission in the US ethanol market. In: Khanna, M., Scheffran, J., Zilberman, D. (Eds.), Handbook of Bioenergy Economics and Policy: Natural Resource Management and Policy. Springer, pp. 55–72 (Chapter 5).
Serra, T., Zilberman, D., Gil, J.M., 2011a. Price volatility in ethanol markets. Eur. Rev. Agric. Econ. 38 (2), 259–280.
Serra, T., Zilberman, D., Gil, J.M., Goodwin, B.K., 2011b. Nonlinearities in the U.S. corn–ethanol–oil–gasoline price system (January) Agric. Econ. 42 (1), 35–45.
Song, W.-M., Di Matteo, T., Aste, T., 2011. Nested hierarchy in planar graphs. Discret. Appl. Math. 159, 2135–2146.
Tabak, B., Serra, T., Cajueiro, D., 2009. The expectation hypothesis of interest rates and network theory: the case of Brazil. Phys. A 388, 1137–1149.
Tabak, B., Serra, T., Cajueiro, D., 2010. Topological properties of commodities networks. Eur. Phys. J. B 74, 243–249.
Tumminello, M., Di Matteo, T., Aste, T., Mantegna, R., 2007. Correlation based networks of equity returns sampled at different time horizons. Eur. Phys. J. B 55, 209–217.
Tyner, W.E., 2010a. Comparison of the US and EU approaches to stimulating biofuels. Biofuels 1 (1), 19–21.
Tyner, W.E., 2010b. The integration of energy and agricultural markets. Agric. Econ. 41 (s1), 193–201.
Vacha, L., Barunik, J., 2012. Co-movement of energy commodities revisited: evidence from wavelet coherence analysis. Energy Econ. 34 (1), 241–247.
Wixson, S.E., Katchova, A.L., 2012. Price asymmetric relationships in commodity and energy markets (February) Presentation at 123rd EAAE seminar in Dublin.
Zhang, Z., Lohr, L., Escalante, C., Wetzstein, M., 2009. Ethanol, corn, and soybean price relations in a volatile vehicle-fuels market. Energies 2 (2), 320–339.
Zhang, Z., Lohr, L., Escalante, C., Wetzstein, M., 2010. Food versus fuel: what do prices tell us. Energy Policy 38 (1), 445–451.
Ziegelback, M., Kastner, G., 2011. European rapeseed and fossil diesel: threshold cointegration analysis and possible implications. Presentation at GEWISOLA conference in Halle, Germany.
Zilberman, D., Hochman, G., Rajagopal, D., Sexton, S. and Timilsina, G., forthcoming. The impact of biofuels on commodity food prices: assessment of findings. American Journal of Agricultural Economics doi:10.1093/ajae/aas037.
Ziolkowska, J., Meyers, W.H., Meyer, S., Binfield, J., 2010. Targets and mandates: lessons learned from EU and US biofuels policy mechanisms. AgBioforum 13 (4), 398–412.