V Praze 18. ledna 2017

Jedná o práci diplomovou, u níž je žádáno o uznání za práci rigorózní.

Mgr. Ondřej Filip

Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

RIGOROUS THESIS

Food vs. Fuel: The Role of Bioenergy

Author: **Mgr. Ondřej Filip**

Supervisor: **prof. Ing. Karel Janda M.A., Dr., Ph. D.**

Academic Year: **2016/2017**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 18, 2017

Signature

Acknowledgments

This thesis benefited from the European Union's Horizon 2020 Research and Innovation Staff Exchange programme under the Marie Sklodowska -Curie, grant agreement No 681228.

I am grateful especially to prof. Ing. Karel Janda M.A., Dr., Ph.D. for his valuable assistance throughout the work on this thesis. I highly appreciate his initial advice regarding the choice of the topic. I would like to thank him for his accurate comments and numerous suggestions which improved my thesis significantly. I would like to thank to PhDr. Tomáš Havránek Ph.D. for a great LaTeX template. A very special word of gratitude belongs to my wife.

Bibliographic entry

Filip, O. (2015): "Food vs. Fuel: The Role of Bioenergy." *(Unpublished master's thesis)*. Charles University in Prague. Supervisor: prof. Ing. Karel Janda M.A., Dr., Ph. D.

Length: 121,976 characters

Abstract

This thesis studies the relationship between the first generation biofuels and selected commodities and assets in the USA, Europe, and Brazil. It is the first attempt to combine the taxonomy and wavelet analyses in a single research application. Our unique dataset comprises 32 weekly price series covering the 2003–2015 time period. First, we employ a method of minimum spanning trees and hierarchical trees to model a biofuel-related price network. We demonstrate a development phase shift between Brazilian and the US/EU biofuel industries. We reveal a strong and stable connection between Brazilian ethanol and its main production factor, local sugarcane. We further find that US ethanol is closely linked to corn. In the contrary, European biodiesel exhibits only moderate ties to its production factors. Subsequent wavelet analysis scrutinizes the identified price connections both in time and frequency domains. Both Brazilian and US ethanols are found to be positively related to their respective feedstock commodities. In particular, feedstock proves to lead the price of the biofuel and not vice versa. Moreover, the dynamics remains qualitatively unchanged when controlled for the influence of crude oil.

JEL Classification	C22, C38, Q16, Q42
Keywords	biofuels, ethanol, biodiesel, taxonomy, minimum spanning tree, hierarchical tree, wavelet coherence

Author's e-mail	ofilip38@gmail.com
Supervisor's e-mail	Karel-Janda@seznam.cz

Abstrakt

Tato práce zkoumá vztah biopaliv první generace k vybraným komoditám a dalším aktivům v USA, Evropě a Brazílii. Jedná se o první aplikaci kombinující taxonomní a vlnkovou analýzu v rámci jedné práce. Unikátní dataset obsahuje 32 týdenních cenových řad, které pokrývají období let 2003 až 2015. Nejdříve používáme metodu minimální kostry grafu a hierarchického stromu, abychom modelovali systém cen souvisejících s biopalivy. Ukazujeme fázový posun mezi brazilskou výrobou biopaliv a vývojem v USA a EU. Nalézáme silné a stabilní spojení mezi brazilským etanolem a jeho hlavním výrobním faktorem, tamní cukrovou třtinou. Dále odhalujeme, že americký etanol je silně spojen s cenou kukuřice. Naproti tomu evropská bionafta vykazuje pouze slabé napojení na své výrobní faktory. Následná vlnková analýza zkoumá zjištěné závislosti v časové a ve frekvenční doméně. Nacházíme, že americký i brazilský etanol jsou dlouhodobě pozitivně svázány s cenami svých výrobních surovin. Ceny biopaliv jsou navíc taženy cenami surovin, nikoli naopak. Naše výsledky zůstávají kvalitativně nezměněny, i když odfiltrujeme možný vliv cen ropy.

Klasifikace JEL	C22, C38, Q16, Q42
Klíčová slova	biopaliva, etanol, bionafta, taxonomie, minimální kostra grafu, hierarchický strom, vlnková koherence

E-mail autora	`ofilip38@gmail.com`
E-mail vedoucího práce	`Karel-Janda@seznam.cz`

Contents

List of Tables

List of Figures

Acronyms

FOB Free on Board

BRL Brazilian Real

HT Hierarchical Tree

JB Jarque–Bera test

SW Shapiro–Wilk test

DAX Deutscher Aktienindex

MST Minimum Spanning Tree

ADF Augmented Dickey-Fuller test

LIBOR London Interbank Offered Rate

EIA U.S. Energy Information Administration

KPSS Kwiatkowski–Phillips–Schmidt–Shin test

Bovespa Bolsa de Valores do Estado de Sao Paulo

FTSE 100 Financial Times Stock Exchange 100 Index

CEPEA Centro de Estudos Avancados em Economica Aplicada

FAO Food and Agricultural Organization of the United Nations

Master's Thesis Proposal

Author	Mgr. Ondřej Filip
Supervisor	prof. Ing. Karel Janda M.A., Dr., Ph. D.
Proposed topic	Food vs. Fuel: The Role of Bioenergy

Topic characteristics A lively global-scale discussion about the effects of biofuels production became particularly intense with the outbreak of the 2007–2008 world food price crisis. Do the efforts to increase biofuels production have to do with the increases in agricultural commodity prices? Is the future food supply threatened by the production of biofuels intensified in the recent years? In this thesis, we are going to explore what items form a food-fuel-biofuel system by identifying relevant price transmission links. We intend to examine and subsequently graphically visualize the nature of identified price interdependencies stressing a broadly defined system of relevant biofuel-related commodities and assets.

Hypotheses

1. There exists a structured biofuels-related system of interdependent commodities and assets.
2. Although biofuels and food are positively linked, we have to distinguish between different biofuels. Biodiesel and ethanol have different places and roles within the price system.
3. The price co-movements within the biofuel-related system evolve and change over time.

Methodology This thesis intends to bring together two empirical approaches which are both rather innovative in biofuel economics. In the first section of the empirical part, we plan to identify the components (commodities as well as other items) which form an interlinked biofuel-related price system. At this point, we are going to utilize the taxonomy methodology of minimum spanning trees and hierarchical trees first employed on financial time series by Mantegna (1999).

We aim to build on Krištoufek *et al.* (2012) who introduced the taxonomy perspective into the biofuels context. We attempt to follow the suggestions given in their 2012 paper and extend the analysis through increasing the scope and complexity of studied price system. In particular, we will gather and employ a large dataset of time series including fossil fuels, biofuels and associated feedstock, food as well as other relevant commodities and assets as outlined for example by Savascin (2011). Through constructing the minimum spanning trees we are going to reveal the most important price connection in the whole framework. We will then describe the hierarchical structure of the biofuel-related system using a perspective of hierarchical trees. We plan to employ the (logarithmic transformation of) price series separately with both weekly and monthly frequencies to see whether the observed relationships occur in the short or medium term.

Once we will have selected the most important links within the biofuels price system, we aim to study their evolution both in time and frequency domains. In order to accomplish this goal we are going to employ the wavelet analysis of the selected time series. In the biofuels context, partial wavelet coherence was first utilized by Krištoufek *et al.* (2014). The method is model-free and allows to study behavior of a given price link and to observe its evolution in time. We will utilize the approach of wavelet coherence to explore, describe, and visualize the time and scale evolvement of the most important pair interdependencies while projecting them into two dimensional charts rich in information.

Expected Contribution Our contribution lies in utilizing the taxonomy approach for identifying the relevant items of the biofuels-related system. The use of both minimum spanning trees and hierarchical trees is still rather an innovative method to study the price transmission in biofuels system. The taxonomy perspective allows for studying a high number of time series simultaneously thus contrasting with previous econometric analyses which mostly focused only on a small group of commodities. Covering a wide range of commodities and other assets we will provide a new insight into the system of biofuel-related assets.

A further step towards a clearer understanding of food-fuel system will be made through the wavelet analysis. Wavelet coherence will allow us to properly describe the behavior of important price co-movements taking into account both time and frequency domains. Potential answers stemming from this work are of interest especially for policy makers dealing with biofuels and setting agriculture and energy-related regulations. They matter on both national and international levels.

Outline

1. Introduction
2. Literature Review
3. Methodology
4. Data
5. Taxonomy Analysis
6. Wavelet Analysis
7. Conclusion

Core bibliography

1. Serra, T. & D. Zilberman (2013): "Biofuel-related price transmission literature: A review." *Energy Economics* **37(0)**: pp. 141–151.
2. Mantegna, R. (1999): "Hierarchical structure in financial markets." *European Physical Journal B - Condensed Matter and Complex Systems*, **11(1)**: pp. 193–197.
3. Kristousek, L., K. Janda, D. Zilberman (2012): "Correlations between biofuels and related commodities: A taxonomy perspective." *IES Working Paper* **15/2012**, IES FSV, Charles University
4. Savaşsçin, Ö. (2011): "The dynamics of commodity prices: A clustering approach." *Univesity of North Carolina, Chapel Hill*
5. Vacha, L., K. Janda, L. Kristoufek, D. Zilberman (2013): "Time-frequency dynamics of biofuels-fuels-food system." *Energy Economics* **40(0)**: pp. 233–241.
6. Kristoufek, L., K. Janda, D. Zilberman (2015): "Co-movements of Ethanol Related Prices: Evidence from Brazil and the USA." *GCB Bioenergy*

Author

Supervisor

Chapter 1

Introduction

This thesis undertakes an innovative research attempt in the field of biofuel economics. Its coverage ranges from a rapid development of global biofuel industry to the most recent period of falling commodity prices. First, we use the taxonomy method of Minimum Spanning Tree (MST) and Hierarchical Tree (HT) to identify potential price transmission channels within a system of biofuel-related commodities and assets. Second, the core price connections revealed between biofuels and feedstock commodities are further explored using the wavelet coherence analysis. Here, we present the first attempt to combine the taxonomy and wavelet analyses within a single research application. Our study focuses on the world's largest biofuel markets; the United States of America, Brazil, and the European Union. On energy basis, we cover a vast majority of the world's biofuel production over the period of the last decade.

Whether we like it or not, biofuels matter. Until today, more than 60 nations worldwide implemented biofuels blending mandates. Most importantly, with 54 billion liters of ethanol generated in 2014, the US itself accounts for more than a half of the world's ethanol production. First generation biofuels, i.e. ethanol and biodiesel, are made from crops grown on agricultural land. About 90% of US ethanol is produced from corn. Ethanol now consumes 40% of the country's total corn production, a huge increase from just 5% in 2000. The US ethanol industry is followed by Brazil, historically the world's pioneering biofuel economy. Brazilian ethanol is made from sugarcane. Annually, over 55% of Brazilian sugarcane harvest is used to produce ethanol (Conca 2014). In contrast, European biofuel industry is based on biodiesel produced from rapeseed and other vegetable oils. The EU produces almost a half of the world's biodiesel. Biofuels production represents a topic far exceeding national borders.

Biofuels have become increasingly important since the oil export embargo and resulting oil crisis of 1970s. Originally, biofuels attracted the attention of policymakers as a way to support energy self-reliance especially as a substitute for liquid fossil fuels in transportation. Biofuels were seen as an alternative mitigating the country's dependence on crude oil imports while fostering domestic employment, promoting technological innovation, and keeping the associated income stream in the country. Fast growth of biofuel industries owes a lion's share of its success to a wide government-backed support. The major biofuel markers have been shaped by targets, blending mandates, tax exemptions, and subsidies. During the last decade, biofuels were often in the public eye because their intensified usage rose new economics, environmental, and ethical concerns.

In the last years, we observed very volatile prices of agricultural commodities. Being made from crops, biofuels effectively compete over land and water supply with the crop's other uses, e.g. food or animal feed. For example, corn represents not only a primary US biofuel feedstock, but also an essential animal feed. At the same time, corn is largely used in both food and beverage industries. Since the US produces about 40% of the world's corn, the US domestic corn demand may significantly influence the commodity's world price. Other associated concerns include indirect land use change, increased use of fertilizers or water scarcity issues. As reviewed in Chapter 2, this area attracted a lot of research attention in the last few years. Due to many unanswered questions, this field of interest remains highly relevant and provides plenty of opportunities for future research.

This thesis makes a major step towards a profound understanding of the core price dependencies between biofuels, feedstock, fossil fuels, and other relevant assets. The objective of this thesis is to analyze how the associated price system is structured and whether it reflects the biofuels production logic in respective geographic markets. We trace the system's evolution in time and elaborate on correlation patterns emerging throughout the analysis. Apart from observing the network as a whole, we pay a special attention to individual price relationships between the biofuels and their respective production factors. Is the price of US ethanol tied to its feedstock price? What other commodities does it depend on? How about the Brazilian or European biofuel markets, do they exhibit any differences? Do the roles of individual biofuels change in time and with different frequencies? These and similar questions will be in the center of our research interest.

Present study takes advantage of two methods that are novel to financial time series analysis. First, we use the taxonomy method of minimum spanning trees and hierarchical trees introduced into biofuels context by Kristoufek *et al.* (2012a). The taxonomy trees help us classify and visualize the system of biofuel-related commodities and assets. Resulting structures highlight the most important connections among studied items based on their mutual correlation. We construct these taxonomy objects for several subperiods and different frequencies (weekly and monthly). Second, particular biofuel–feedstock price pairs resulting from taxonomy perspective get further explored using the wavelet coherence analysis. Wavelet analysis was first applied on biofuels data by Vacha & Barunik (2012). This model-free approach allows for studying the correlation between two time series in both time and frequency domains. Moreover, wavelet framework evaluates the direction of the studied relationship. This feature makes up for a major limitation of taxonomy trees. Additionally, partial wavelet coherence enables us to remove a possible influence of other variables that may potentially correlate with both of the studied series. Here, we control for a possible influence of crude oil. Our novel combination of taxonomy and wavelet analyses is not limited by prior model assumptions (except for stationarity required by correlation in the taxonomy part). Therefore, it is viable to implement several recommendations of prior research and to provide a broad data coverage.

The rest of the thesis is structured as follows: Chapter 2 reviews relevant literature recently dealing with food–energy price links and economic impacts of biofuels. Chapter 3 introduces the employed methodology and develops the theoretical background of our toolbox. Chapter 4 describes both qualitative and quantitative features of our dataset. Chapter 5 performs the taxonomy analysis subsequently followed up by the wavelet analysis in Chapter 6. Finally, Chapter 7 concludes.

Chapter 2

Literature Review

This chapter aims to deliver a review of recent literature focused on studying the economic impacts of biofuels. There has been a variety of research interests as well as methods employed to investigate the economics of biofuels worldwide. We will pay our primary attention to reviewing the literature dealing with food–energy price links. We intend to provide a concise overview of the latest attempts elaborating on the system of biofuel related commodities.

Notably, Ciaian & Kancs (2011) combine a theoretical framework with empirical evidence to scrutinize the linkages between energy, biofuel, and agricultural markets. Their aim is to address the role of biofuels for agricultural prices. While employing a Vector Autoregressive Model, their cointegration analysis uses 1993–2010 time series data on crude oil and frequently traded agricultural commodities. The results are interpreted for three individual subperiods and confirm that *"energy prices do affect prices for agricultural commodities and the interdependencies between the energy and food markers are increasing over time"*(Ciaian & Kancs 2011, p.15). For earlier studies focused on competitive food–biofuel–fuel markets refer among others to de Gorter & Just (2009a;b) and Drabik (2011). Although there has been a vivid research exploring the international biofuel related markets, Serra & Zilberman (2013) note that a majority of studies does not in fact include biofuel price data into their analyses.

As pointed out by Serra & Zilberman (2013), modeling the price level patterns has received the main research attention as compared to price volatility interactions. To introduce the price links between agricultural and energy commodities, the concept of partial equilibrium is often found useful. Kristoufek *et al.* (2012a) use a partial equilibrium framework based on Serra *et al.* (2010)

to motivate the theoretical background of their biofuel related price system. In the model, the price of a particular biofuel would be simply determined by forces of supply and demand. However, there exist two exogenous limitations imposed on the model. A regulatory constraint (such as mandates, subsidies, blending obligations or similar regulatory support) sets the minimum quantity of a biofuel produced. On the other hand, technological feasibility (such as production capacities) restricts the maximum biofuel amount available on the market. While the prices of feedstock and fossil fuels are exogenous, the price of biofuel is determined within the market operating under a constrained equilibria setting. Kristoufek *et al.* (2012a) further remind us that both technological and regulatory constraints may be responsible for preventing high and positive correlation among prices in a widely defined biofuels commodity system.

In an effort to explore the impact of US biodiesel production on the price level and volatility of agricultural commodities, Hao *et al.* (2013) study both the short– and long–run relationships between fuel and agricultural commodity markets. Under the Vector Error Correction Model, the study employs 2006–2011 weekly price series representing biodiesel, its feedstock, and fossil fuels. Biodiesel is found to have a long-run price connection with soybeans. The results indicate that an increase in soybeans' price translates into a higher biofuel price, although the causal relationship is rather week. Oil price movements, however, represent an influential driving force for both agricultural and fuel prices.

de Gorter *et al.* (2013b) also focus on the US biofuel market. They develop an empirical model including biodiesel production from both soybeans and canola feedstock. It is stressed that soybeans and canola are used both in biofuel and meal production. A competing commodity allocation (feedstock versus food) alters a usual direct link between biofuel and its feedstock. Market equilibrium changes subject to volatility in the crude oil price. If biofuel enjoys a tax exemption, rising crude oil prices increase also the price of biodiesel. The opposite effect of crude oil price on biodiesel applies in case of a blending mandate requirement. However, the effect of an oil price shock on feedstock itself is not straightforward. Nonetheless, in case of blending mandate, the shock translates into a smaller oilseed price change relative to tax exemption setting. Other works focusing on the US food–biofuel market include Miller *et al.* (2012); Wang & McPhail (2012) or Du & Hayes (2012).

Specifics of recent development on the Brazilian biofuel market have been investigated by de Gorter *et al.* (2013a). The paper specifies an empirical model of a unique Brazilian sugar–ethanol fuel market. The analysis yields surprising results. The authors clearly identify two policies that actually harm the ethanol industry instead of helping it as generally believed. Namely, a low tax on gasoline and a high tax exemption for anhydrous ethanol were found to translate into lower ethanol prices which contrasts with the mechanism known from the US.

Drabik *et al.* (2014) study how biofuels affect the price transmission within the food chain. Their analysis focuses on the US corn and ethanol markets. Depending strongly on a source of the market shock and policy regime, the existence of biofuels is found to considerably impact the elasticity of price transmission. Interestingly, the authors find that the presence of biofuels substantially tempers the reaction of corn and food prices to shocks in their respective markets. The dependence of agricultural production on food markets may be thus reduced through the existence of biofuels.

Other environmental benefit is attributed to biofuels by Piroli *et al.* (2014). The authors study an impact of a rising bioenergy production on global CO_2 emissions. Employing a structural vector autoregression framework the study covers 1961–2009 time series data on both global biofuel production and CO_2 emissions. Although biofuels may increase CO_2 emission in the short–term, they are associated with a global significant CO_2 reduction in medium to long-term.

Biofuels have been recently studied in other environmental contexts, too. For example Rajcaniova *et al.* (2014) explore how a global production of biofuels affects the land use worldwide. An econometric analysis of time series on fuel, biofuel and agricultural commodities (both price and production) is performed. The results indicate that in the context of rising energy prices, the increasing production of bioenergy adds to a land use change. In particular, the study reports two effects. First, agricultural area is increasing due to a rising biofuel production. Second, there exists a substitution effect from food to energy crops being planted on the existing agricultural land.

Various implications identified during the years since the introduction of biofuels are brought together by Hochman (2014). An indirect land use change is one of the most evident phenomena accompanying biofuels commercialization. Not only biofuels compete with food, but their commercial production still requires further technological progress to become competitive. The process of biofuels commercialization proves more costly than expected.

In order to understand both the existing and upcoming food-price challenges, a lot of literature focus predominantly on food, energy, and environmental policy regulations and their mutual coordination (refer to Carter *et al.* 2012; McPhail & Babcock 2012; Abbott 2013; Peri & Baldi 2013; Rausser & de Gorter 2013; Drabik *et al.* 2014, inter alia). In particular, a recent article by Zilberman *et al.* (2014) focuses on both macro and micro level aspects of the political economy of biofuels.

An innovative line of research is represented by Kristoufek, Janda, & Zilberman (2012a;b; 2013; 2014; 2015). They have studied various aspects of the biofuel–fuel–food system. Adopting an international perspective, they cover the US, Brazilian and European biofuel markets instead of limiting their attention to one market only. While employing the actual biofuel price data which is rather rare, the authors have introduced unconventional tools into the biofuels research. Kristoufek *et al.* (2012a) are the first to apply the taxonomy methodology of MST and HT in the context of biofuels. Price series on biofuels, respective feedstock and fossil fuels are analyzed for 2003–2011 time period. The paper identifies a structured price system. The links among the commodities are weaker in the short term. However, the connections become stronger in the medium term. While ethanol tends to the food part, biodiesel is more prone to belong to the fuel part.

Another innovative attempt by Vacha *et al.* (2013) introduces the wavelet methodology to biofuels research. Wavelets–not a traditional tool in economics–are usually used to analyze the information contained in signals. This model free approach allowed the authors to study the correlation between fossil fuels, biofuels, and agricultural commodities in both time and frequency domains. The analysis reveals two strongly correlated pairs: ethanol–corn, biodiesel–German diesel and shows that during a food crisis period biofuels do react more rapidly to changes in the price of their producing factors.

The paper has been recently followed by another wavelet application. Kristoufek *et al.* (2015) use wavelet coherence methodology to explore the relationship between both Brazilian and US ethanols and their respective feedstock commodities. Their results illustrate that there exists a strong and stable relationship between both US ethanol–corn and Brazilian ethanol–sugar price pairs. Moreover, both pairs are characterized by feedstock leading the price of ethanol.

Chapter 3

Methodology

This chapter introduces individual items of a quantitative toolbox employed throughout the empirical part of the thesis. We use analytical methods that are not common in economics. In a consecutive manner, we describe the principles used in the construction of Minimum spanning trees and Hierarchical trees before arriving at the methodology of Wavelet coherence analysis. A substantial part of this section is attributed to previous works of Mantegna (1999); Mantegna & Stanley (2000); Kristoufek *et al.* (2012a), and Vacha *et al.* (2013). Mantegna (1999) was the first to use the taxonomy method of MST and HT on financial data to study the structure formed by components of S&P 500 index. This thesis further benefits from Kristoufek *et al.* (2012a) who introduced the taxonomy technique into the context of biofuels. The use of wavelet coherence approach on biofuels data was pioneered by Vacha *et al.* (2013). In order to extend previous research and to allow for comparability of results, presented methodology will be motivated by the above mentioned studies.

3.1 Distance metric

We are interested in establishing a metric that we will use to measure the interconnections within a group of commodities and assets. Before eventually arriving at a distance measure, we introduce important concepts which we will build our methodology upon. Following definitions and concepts are provided based on Wooldridge (2008).

Let $X(t)$ be a *stochastic process*. $X(t)$ is then a time ordered sequence of random variables given as $\{X(t) : t \in T,\ T \subset N\}$.

Definition 1. **(Stationarity)**

A stochastic process $\{X(t) : t = 1, 2, \dots\}$ is *stationary* if for every collection of time indices $1 \leq t_1 < t_2 \dots < t_m$, the joint distribution of $X(t_1), X(t_2), \dots, X(t_m)$ is the same as the joint distribution of $X(t_1+h), X(t_2+h), \dots, X(t_m+h)$ for all integers $h \geq 1$.

Definition 2. **(Weak Stationarity)**

We say that a stochastic process $\{X(t) : t = 1, 2, \dots\}$ with finite second moment $\left[\mathbb{E}(X(t)^2) < \infty\right]$ is *weakly stationary*[1] if it holds that:

(1) $\mathbb{E}\big(X(t)\big)$ is constant,

(2) $Var\big(X(t)\big)$ is constant,

(3) for any $t, h \geq 1, Cov\big(X(t), X(t+h)\big)$ depends only on h and not on t.

It is a standard practice to employ the sample correlation coefficient[2] to measure linear dependence between two time series. Having a pair of assets i and j with values X_{it} and X_{jt} where $t = 1, \dots T$ we obtain the *sample Pearson correlation coefficient* as:

$$\widehat{\rho_{ij}} = \frac{\sum_{t=1}^{T}(X_{it} - \overline{X_i})(X_{jt} - \overline{X_j})}{\sqrt{\sum_{i=1}^{T}(X_{it} - \overline{X_i})^2 \sum_{i=1}^{T}(X_{jt} - \overline{X_j})^2}}, \tag{3.1}$$

where $\overline{X_i}$ and $\overline{X_j}$ stand for the time series averages defined as $\overline{X_i} = \frac{1}{T}\sum_{t=1}^{T} X_{it}$ and $\overline{X_j} = \frac{1}{T}\sum_{t=1}^{T} X_{jt}$. Value of the correlation coefficient ranges from -1 to 1 with the following logic:

$$\rho_{ij} = \begin{cases} 1 & \text{perfect positive correlation} \\ 0 & \text{no correlation} \\ -1 & \text{perfect negative correlation} \end{cases}$$

It needs to be stressed that a sample correlation coefficient can only be meaningfully computed for series with well defined means and variances. Thus, the above defined (weak) stationarity is required (Kristoufek *et al.* 2012a).

[1] A term *covariance stationary* is preferred by Wooldridge (2008).

[2] A mention about correlation coefficient will always refer to the concept of Pearson's product-moment correlation coefficient.

Correlation matrix $\mathbb{C}$ is a matrix of correlation coefficients given as:

$$\mathbb{C} = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & 1 \end{pmatrix}, \tag{3.2}$$

where n is the number of assets. Correlation matrix $\mathbb{C}$ fulfills following properties that will be important for further analysis, see Mantegna (1999).

(1) With n rows and n columns $\mathbb{C}$ is a $n \times n$ square matrix,

(2) All items on the diagonal of $\mathbb{C}$ are equal to 1, thus $\forall \rho_{ij}$, $i = j$, it holds that $\rho_{ii} = 1$,

(3) Since $\rho_{ij} = \rho_{ji} : \forall i, j \in N$ it follows that $\mathbb{C}$ is a symmetric matrix,

(4) It follows from (1),(2) and (3) that $\frac{n \cdot (n-1)}{2}$ correlation coefficients suffice to fully describe $\mathbb{C}$.

Mantegna (1999) argues that a simple correlation coefficient cannot be used as a measure of distance because it violates the axioms of the Euclidian metric. However, he shows how to transform correlation into a distance measure. In line with Kristoufek *et al.* (2012a), we employ the following transformation of correlation coefficient:

$$d_{ij} = \sqrt{2(1 - \rho_{ij})}. \tag{3.3}$$

Unlike ρ_{ij}, d_{ij} fulfills the axioms of Euclidian distance:

(1) *Identity:* $d_{ij} = 0 \iff i = j, \forall i, j \in N$

(2) *Symmetry:* $d_{ij} = d_{ji}, \forall i, j \in N$

(3) *Triangle inequality:* $d_{ij} \leq d_{ik} + d_{kj}, \forall i, j \in N$

Once the Euclidian axioms are satisfied, the proposed non-linear transformation d_{ij} can be used as a suitable *measure of distance.* Since the correlation coefficient ranges between -1 and 1, it follows that d_{ij} takes corresponding values between 0 and 2.

$$d_{ij} = \begin{cases} 0 & \text{perfect positive correlation} \\ \sqrt{2} & \text{no correlation} \\ 2 & \text{perfect negative correlation} \end{cases}$$

Using d_{ij}, we transform each correlation coefficient ρ_{ij} in matrix $\mathbb{C}$ into a distance measure and obtain the following *distance matrix* $\mathbb{D}$:

$$\mathbb{D} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{pmatrix} = \begin{pmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & 0 \end{pmatrix} \tag{3.4}$$

Having originated from $\mathbb{C}$ through a transformation of individual correlation coefficients, $\mathbb{D}$ shares the properties that have been stipulated for $\mathbb{C}$. In particular, it immediately follows from (2) that all items on the main diagonal of $\mathbb{D}$ are equal to zero.

3.2 Minimum Spanning Tree

In order to become acquainted with concepts of MST and HT, let us first introduce a few principles from the graph theory. For this purpose, we refer to Bondy & Murty (1976), Diestel (2000) and Matousek & Nesetril (2008).

A *graph* is an ordered pair $G = (V, E)$ of sets such that $E \subseteq [V]^2$. While V represents a nonempty set of *vertices* (also *nodes* or *points*), M stands for a set of *edges* (also called *lines* or *links*) that is disjoint from V. Set E forms a 2-element subset of V. A graph G is said to be *non-empty*, if it at least $V \neq \emptyset$.

We say that a non-empty graph G is *connected*, if any two of its vertices are linked by a path in G. An edge both beginning and ending in the same vertex forms a *loop* since it does not connect two different vertices. An *acyclic graph*—the one without a loop—is called a *forest*. Combining the concept of a connected graph with that of a forest we arrive at the following definition.

Definition 3. **(Tree)** A connected forest is called a *tree*.

Following Diestel (2000), we say that a graph G' is a *subgraph* of G (or $G' \subseteq G$), if $V' \subseteq V$ and $E' \subseteq E$. Moreover, if $V' = V$, than $G' \subseteq G$ is a *spanning* subgraph of G.

Definition 4. **(Spanning Tree)** A tree that is a spanning subgraph of G is said to be a *spanning tree* of G.

Diestel (2000) shows that for every connected graph there exists at least one spanning tree. However, we strive to find a specific spanning tree among all the available trees—a Minimum Spanning Tree, MST. Edges of such a tree represent then the most important links in the underlying network. To allow edges of the same graph to be of different *importance* we introduce a concept of *weight*. We assume a real-valued function $w : E \to \mathbf{R}$ defined for every edge $e \in E$. This *weight function* assigns a weight $w(e)$ to every edge of the graph.

A Problem of Finding a Minimum Spanning Tree

We formulate the task of finding a MST and describe the associated algorithm based on Matousek & Nesetril (2008), pp.171-174 as follows:

Given a connected graph $G = (V, E)$ with a nonnegative weight function w on the edges, find a spanning tree $T = (V, E')$ of the graph G such that the sum of edges $w(E') = \sum_{e \in E'} w(e)$ has the minimum possible value.

There exist several algorithms, that can be used to find a MST of a graph. These include *Kruskal's*, *Jarnik's* (also called *Prim's*) or *Boruvka's* algorithms. In accordance with Mantegna (1999) and Kristoufek *et al.* (2012a) we are going to employ the **Kruskal's algorithm** introduced by Kruskal (1956).

The algorithm begins with a connected graph $G = (V, E)$ with weight function w defined on its edges. Having ordered the edges according to their weight in a nondecreasing order, the algorithm marks the edge with the minimum weight. Should there be more than one edge of minimum weight, the algorithm selects one of them randomly. The algorithm keeps choosing the edges of minimum weight among the so far unmarked edges. The next edge can only be chosen if it does not create a loop. This step-by-step process continues until the selected edges connect all vertices and form a MST of G. Computations associated with the construction of the taxonomy objects will be processed in R software. MSTs and HTs will be visualized using the igraph package.

3.3 Hierarchical Tree

The methodology necessary for construction of HT further elaborates on the toolbox of MST. This section is primarily based on Mantegna (1999) and Mantegna & Stanley (2000) from where the methodology of HTs was sourced.

First, we need to become familiar with the concept of *ultrametric distance*. In addition to satisfying the first two properties of the Euclidian metric distance, the usual triangular inequality changes now into an *ultrametric inequality*, given in (iii), which is even stronger. Triangular inequality is implied by current ultrametric inequality.

(i) *Identity:* $d_{ij} = 0 \iff i = j, \forall i, j \in N$,

(ii) *Symmetry*: $d_{ij} = d_{ji}, \forall i, j \in N$,

(iii) *Ultrametric inequality:* $d_{ij} \leq \max\{d_{ik}, d_{kj}\}, \forall i, j \in N$.

Second, we say that an *ultrametric space* is such a space in which the distance between objects is given by the ultrametric distances. We learn that: "*Ultrametric spaces provide a natural way to describe hierarchically structured complex systems, since the concept of ultrametricity is directly connected to the concept of hierarchy*" (Mantegna & Stanley 2000, p.107).

Third, among all available ultrametric distances we identify the longest one, called *subdominant ultrametric distance*, d^*_{ij}. In an arbitrary MST, this distance corresponds to the longest edge one comes across on the way from vertex i to vertex j. Formally, this is written as:

$$d^*_{ij} = \max(d_{kl}), \tag{3.5}$$

where k and l represent all vertices of a MST along the way from i to j, including i and j.

Fourth, there are $n-1$ links connecting n vertices in every MST. Therefore there exist a maximum of $n-1$ subdominant ultrametric distances associated with a particular MST. All values of d^*_{ij} in a given MST constitute the *subdominant ultrametric distance matrix*, $\mathbb{D}^*$.

Inspecting $\mathbb{D}^*$, the first pair of HT originates by linking the two items connected by the minimal d^*_{ij}. With increasing mutual distance, we keep matching the other items, clusters with equal d^*_{ij} get connected together. Proceeding along these lines, we eventually arrive at the final form of HT.

3.4 Stability of Links

Kristoufek *et al.* (2012a) pointed out the major weakness of described taxonomy methodology, an issue of potential link instability. When inspecting the taxonomy objects, one cannot be certain whether the observed links are relevant for the network or have appeared merely coincidentally. In order to assess stability (importance) of individual links, we employ the bootstrapping technique presented by Tumminello *et al.* (2007).

Basically, we are interested to find out whether the established links remain and prove to be stable as the procedure is repeated with reordered samples. Once the MST has been constructed, we take the underlying time series and create its bootstrapped version. Although its length stays fixed (studied period remains unchanged), its items get randomly reordered while allowing for repetitions[3]. A new MST is then constructed based on this bootstrapped time series and its links are recorded. Repeating this procedure 1,000 times, for each edge we end up with a number indicating how many times out of a thousand repetitions that particular link appeared in MST. Resulting *bootstrap values* are reported for each edge in form of a ratio b_{ij} that divides the number of occurrences by the total of bootstrapped realizations, thus $b_{ij} \in [0; 1]$.

3.5 Wavelet Coherence

Wavelet coherence represents an analytical framework that is not technically related to that of MST or HT. It is a model-free approach that allows for exploring the relationship between two time series. Specifically, we will use this tool to further study the connections identified through the taxonomy perspective. Without imposing any prior assumptions, wavelet coherence enables us to study correlation between two series both in time and across frequencies.

In the rest of this chapter, we are going to briefly introduce the wavelet framework. We aim to deliver a guideline instructing readers on how to understand the concept and use it as a tool. This section is primarily based on Grinsted *et al.* (2004). We also permanently refer to Vacha *et al.* (2013); Vacha & Barunik (2012), and Kristoufek *et al.* (2015) who pioneered the use of wavelets in the context of biofuels.

Traditionally, financial series are studied from the time perspective. In fact,

[3]Thus, some observations may be included repeatedly in bootstrapped series while others may be missing completely.

a time series can also be understood as a signal having several components of different properties. Generally, a signal may be composed of individual waves cycling with different speed, i.e. with different frequencies. Individual components of the signal get separated in the frequency domain. A central feature of wavelet analysis is that it captures both time and frequency domains.

A *wavelet* $\psi_{u,s}(t)$ is a real or complex-valued function given as

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right), \tag{3.6}$$

with a scale (dilation) parameter s and a location (translation) parameter u. Under certain conditions, in detail discussed by Daubechies (2004), the original series $\{x_t\}$ can be fully reconstructed from its *wavelet transform* $W_x(u,s)$

$$W_x(u,s) = \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{s}}\psi^*\left(\frac{t-u}{s}\right)\mathrm{d}t, \tag{3.7}$$

where * stands for a complex conjugate operator preventing an information loss in the transformation. The degree of similarity between $\{x_t\}$ and the shape of wavelet is measured by the integral above. The studied event is further described by parameters u and s. While u specifies its location in time, s indicates its time length. Most often, economic applications of wavelet analysis use Morlet wavelet since it enables studying multivariate relationships between series. To ensure comparability with previous research of Kristoufek *et al.* (2015), we employ Morlet wavelet with central frequency of $\omega_o = 6$ defined as

$$\psi(t) = \pi^{-1/4}e^{t(6i-t)/2}. \tag{3.8}$$

Our analysis is going to focus on relationships between pairs of time series. Under a bivariate setting, the *cross wavelet spectrum* is given by

$$W_{xy}(u,s) = W_x(u,s)W_y^*(u,s), \tag{3.9}$$

where $W_{xy}(u,s)$ stands for the continuous wavelet transform of series $\{x_t\}$ and $W_y^*(u,s)$ marks a complex conjugate of continuous wavelet transform (Torrence & Compo 1998). Since the cross wavelet spectrum is complex, the cross wavelet power is given by $|W_{xy}(u,s)|$. It is usually understood to be a measure of local covariance between two series at a given frequency. Nonetheless, we cannot easily assess the strength of a detected co-movement since the cross-wavelet power is not bounded. Thus we introduce *squared wavelet coherence* defined as

$$R^2_{xy}(u,s) = \frac{|S\left(\frac{1}{s}W_{xy}(u,s)\right)|^2}{S\left(\frac{1}{s}|W_x(u,s)|^2\right) S\left(\frac{1}{s}|W_y(u,s)|^2\right)}, \tag{3.10}$$

with S being a smoothing operator (Torrence & Webster 1998). By definition, the value of squared coherence varies between 0 and 1. Moreover, squared wavelet coherence corresponds to the usual squared correlation coefficient for a specific time and frequency. As the cross-wavelet spectrum translates into a the squared coherence, the information about the direction of the relationship is lost. Therefore, we need to study the *phase difference* specified as

$$\varphi_{xy}(u,s) = \tan^{-1}\left(\frac{\Im\left[S\left(\frac{1}{s}W_{xy}(u,s)\right)\right]}{\Re\left[S\left(\frac{1}{s}W_{xy}(u,s)\right)\right]}\right), \tag{3.11}$$

with $\Re$ and $\Im$ representing a real and imaginary part, respectively. Furthermore, we test statistical significance of the coherence using the Monte Carlo simulation method. For technical details please refer to Grinsted *et al.* (2004).

As pointed out by Kristoufek *et al.* (2015), wavelet coherence is limited by the same technical constraint as usual correlation. It may suffer from the omitted variable bias since it does not control for a possible influence of other variables. Thus, we may observe a (seemingly) high coherence between two price series; however, the observed relationship can in fact be caused by their mutual ties to a third variable. To overcome this issue, we follow Kristoufek *et al.* (2015) in using *partial wavelet squared coherence*, an analogy of partial correlation defined as

$$RP^2_{y,x_1,x_2} = \frac{|R_{yx_1} - R_{yx_2}R^*_{yx1}|^2}{\left(1-R^2_{yx_2}\right)^2\left(1-R^2_{x_2x_1}\right)^2}. \tag{3.12}$$

Partial wavelet coherence evaluates the relationship between $\{y\}$ and $\{x_1\}$ while controlling for the effect of $\{x_2\}$, please see Mihanovic *et al.* (2009) for details.

An outcome of wavelet analysis is presented in a two-dimensional chart. The way of interpreting these charts is demonstrated in Chapter 6 where the results for both the wavelet squared coherence and the partial wavelet coherence are delivered. The whole computational process of wavelet analysis was processed in MATLAB R2014b (version 8.4) using packages by Aslak Grinsted[4] and E. K. W. Ng and T. W. Kwok[5].

[4]Wavelet coherence package was provided by A. Grinsted.

[5]The software for the partial wavelet coherence was provided by E. K. W. Ng and T. W. Kwok and is available at: cityu.edu.

Chapter 4

Data

In this chapter, we provide a description of the dataset employed in our study. Together with motivating the data choice we introduce the sources used. We further examine statistical properties of the dataset and describe the adjustments necessary for our empirical approach. This unique dataset was gathered entirely for the purpose of this thesis and belongs to its core contributions.

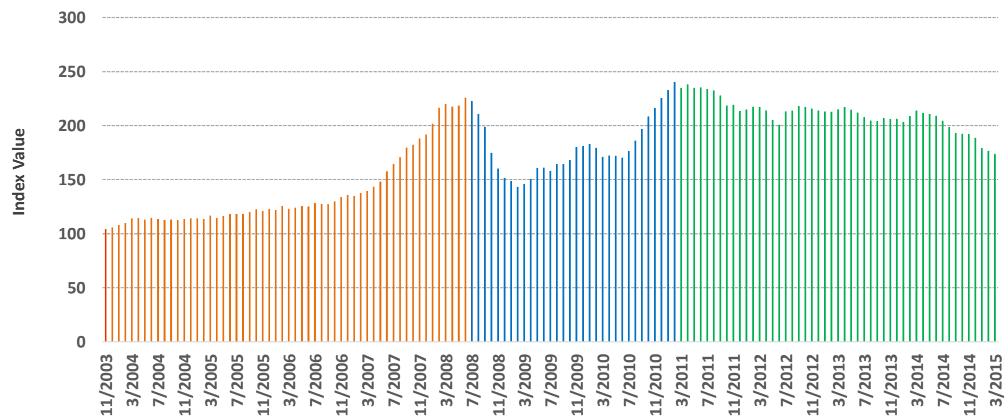
4.1 Dataset Description

In search for a system of commodities and assets that are related to biofuels, we decided to choose a comprehensive approach. Our dataset was gathered from various sources in order to systematically include representative items of the following asset classes: *biofuels* (both ethanol fuel and biodiesel), *ethanol feedstock*, *biodiesel feedstock*, *fossil fuels* including the crude oil, *food*, *stock indices*, *exchange rates*, and *interest rates*. In contrast with previous studies we intend to substantially increase the number of items considered in the analysis. Our attempt results in a dataset composed of 32 price series.

Throughout the thesis we analyze time series data. Each series carries an information about price development of a specific item over time. We work with weekly price data covering an 11 year period starting from November 24, 2003 until January 19, 2015. Some of the data was available even on higher frequency, e.g. daily data from stock exchanges. In such a case, we kept using the prices observed for Mondays. For exchange traded commodities we used Monday closing prices. Occasionally, when a particular exchange was closed on Monday, e.g. due to a bank holiday, we included the closing price recorded on the previous business day.

The time period we are about to analyze is rather long. Over those 11 years, commodity markets went through a number of structural changes. Resulting price development exhibits several different patterns. In order to account for various market environments we decided not to analyse the period as a whole. Instead, we divided it into individual subperiods. For this purpose, the Food Price Index was used. The index published by Food and Agricultural Organization of the United Nations (FAO) provides a broader notion about the development of agricultural commodity prices. The FAO Food Price Index (measured in points) captures the monthly change in international prices of food commodities. It is a weighted average of the five commodity group price indices. Weights are represented by the average export shares of each of the groups during 2002-2004[1]. Figure 4.1 shows the development of the index over the 11/2003–03/2015 period. One can observe an upward sloping trend resulting in the 2007-2008 world food price crisis with the index value peaking in June 2008. Subsequently, the agricultural prices fell bottoming in September 2009 before taking up again for a new food commodity rally. The index reached its new peak in February 2011 without attacking it since then.

Figure 4.1: Food Price Index



Data Source: Food and Agricultural Organization of the United Nations

Author's Layout

Having inspected the index's historical behaviour, we assigned the two aforementioned index peaks to be our dividing points. This way we obtained three subperiods of unequal lengths as depicted in Figure 4.1 in color.

[1]Detailed description together with the underlying data can be found under: www.fao.org

Hence, our analysis will be carried out with the following time perspective:

I. Period: November 24, 2003 — June 30, 2008; 241 weekly observations

II. Period: July 7, 2008 — February 28, 2011; 139 weekly observations

III. Period: March 7, 2011 — January 19, 2015; 203 weekly observations

In the taxonomy analysis we study three subperiods as we hope to distinguish between development stages of the biofuel industry. Moreover, we found several structural changes which are likely to have affected the connections within the price system. The US Energy Policy Act of 2005 or the EU's Renewables Directive 2009/28/EC may serve as an example[2]. Our dividing points exactly correspond to those the World Bank identified as terminal points of the two global food price crises in 2008 and 2011. As summarized by Cuesta *et al.* (2014) for this purpose, the World Bank developed a methodological approach to identify a situation resulting into a potential food price crisis.

Having commented on the time perspective of our analysis, we are going to provide a detailed description of the commodities and assets employed. The items are grouped according to their specific type. To ensure a quick orientation in resulting taxonomy objects, individual groups are graphically differentiated by colors. Table A.1 in Appendix A provides a summary of data and sources.

1. Biofuels

Since our primary focus is on the first generation biofuels, their inclusion is clearly justified. Although the prices are not volatile enough on daily basis, studied biofuel markets exhibit sufficient liquidity to be analyzed on weekly frequency. In MST structure, biofuels will be colored in green.

- **US and Brazilian Ethanol**

  A majority of the world's ethanol is produced in the USA, followed by Brazil. Therefore, we include prices of both US ethanol and Brazilian ethanol represented by the New York Harbor Price and Centro de Estudos Avancados em Economica Aplicada (CEPEA) Ethanol Index, respectively. New York Harbor price is a spot price Free on Board (FOB) quoted in US cents per gallon. The data was obtained from Bloomberg database under the ticker ETHNNYPR Index. It is a denaturated anhydrous fuel

[2] The legislations set mandates for blending biofuels into US and EU fuels, respectively.

ethanol purposed for blending with gasoline as disclosed by Kristoufek *et al.* (2012a). Brazilian ethanol price is reported by CEPEA for anhydrous fuel ethanol. The data was downloaded from CEPEA's website.

- **Biodiesel**

  As opposed to ethanol which leads the biofuels production in the USA and Brazil, biodiesel is primarily produced in Europe. Biodiesel stands for the most important biofuel in the European Union accounting for about 80% of its biofuel used in transportation. Therefore, we searched for suitable data which would represent the price of European biodiesel. At this point, we had to cope with a severe data unavailability. The variety of suitable European biodiesel tickers available from a standard Bloobler terminal is very limited. Many of the time series are too short to be meaningfully employed. We finally solved the issue by including the data from two different sources. Spot price of German consumer biodiesel (sourced from the Bloomberg terminal as BIOCEUGE ATPU FOL Index) is used for periods I. and II. Period III. is then covered by Dutch biodiesel data which was gathered and provided by Reuters. Specifically, we employ a price series labeled as FAME 0 FOB ARA Spot. The label corresponds to FOB spot price of Fatty Acid Methyl Ester traded OTC in harbors of Amsterdam, Rotterdam, and Antwerp. Being quoted in USD per metric ton, this type of biodiesel conforms to EN 14214 norm set at 0°C with a maximum water content of 350 ppm.

2. Ethanol Feedstock

Ethanol is produced from crops that are rich in sugars. Most of the world's ethanol is obtained from corn (in the USA), followed by sugarcane (in Brazil). Other frequent ethanol feedstock include wheat and sugar beets. Still other agricultural commodities may be technically used to produce ethanol, for example cassava, potatoes, cotton or sorghum. In case of the US, we consider the following American commodities: corn, wheat, sugarcane, and sugar beets since they account for a vast majority of American ethanol production. The data represents USD prices and come from Bloomberg as specified in Table A.1. In Brazil we include the price of Brazilian sugar to proxy for local sugarcane price. The data is provided by CEPEA[3]. In MST structure, ethanol feedstock will be represented by red color.

[3] Price series on Brazilian sugar may be downloaded from CEPEA's website.

3. Biodiesel Feedstock

As mentioned above, biodiesel represents the primary European biofuel. The EU is the world's biggest producer of biodiesel. Nonetheless, the global volume of produced biodiesel is substantially smaller than the the world's ethanol production. Technically, biodiesel may be produced from a variety of both vegetable oils and biolipids. However, rapeseed and soybean oils are the most frequent feedstock commodities. In addition to rapeseed and soybean oils, we include also sunflower[4] and palm oil in our dataset. The data was obtained from Bloomberg as specified in Table A.1. In MST, biodiesel feedstock will be visualized in pink color.

4. Fossil Fuels

Biofuels represent an alternative to traditional fossil fuels that are their substitutes. Our dataset thus contains crude oil price. Crude oil is not only the main input into the other fuels' production, but it also stands for a very lively traded commodity. Since our focus is on ethanol and biodiesel we include those fossil fuels that compete with our biofuels from a local perspective. Thus, German diesel and German gasoline are considered because of their competitive relation to European biodiesel. Similarly, we include US gasoline and US diesel as well as Brazilian gasoline and Brazilian diesel to serve as a counter party for US and Brazilian ethanol, respectively. The price of (Brent) crude oil comes from Bloomberg. Details are disclosed in Table A.1. Retail prices of both US/German gasoline and US/German diesel were obtained from the website of U.S. Energy Information Administration (EIA). Prices were quoted in USD per gallon, excluding taxes. For the Brazilian fuel prices we referred to the National Agency of Petroleum, Natural Gas and Biofuels[5]. We employ the weekly weighted average consumer prices for gasoline and diesel which we previously converted to US dollar prices per gallon. In MST structure, fossil fuels will be differenced by gray color.

[4]Due to data unavailability, sunflower seeds (Bloomberg ticker SU1) are used instead of sunflower oil.

[5]Agencia Nacional do Petroleo, Gas Natural e Biocombustiveis - ANP. Price series quoted in Brazilian real per liter are available from www.anp.gov.br

5. Food

In addition to agricultural commodities that are used as biofuel feedstock both in Americas and in Europe, we cover selected purely food commodities as well. Our dataset contains coffee, cocoa, rice, and oranges. First, these commodities cannot be used to produce biofuels. Second, all of them are frequently traded agricultural products that compete with biofuel feedstock over the cultivated land. Food commodity prices come from Bloomberg with details provided in Table A.1. Food commodities will be visualized in purple.

A comprehensive overview of non-energy commodities by Savaşsçin (2011) serves as a good inspiration when sourcing the data for similar purposes. As proposed by Serra & Zilberman (2013), an analysis investigating potential price links between biofuels and other commodities should not omit external factors that might affect price links within the food–energy system. These factors include price development of stocks or futures, policy regulations, and macroeconomic conditions, e.g. exchange rates or interest rates. Similarly, Kristoufek *et al.* (2012a) recommend to extend the taxonomy analysis not only in terms of goods or commodities but also by inclusion of assets such as stocks, exchange rates or interest rates. These recommendations motivated our decision to experimentally increase the complexity of our price system.

6. Stock Indices

Our data set contains a group of five frequently quoted stock indices. A national stock index may serve as a proxy of GDP reflecting the atmosphere in a particular market at a given point in time. The choice of stock indices is suitable because usual GDP data is not available at a weekly frequency. Stock indices can provide information about the overall state of economic performance. We cover the major indices that geographically correspond to the markets of our interest. Namely, we include Dow Jones Industrial Average and S&P 500 to represent the US stock market, Financial Times Stock Exchange 100 Index (FTSE 100) and Deutscher Aktienindex (DAX) to account for the British and German stock markets, respectively. Moreover, due to our interest in Brazilian ethanol, we also include the Brazilian Bolsa de Valores do Estado de Sao Paulo (Bovespa) index. Market data for all the indices was obtained from Bloomberg platform as summarized in Table A.1. The group of stock indices will be differenced by orange color.

7. Interest Rates

From a general perspective, interest rates reflect the nature of macroeconomic conditions. In our case though, there arises an issue with data frequency since a lot of key interest rates are not set on a weekly basis. However, we chose two interest rates that are set daily. US Federal funds rate represents the base interest rate of the US Federal Reserve. Fed funds is the interest rate one bank uses for the overnight lending to another bank and results from the open market. The data was downloaded from the Fed's website[6].

London Interbank Offered Rate (LIBOR)–former BBA LIBOR–is now quoted daily by the Intercontinental Exchange. It serves as a global benchmark for short term interest rates. Out of the variety of currencies and borrowing periods we take 3 months USD LIBOR as it is supposed to be the most frequent one. The data was obtained from the ECONSTATS's website[7]. Interest rates will be visualized in blue.

8. Exchange Rates

Our analysis focuses on three geographical markets where a majority of the world's biofuels is produced; the USA, the EU, and Brazil. For this reason, we consider the USD/EUR and USD/BRL exchange rates. As stressed by Algieri (2014), international food (as well as other commodity) prices are denominated in US dollars. However, since consumers pay for commodities in their local currency, changes in dollar exchange rate affect supply and demand which translates in price changes. We may observe that a strengthening dollar means falling commodity prices. Historical data on USD/EUR was gathered from the European Central Bank[8]. The USD/BRL rate was obtained from the US Federal Reserve web page[9]. Exchange rates are going to be depicted in yellow color.

[6]Federal funds rate available under www.federalreserve.gov/
[7]ECONSTATS data available under www.econstats.com/
[8]Data downloadable from sdw.ecb.europa.eu/
[9]Data can be downloaded from www.federalreserve.gov/

4.2 Descriptive Statistics

Table A.2 delivers basic descriptive statistics of our dataset. For the sake of simplicity, the term *price* will be also used in a natural reference to stock index values, interest rates or exchange rates throughout this paper. Figure 4.2 presents the development of selected weekly logarithmic USD prices. In the price chart (a), we can observe that both US and Brazilian ethanols follow a similar path. To ensure legibility of the figure, crude oil price was plotted together with biofuels instead of fossil fuels. All fossil fuels in (b) exhibit a strong co-movement. However, both Brazilian fuels still stand a bit aside showing a relatively milder price development during the great recession. Ethanol feedstock prices are plotted in two separate graphs. In (c), corn exhibits a very strong co-movement with wheat. Similarly, there is a natural similarity between price movements of sugar beets and sugarcane in (d). Vegetable oils which are feedstock for biodiesel all follow a similar path at different price levels (e). Furthermore, the development stock indices is captured in chart (f).

For the purpose of our analysis, we will convert our price series data P_t into logarithmic returns r_t defined as:

$$r_t = \log(P_t) - \log(P_{t-1}) = \log\left(\frac{P_t}{P_{t-1}}\right) \tag{4.1}$$

The use of logarithmic returns is suitable due to their symmetry as discussed by Hudson & Gregoriou (2010). We also benefit from an earlier analysis by Kristoufek *et al.* (2012a) who chose the same transformation. Thus, we ensure comparability of results. Moreover, the use of logarithmic returns instead of simple prices is beneficial for a technical reason. As will be further explained bellow, when transforming prices into returns, we technically obtain first differences. This fact turns out to be crucial for the discussion of time series stationarity.

4.3 Stationarity Tests

Our analysis will process a construction of MST and HT using a distance metric. Strictly speaking, our distance metric d_{ij} is merely a transformed correlation coefficient. As we are going to compute correlations, stationarity plays a vital role. Thus, we need to check for stationarity of our time series. For this purpose, we will employ the following two tests.

Augmented Dickey-Fuller test

Augmented Dickey-Fuller test (ADF) is an augmented version of an original test introduced by Dickey & Fuller (1979) which tests for the presence of a unit root. To establish stationarity, we aim to reject the null hypothesis of a present unit root. Observed ADF statistic will always be a negative number. The more negative statistic we obtain the stronger is our ability to reject the null.

Kwiatkowski–Phillips–Schmidt–Shin test

Kwiatkowski–Phillips–Schmidt–Shin test (KPSS) was developed by Kwiatkowski *et al.* (1992). It tests a null hypothesis of stationarity against an alternative of a unit root. At this point, we wish not to reject the null hypothesis in order to establish the stationarity of a series.

An attempt to run the stationarity tests on the price series before transforming them into logarithmic returns results in an undesirable outcome. Both ADF and KPSS tests yield non-stationarity for all but one logarithmic price series. Table A.3 in Appendix A summarizes the outcomes of performed stationarity tests. The only exception where stationarity seems not to get rejected on reasonable p-values is US gasoline under the ADF test. However, even this series does not prove to be stationary under both tests. Therefore, we have to state that all of our logarithmic price series are non-stationary.

Further adjustment is needed to achieve stationarity in order to be able to compute correlations. This urge for a technical adjustment constitutes another justification for the use of returns (first differences) instead of simple prices. Both stationarity tests performed on the series of logarithmic returns (582 observations) yield straight and satisfactory results. All analyzed series turn out to be stationary under both ADF and KPSS tests. Non-stationarity is strongly rejected without any exception. Tests' results can be inspected in Table 4.1.

Table 4.1: Stationarity Tests – Log Returns

	ADF	**p-value**	**KPSS**	**p-value**
Biodiesel	-8.2985	< 0.01	0.1045	> 0.1
US Ethanol	-9.1553	< 0.01	0.0918	> 0.1
BR Ethanol	-8.2474	< 0.01	0.0797	> 0.1
Crude Oil	-5.6530	< 0.01	0.3120	> 0.1
Corn	-6.5972	< 0.01	0.1209	> 0.1
Wheat	-8.0847	< 0.01	0.1036	> 0.1
Sugarcane	-7.4650	< 0.01	0.1691	> 0.1
Sugar Beets	-7.6338	< 0.01	0.2776	> 0.1
Soybeans	-6.7541	< 0.01	0.0845	> 0.1
Sunflower	-7.5821	< 0.01	0.0586	> 0.1
Rapeseed	-6.9994	< 0.01	0.0935	> 0.1
Palm Oil	-6.7438	< 0.01	0.0682	> 0.1
US Gasoline	-7.2226	< 0.01	0.1795	> 0.1
US Diesel	-6.3976	< 0.01	0.2315	> 0.1
DE Gasoline	-6.8158	< 0.01	0.1998	> 0.1
DE Diesel	-6.3632	< 0.01	0.2932	> 0.1
BR Gasoline	-7.0856	< 0.01	0.1641	> 0.1
BR Diesel	-7.3788	< 0.01	0.2072	> 0.1
Coffee	-7.5110	< 0.01	0.1472	> 0.1
Cocoa	-9.0287	< 0.01	0.0636	> 0.1
Rice	-7.3896	< 0.01	0.1202	> 0.1
Oranges	-6.2106	< 0.01	0.2487	> 0.1
Dow Jones	-7.8402	< 0.01	0.1297	> 0.1
S&P 500	-7.5160	< 0.01	0.1452	> 0.1
FTSE 100	-8.2013	< 0.01	0.0774	> 0.1
DAX	-8.2896	< 0.01	0.0791	> 0.1
Bovespa	-7.8316	< 0.01	0.2887	> 0.1
Fed Funds	-7.0904	< 0.01	0.1874	> 0.1
Libor	-5.3280	< 0.01	0.2855	> 0.1
USD/EUR	-8.3592	< 0.01	0.1674	> 0.1
USD/BRL	-6.8069	< 0.01	0.3493	> 0.1

Source: Author's Computation

4.4 Normality Tests

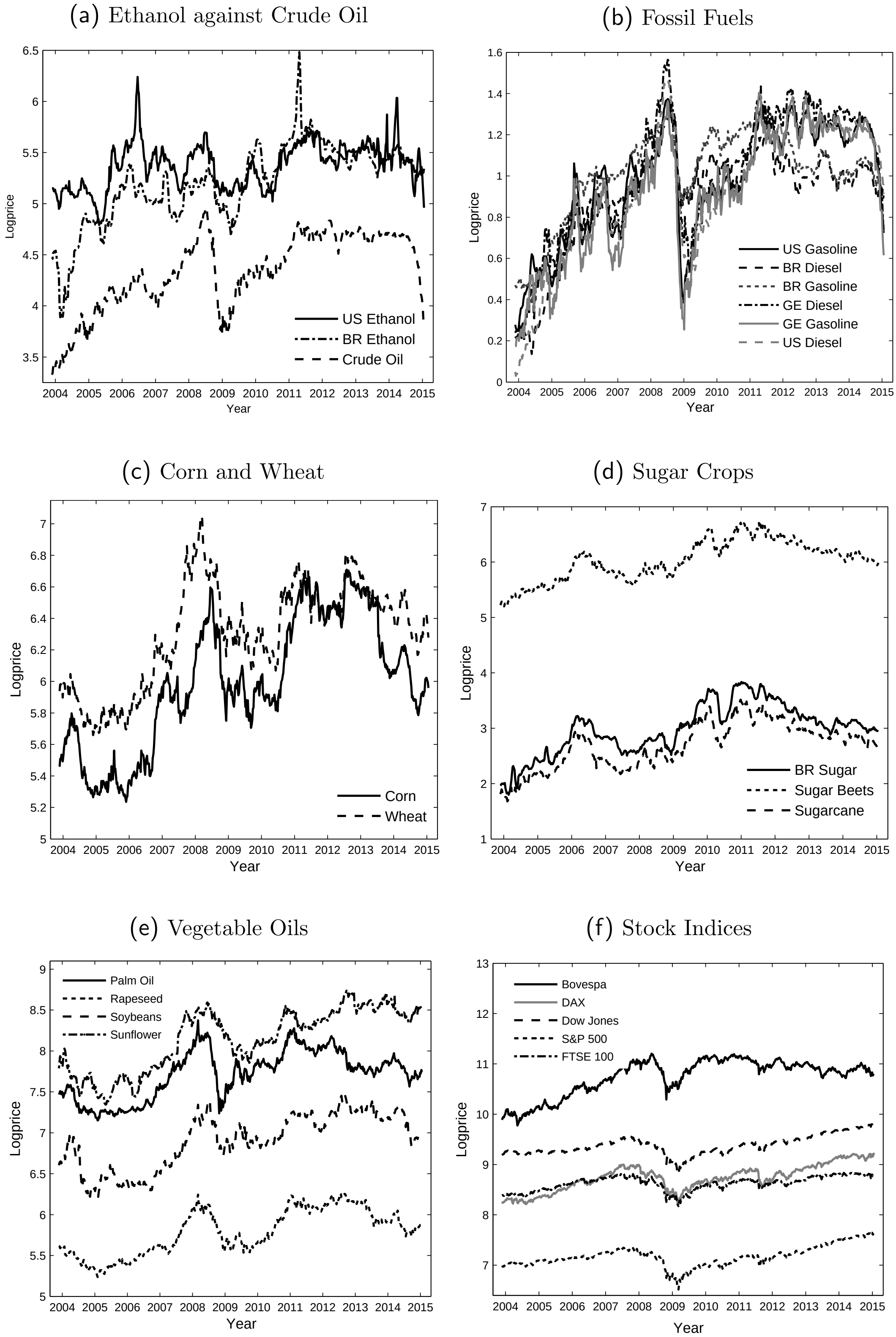
In our normality testing, we combine Shapiro–Wilk test (SW) and Jarque–Bera test (JB). The SW test was introduced in 1965 by Samuel Sanford Shapiro and Martin Wilk. The test is based on an estimated variance and can be used even for relatively small samples (Shapiro & Wilk 1965). In 1987, Carlos Jarque and Anil K. Bera introduced their goodness–of–fit test that checks whether the third and fourth central moments correspond with the normal distribution (Jarque & Bera 1987). Both test share the same null hypothesis of the data coming from the normal distribution. The null is tested against a contradictory alternative.

We performed both test on our logarithmic returns with very similar results. The tests strongly reject the null hypothesis of normality for all of our price series without any exception. Having supplied the tests with a sufficient amount of observations we have to state that our time series do not come from the normal distribution. Detailed results of JB and SW tests together with associated p-values may be inspected in Table A.4 in the Appendix A.

Having said this, the normality assumption being made by Pearson correlation turns out to be very complicated here, although required by standard statistical literature. Usually, the normality assumption has to be disregarded as a frequent issue in social sciences. Earlier, Chance (1986) showed that the distribution of linear correlation coefficient does not depend on either of the underlying data distributions. In the same vein, Good (2010) demonstrates that non-normality of the underlying data does not come at cost of validity or precision of the correlation coefficient. Nonetheless, we need to keep in mind this specific property of our data especially when interpreting the results.

In this section, we provided a detailed description of our dataset including its sources. A summary can be inspected in Table A.1 in Appendix A. We discussed our motivation for the use of individual price series. Due to their revealed statistical properties, we chose to transform simple prices into logarithmic returns. The data is provided on weekly frequency. In order to exploit the highest possible amount of information, we will also use the same data set on monthly frequency. The monthly data is extracted from our weekly series by simply taking one observation every four weeks.

Figure 4.2: Selected Logarithmic Prices



Source: Refer to Table A.1, Author's Layout

Chapter 5

Taxonomy Analysis

Having previously described both the employed methodology and the dataset we are now going to elaborate on the results of our taxonomy analysis. Following a step-by-step approach, this chapter builds and visualizes an interconnected system of biofuel-related commodities and assets. The chapter begins by outlining an experimental system of assets before proceeding to its simplified version that contains physical commodities only. Construction and visualization of MST and HT structures were performed using R software.

5.1 Experimental Price System

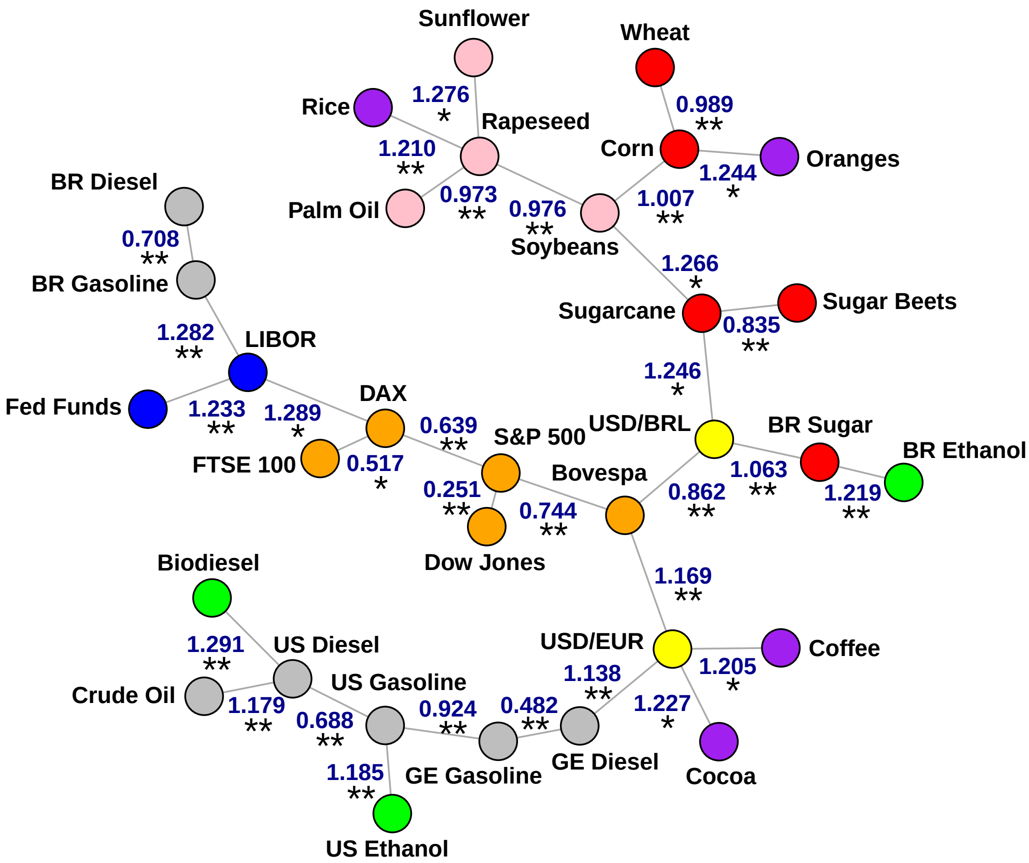
To our knowledge, we are the first to construct a similarly complex system of commodities and assets that are associated with the global production of biofuels. As described in Chapter 4, the complexity of our price system increases due to employing a comprehensive pool of items as well as through covering an exceptionally long period of time. The inclusion of purely food commodities was inspired by a recommendation of Serra & Zilberman (2013). We also took an advice from Kristoufek *et al.* (2012a) who had proposed considering relevant financial series such as exchange rates, interest rates or stocks indices.

5.1.1 Period I, 2003 – 2008

In order to explain a practical use of the taxonomy methodology described in Chapter 3 we start with a description of how a particular MST and an associated HT come to existence. For each period, we employ logarithmic returns on weekly and monthly frequency to differentiate short term and medium term effects.

When constructing a MST, we are interested in visualizing the most *important* connections among the vertices. The *importance* of an edge is determined by the strength of correlation between the two given vertices. Simple correlation is transformed into the distance measure, d_{ij} – the stronger the correlation, the shorter the edge. Actual values of realized distances d_{ij} are indicated by blue bold numbers attached to the edges.

Figure 5.1: MST – Experimental Taxonomy, Period I, Weekly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

Source: Author's Computation

Apart from distances between the vertices, we also test the stability of the links using the bootstrap method. We then inspect how many times out of a thousand bootstrapped realizations a particular connection appeared in the MST. A robust link being present more than 500 times is marked by a double asterisk sign. On the contrary, single asterisk designates a rather unstable link.

Exploring the weekly logarithmic returns, we found the shortest edge ($d_{ij} = 0.251$) between the US stock indices. Hence, Dow Jones and S&P 500 create the first pair a nascent MST. The second shortest link ($d_{ij} = 0.482$) was identified between German gasoline and German diesel. The third strongest correlation exists between the European stock indices, FTSE 100 and DAX, ($d_{ij} = 0.517$). At this point, our MST consists of three separate pairs: Dow Jones–S&P 500, GE gasoline–GE diesel, and FTSE 100–DAX. The next shortest distance ($d_{ij} = 0.639$) is then found between DAX and S&P 500 which are both already present in the MST. By matching these vertices we create a quadruple through connecting the two pairs. The fifth highest correlation exists between US gasoline and US diesel, ($d_{ij} = 0.688$). We add a new separate pair since neither of these fuels has yet been present in the MST. The connection between Brazilian gasoline and Brazilian diesel ($d_{ij} = 0.708$) creates a separate fuel pair. The next shortest edge ($d_{ij} = 0.744$) connects Bovespa to the S&P 500 on the already existing quadruple of the remaining stock indices.

Our MST consists now of a quintuplet of stock indices and three separate (gasoline-diesel) retail fuel pairs. Further steps add a separate sugarcane–sugar beets pair and then connect USD/BRL exchange rate to the Bovespa index. Each time, before a potential new edge is constructed, we need to make sure it will not create an undesirable loop in the MST. Next steps form a fuel quadruple by linking the US and German gasolines ($d_{ij} = 0.924$), before establishing the first vegetable oilpair, rapeseed–palm oil($d_{ij} = 0.973$). Following this logic we eventually obtain a complete MST as shown in Figure 5.1.

A HT is paired with a particular MST for the purpose of classification and visualization of the MST's hierarchical structure. With a form of an inverted tree, the composition of a HT uses a following logic which corresponds to that of MST. Construction of the HT depicted in Figure A.3 begins with matching Dow Jones and S&P 500, the closest pair from the underlying MST in Figure 5.1. Intuitively, next pairs arise from connecting German gasoline to German diesel and DAX to FTSE 100. The fourth pair is formed by the already present DAX and S&P 500. Graphically, we connect the two relevant pairs. Assigned distance ($d_{ij} = 0.639$) applies now to all potential connections among these four vertices. After establishing the US and Brazilian fossil fuel pairs, respectively, Bovespa is connected to the existing cluster of stock indexes. The associated distance ($d_{ij} = 0.744$) is now assigned to all the four possible pairs. Consecutive steps according to this logic result eventually in the complete HT in Figure A.3.

Inspecting Figure 5.1 we observe the visualized characteristics of our experimental network. A core of the MST is formed by a compact group the stock indices. It follows from the interconnected nature of stock markets that the indices are strongly correlated with stable mutual links. We see that not only stock indices but also a number other items gather according to their type. In broader terms, there seems to be a group of agricultural commodities–a food branch– and fuel branch of the MST at the opposite sites of the network. A cluster of vegetable oils together with ethanol feedstock commodities constitute the food part of the tree. On the other hand, fossil and biofuels form the fuel branch. While no other biofuel is connected to its feedstock, Brazilian ethanol makes a notable exception. Its robust link to Brazilian sugar is evident already during the first studied period. US ethanol and even biodiesel have stable links to their US fossil substitutes.

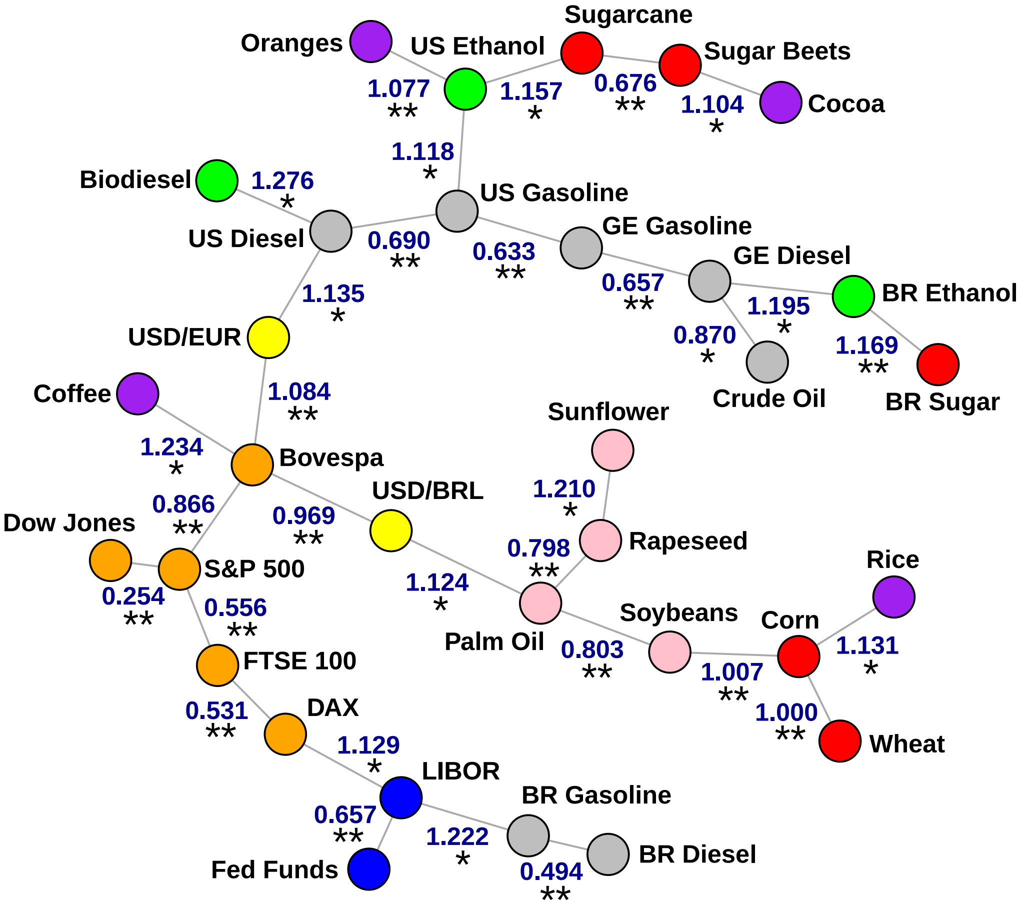
As opposed to the exchange rates which bridge the stock market cluster with fuel and food parts, interest rates do not seem to interact a lot. Brazilian retail fuels are not integrated into the fuel branch, they stand at the edge of the network being only linked to Libor. The isolated position of Brazilian fossil fuels is implied by a specific setting of national fuel market in Brazil. Due to a decisive influence of Petrobras[1] on local fuel prices, Brazilian fossil fuels do not necessarily follow the global markets' development. As we will see during the whole studied period, they do not usually integrate into the fossil fuel cluster. We should not forget about the four purely food commodities which cannot be used to produce biofuels. These items do not form any cluster and are only individually connected to different nodes of the network.

In contrast with the MST on weekly frequency, the MST depicted in Figure 5.2 was based on monthly data. Although reflecting the same period of time these two MSTs differ in some details. We still observe a strongly connected stock market cluster linked together with interest rates and Brazilian fuels. The food branch shrunk somewhat but vegetable oils cluster and cereals remain almost untouched. On the other hand, lower frequency brings interesting changes into the fuel part of the MST. Not only is US ethanol still linked to US gasoline but it also becomes connected to sugars. Biodiesel and Brazilian ethanol keep their links to US diesel and Brazilian sugar, respectively. Specifically, Brazilian ethanol moves with its feedstock to the fuel part of the network getting attached to German diesel. Here we need to point out a natural feature

[1]Petroleo Brasileiro S.A is the Brazilian largest energy corporation with multinational presence. A share of up to 64% is directly or indirectly controlled by the government.

of our taxonomy approach. When interpreting the results, one needs to keep in mind that the longer is the link the weaker is the mutual correlation. In particular, a link of length close to $\sqrt{2}$ effectively refers to no correlation.

Figure 5.2: MST – Experimental Taxonomy, Period I, Monthly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

Source: Author's Computation

A first look separately at HTs depicted in Figure A.3 and Figure A.4 in Appendix A does not immediately tell us an intuitive story. When comparing these two HTs though, the difference becomes more evident. Foremost, we may observe that in the monthly HT realized connections take place on relatively lower levels. Thus, the whole price system gets more closely interconnected on a lower frequency – a sign we do not observe for the last time.

5.1.2 Period II, 2008 – 2011

Our first period covered an era of rising food prices and increasing global significance of biofuels preceding the first world food price crisis. We now continue investigating our network's development in a changed market environment. After several years of accelerating agricultural and energy prices, these slumped quickly during the second half of 2008 hand in hand with a global economic crisis. After the bottom was reached in 2009, both energy and agricultural prices started a new rally until approaching a new peak in February 2011. At this period, we expect to see signs of an established biofuel production not only in Brazil but also in the US and Europe. Figure A.5 delivers a MST generated from weekly data. Compared to the previous period, US ethanol moved to the food branch connecting to corn. The US ethanol–corn link is relatively short ($d_{ij} = 0.930$) and stable. It reflects an important connection between main US biofuel and its primary feedstock.

The lower frequency (Figure A.7) turns out to be suitable since it reveals some essential connections within the network. The MST becomes meaningfully structured and its links get shorter. All three biofuels connect to their feedstock reflecting the underlying production logic. In case of biodiesel, this is observed for the first time. In particular, biodiesel gets attached to rapeseed oil which is its main European feedstock commodity. US ethanol preserves its link to corn and their mutual distance even reduces from $d_{ij} = 0.930$ to $d_{ij} = 0.659$. Notably, Brazilian ethanol and sugars form a relatively isolated cluster reflecting strong ties of Brazilian ethanol production to local sugar prices.

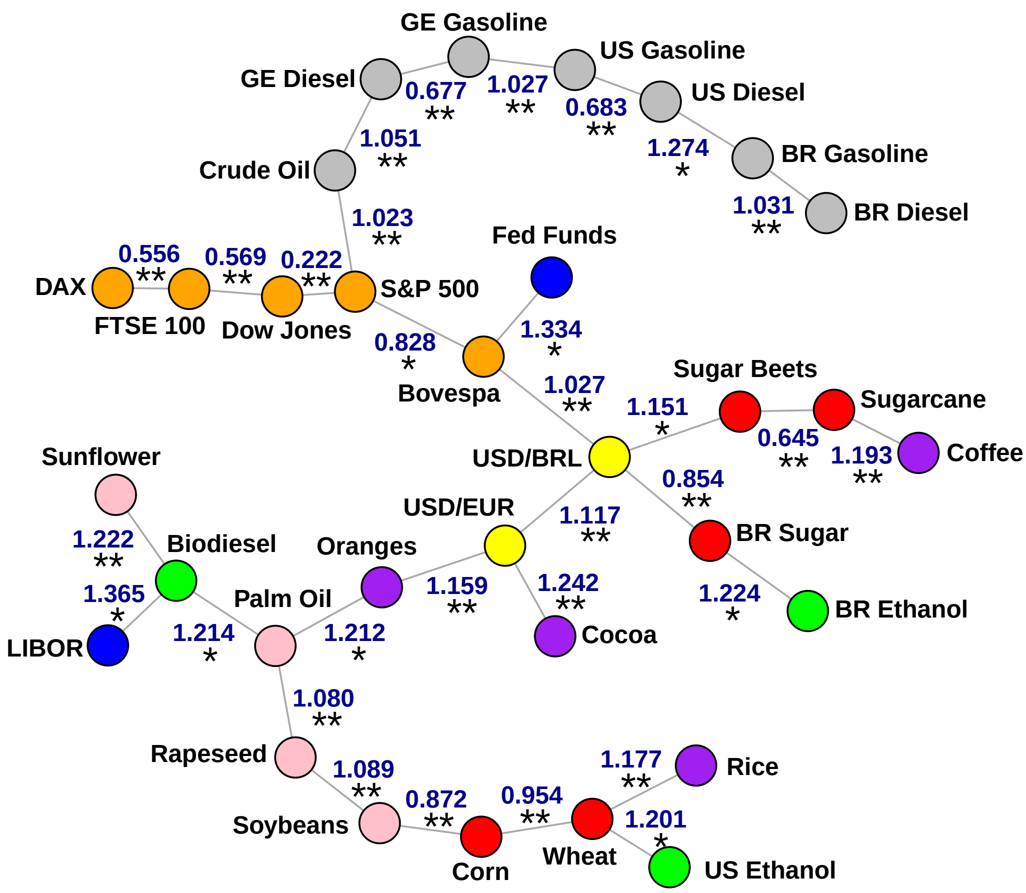
Let us briefly compare the HTs depicted in Figure A.6 and Figure A.8 visualizing weekly and monthly hierarchical structures, respectively. We quickly notice an obvious difference. It becomes evident now that the network gets considerably more interconnected when data frequency is lowered. The realized links get shorter. The monthly HT appears to be much more readable. We notice that a number of links are now established at heights of about $0.6 - 0.8$ rather than previous $1 - 1.2$.

The second period further develops our knowledge of the biofuel-related price system. Especially the taxonomies resulting from the monthly data reflect a well established biofuel production in the US, Brazil, and Europe. We observe all biofuels connected to their respective productions factors. In important move in the European context is biodiesel getting eventually connected to rapeseed.

5.1.3 Period III, 2011 – 2015

During the last studied period, the values of FAO Food Price Index (Figure 4.1) experienced a gradual slowdown continuing until and including January 2015. Considered food and energy prices went through a volatile season during which they approached considerably lower values. At the same time, stock indices grew reaching new all time highs.

Figure 5.3: MST – Experimental Taxonomy, Period III, Weekly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

Source: Author's Computation

Inspecting the weekly MST in Figure 5.3 we may draw a clear line separating the fuel from the food branch with the stock indices just in between. Now the visual difference between fossil fuels and agricultural commodities becomes obvious. Exceptionally, both Brazilian fossil fuels get integrated into the fuel cluster. In accordance with the previous findings, the biofuels stay attached to their respective feedstock clusters; however, their mutual links are rather weak. There exist very stable interconnections between palm oil, rapeseed, and soybeans with biodiesel being attached to both sunflower and palm oil.

Although the monthly MST (Figure A.10) is very similar to the previous weekly MST, there are several structural differences. Foremost, the whole tree seems to be more complicated. Suddenly, it becomes challenging to make a clear cut between the fuel and food regions. Even the food branch itself gets divided into two parts as the sugar cluster moves away with Brazilian ethanol. Biodiesel still remains connected to the vegetable oils cluster. In each of the trees, we have observed a stable link between soybeans and corn. The unstable US ethanol–soybeans connection may look little surprising here. However, since US ethanol is also correlated with wheat and corn it oscillates among them in bootstrapped realizations.

Unlike the previous period, the weekly (Figure A.9) and monthly (Figure A.11) HTs are now somewhat more similar to each other. We again observe that the links get shorter with lower frequency, but the difference is smaller now. In both HTs, we identify the clusters formed by stock indices, US and German fossil fuels, and food commodities grouped around the US ethanol.

Having modeled the experimental price system over the period of 11 years, we have learned about several patterns that emerge from MST and HT structures. Nonetheless, we still wonder whether it is possible to reveal even more dynamics with use of the taxonomy method. After all, one can never be entirely sure their asset selection was optimal. In order to find out, we decided to adjust our price system by reducing the selection of analyzed items. This way, we will be able to compare the outcomes of both systems. During the adjustment procedure, we tried several different settings before eventually arriving at the preset form. Strictly speaking, the final design eliminates four out of the previously employed asset groups. We excluded the group of stock indices, pure food commodities, exchange rates, and interest rates in order to find out whether these are vital for the system.

Looking back at the constructed MSTs, it comes natural that some components are more important for the network than others. The food commodities that cannot be used in biofuel production (i.e. rice, oranges, cocoa, and coffee) have not been effectively integrated into the network. Very often, they got only unstably connected to various vertices at the edge of the tree. These food commodities do not form any cluster. Similarly, the interest rates do not exhibit any reasonably strong position within the network. Judging from the HTs, their links to other items belong among the longest (i.e. weakest). The group of stock indices, on the other hand, does form a very strongly interconnected cluster of an exceptionally stable design. However, this stocks cluster merely separates the food from the fuel branch without actually interacting with either part. It further makes a good sense that the exchange rates are always attached to the stocks cluster. No wonder that Bovespa index (quoted in Brazilian reals) establishes a stable link to USD/BRL exchange rate. Let us now explore what the taxonomy looks like if we exclude the above mentioned four groups. By doing so we allow for including only those commodities that are physically related to the process of biofuels' production and consumption.

5.2 Adjusted Commodity System

Using only physical commodities, the adjusted system aims to provide a more structured view delivered and interpreted in the context of the individual biofuel markets under consideration. All items in the following taxonomy objects play a physical role in either production or sales process of the respective biofuels. We include the biofuels together with their production factors and their fossil fuel alternatives[2]. This way, we hope to increase precision and foster the interpretability of resulting taxonomies.

For the sake of comparison, we keep presenting the MSTs and HTs on both weekly and monthly frequencies while also preserving all the previous technical features of our tree objects. Thanks to a smaller number of analyzed items, the length of the link (d_{ij}) is now supplemented by an exact *bootstrap value* (b_{ij}) indicated in square brackets. The bootstrap value signals how many times out of thousand repeated realizations a particular link appeared in the MST.

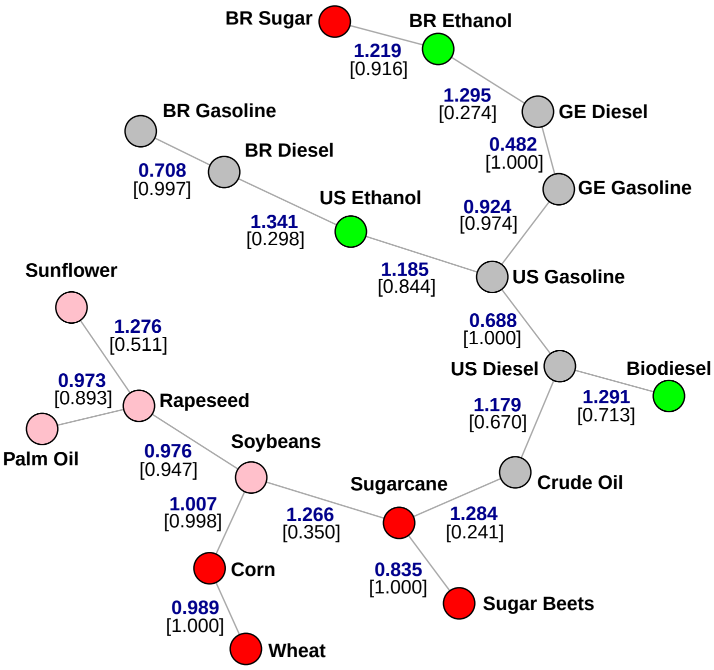
5.2.1 Period I, 2003 – 2008

The MST in Figure 5.4 results from weekly data. Our network is composed of stable food and fuel branches that are relatively well separated in the short term. Individual items within these branches are closely interconnected with very stable links. For instance, in the food region, corn–wheat and sugarcane–sugar beets pairs were found in every bootstrapped case. Similarly, connections established between the pairs of national fossil fuels have also very high stability.

It follows from the design of the present MST that by excluding some of the previous items we actually made no substantial changes to the rest of the network. Compared to the corresponding MST in the experimental system, the structure shrank by 13 vertices. However, realized connections among the remaining items were not actually affected. Put simply, by removing some of the vertices we simplified the system without altering the links existing among the remaining vertices. We can still compare structural composition of whole branches between the experimental and adjusted system with a high accuracy. The gaps resulting from removal of financial and food items are bridged by three new links. Namely, crude oil connects to sugarcane, Brazilian diesel to US ethanol, and Brazilian ethanol to German diesel. However, these edges are very long and represent the most unstable three links within the MST ($b_{ij} < 0.3$).

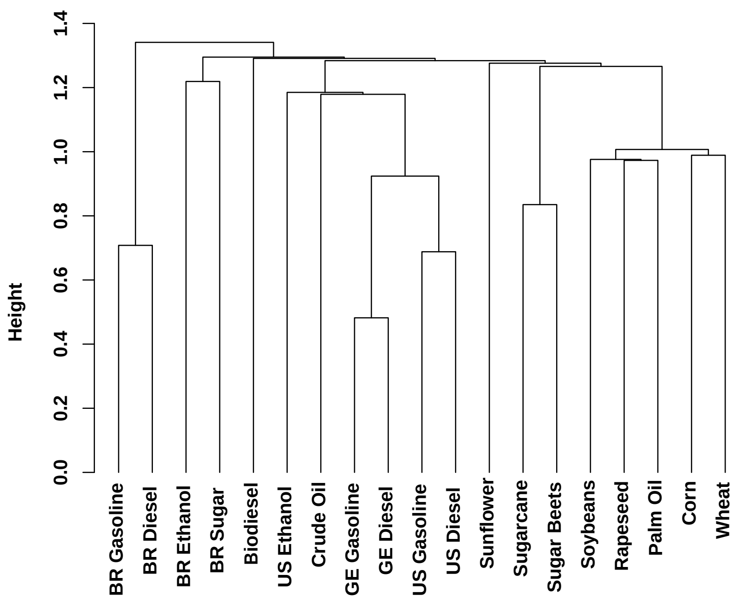
[2]Although not a consumer fuel itself, crude oil represents the main production factor for all fossil fuels irrespective of the geographic market.

Figure 5.4: MST – Adjusted Taxonomy, Period I, Weekly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.5: HT – Adjusted Taxonomy, Period I, Weekly



Source: Author's Computation

All biofuels belong now to the fuel branch. Biodiesel and US ethanol are connected to their US fossil counterparts with weak but stable links (b_{ij} = 0.713 and b_{ij} = 0.844, respectively). Brazilian ethanol is tied to its BR sugar feedstock. Their link is rather long but stable b_{ij} = 0.713. While US and European biofuels connect to fossil fuels and not to their feedstock, Brazilian ethanol behaves differently. It primarily depends on its feedstock while the connection to fossil fuels is unimportant.

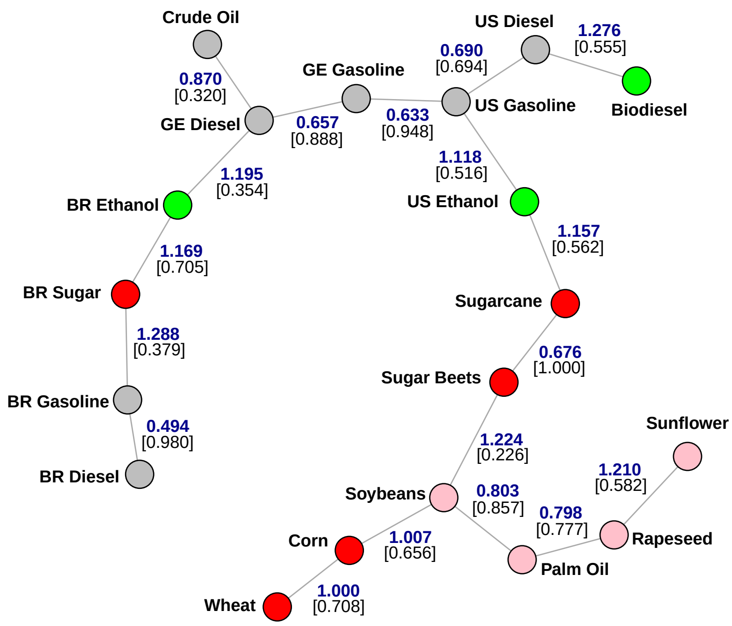
In terms of complexity, HTs generated for the adjusted system will obviously be more plausible to read and to interpret. In Figure 5.5 we observe several clusters. There are two large clusters of fuels and feedstock that correspond to the above mentioned main tree branches. A little apart, Brazilian commodities create two separate pairs. We see that Brazilian fossil fuels are tight together but practically uncorrelated with other items, they connect to the rest of the network at d_{ij} = 1.341. This behavior reflects a heavily regulated Brazilian fuel market with local prices that do not necessarily follow the price of oil.

The MST in Figure 5.6 originated from monthly data. The tree shares a majority of characteristics described for the higher frequency. Its links are highly similar to those observed within the experimental system. In the medium term, they also become stronger than in the short run. US ethanol and biodiesel keep their connections to corresponding US fossil fuels. However, the stability of these fuel–biofuel links decreases considerably. In the medium term, biofuels depend on other commodities as well. Now, US ethanol moves closer to the food branch by connecting to sugarcane. We cannot make any strong judgment as the link is rather weak d_{ij} = 1.157 and only moderately stable b_{ij} = 0.562.

HT in Figure 5.7 shows more details. In particular, we can distinguish five clusters. A big feedstock cluster and a fuel cluster are the main ones. Then there are three separate pairs: sugarcane–sugar beets, BR ethanol–BR sugar, and Brazilian fossil fuels. Apparently, biodiesel stands aside not interacting with other items both in short and medium term. During the first period, we observe different patterns in behavior of Brazilian and the other biofuels.

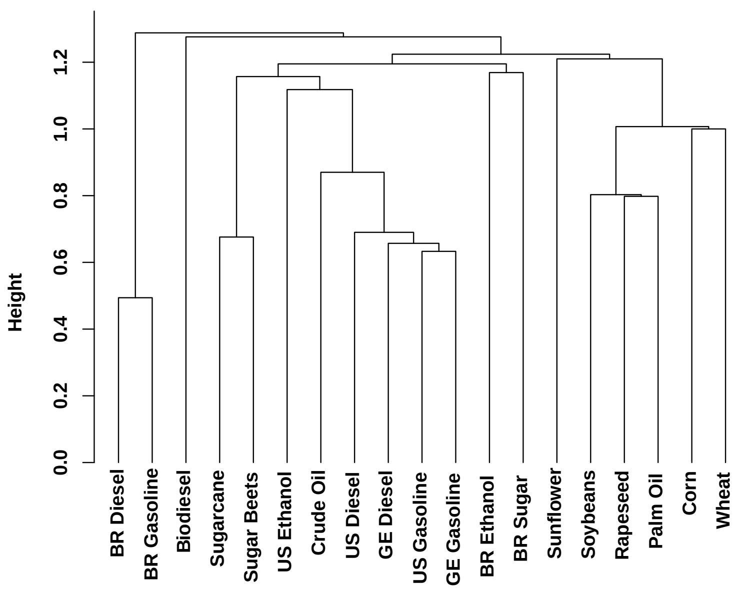
We identify a significant phase shift in development stages of Brazilian and the US/EU biofuel industries. In the US but especially in Europe, this period is associated with a rise of an immature biofuel industry. In Brazil though, the ethanol–sugar connection represents a traditional but still growing ethanol production. In 2000, the US ethanol production stood at 6.1 billion liters while Brazil produced 10.7 billion liters. Five years later, the USA already generated 14.8 billion liters compared to 16 billion liters produced in Brazil.

Figure 5.6: MST – Adjusted Taxonomy, Period I, Monthly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.7: HT – Adjusted Taxonomy, Period I, Monthly



In the global context, Brazil represents a unique scenario. The dawn of Brazilian ethanol industry dates back to the 1970s. At the time of the oil crises, Brazil was covering as much as 80% of its oil demand by imports. A rocketing oil price triggered an enormous Brazilian biofuel effort. Due to the country's large size and favorable geographic location, Brazil enjoys perfect conditions to grow sugarcane, its main ethanol feedstock. Being part of a broader energy security scenario, Brazilian ethanol industry has been massively subsidized. For decades, various forms of governmental support range from blending mandates and infrastructural subsidies to promoted sales of ethanol fueled vehicles. Although the demand for ethanol decreased temporarily after the biofuel market was deregulated in 1990s, Brazil experienced a renewed ethanol boom in early 2000s. Starting from 2003, *flex fuel vehicles*[3] were massively introduced to Brazilian market. A decade later, flex fuel cars account for about a half of the local vehicle fleet and almost 90% of new cars sold in Brazil.

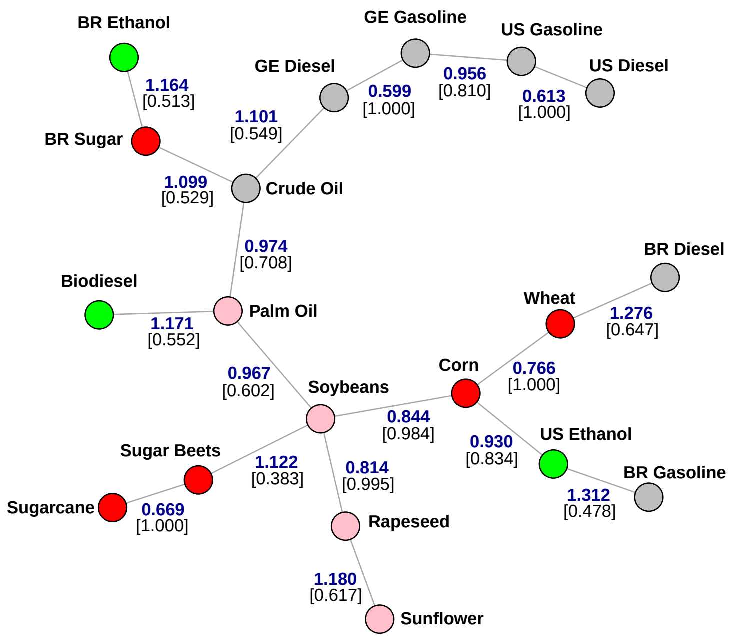
5.2.2 Period II, 2008 – 2011

Our analysis continues through an era of very volatile agricultural prices between the first and second food price crisis. The weekly MST in Figure 5.8 brings several new features along with those we already know. As a matter of course, we again find the BR ethanol–BR sugar link. However, the stability of this link decreases considerably compared to the previous weekly MST. In about 50% of bootstrapped cases, BR ethanol interrupts the connection with its feedstock. At that time, unfavorable weather conditions caused several poor sugarcane harvests. In turn, high sugarcane prices diverted the industrial attention from ethanol to production of sugar. Moreover, artificially low price levels of local fuels contributed to a temporary crisis of Brazilian ethanol industry.

At the same time, we observe notable changes for the US and European biofuels. First, US ethanol moves to the food branch establishing a relatively strong ($d_{ij} = 0.930$) and stable link ($b_{ij} = 0.834$) to its primary US feedstock, corn. Second, European biodiesel connects to the cluster of vegetable oils by attaching to palm oil. Thus, we begin to see signs of the phase shift approaching the US and also more delayed European biofuel markets. During the second period, the short run behavior of all studied biofuels depends on feedstock commodity prices.

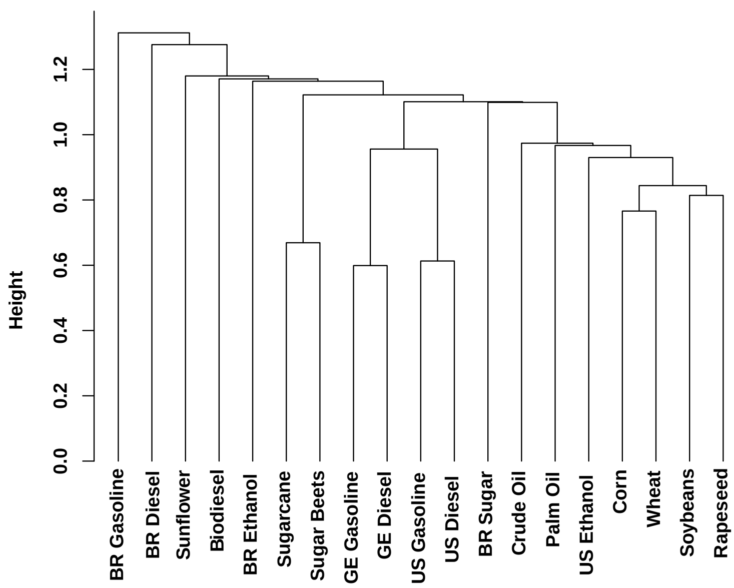
[3]Engine of a flexible-fuel vehicle is capable of running either on gasoline or ethanol or any combination of the two.

Figure 5.8: MST – Adjusted Taxonomy, Period II, Weekly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.9: HT – Adjusted Taxonomy, Period II, Weekly



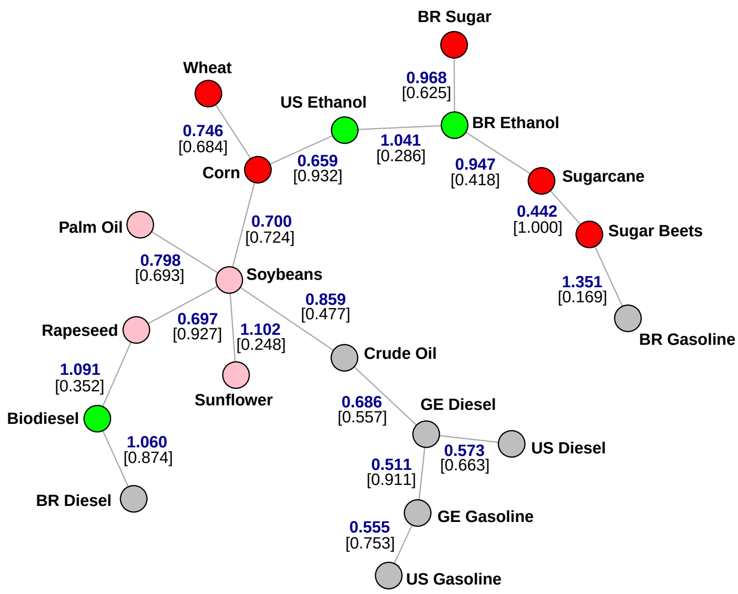
Source: Author's Computation

The associated HT (Figure 5.9) is less structured compared to the previous period. It is dominated by several closely linked pairs formed by vegetable oils, cereals, US and German retail fuels, and sugars. Separation into food and fuel clusters is less clear now. Interestingly, US ethanol moves away from the fuel cluster and even crude oil is quite far from the retail fuels. This finding supports our claim that the US have already developed a mature biofuel industry. Indeed, as of 2010 the US produced 50.3 billion liters of ethanol, an increase by 240% from 2005 levels and far ahead of Brazil's 26 billion liters.

When the data frequency is lowered to one month the resulting MST (Figure 5.10) extends our findings to the medium term. Apparently, the tree gets clearly and logically structured. One can now distinguish several commodity groups. Apart from a typical fossil fuel group, we see a cluster of vegetable oils with soybeans in the position of a central node with five vertices. Biodiesel is now connected to rapeseed, its primary European feedstock. Compared to weekly frequency, biodiesel gets closer to vegetable oils but its connection (to rapeseed) becomes less stable ($b_{ij} = 0.352$) since it is correlated with multiple items from that cluster. Biodiesel is now further linked to Brazilian diesel which is a change from US diesel in the previous period. Starting from scratch, the EU biodiesel production more than quadrupled during 2000-2005, reaching 3.2 megatonnes in 2005. Then, over the course of the next five years, the production volumes of EU biodiesel tripled reaching 9.6 megatonnes in 2010. Also US ethanol firms its ties to corn both in terms of distance and link stability. Their mutual distance decreases considerably from $d_{ij} = 0.930$ to $d_{ij} = 0.659$. At the same time, stability of the link rises from $b_{ij} = 0.834$ to $b_{ij} = 0.932$.

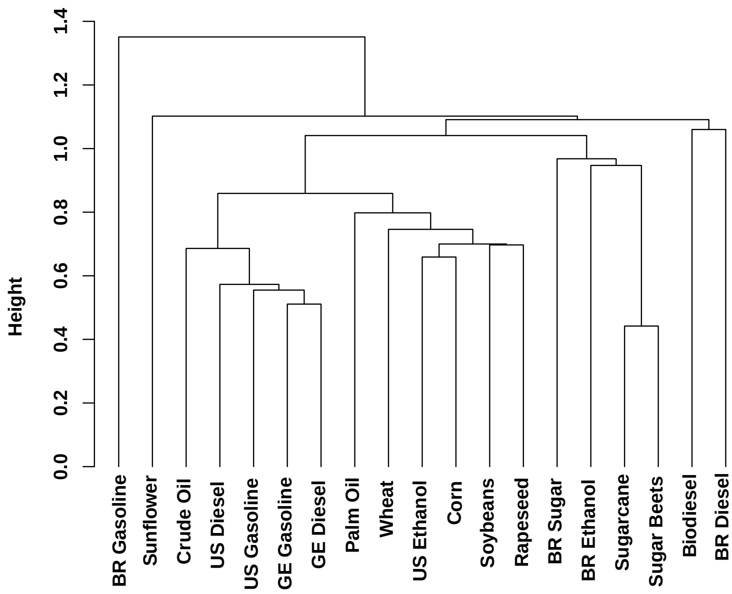
We are able to learn more details from the corresponding HT (Figure 5.11). The overall length of the links shortens as individual correlations strengthen hand in hand with a lower data frequency. We observed this phenomenon already within the experimental system. Hence, we can now argue with more certainty that commodities get more interconnected in the medium term. The HT shows three closely linked clusters which correspond with the structure of the MST. We notice that biodiesel still behaves differently from both US and Brazilian ethanols. In particular, ethanols keep their close ties to feedstock, biodiesel, on the other hand, remains far from the rest of the network. Altogether, the monthly taxonomy is in accordance with our findings for the weekly frequency. During the second period, not only Brazilian ethanol but also the other biofuels lively interact with their production factors in both short and medium terms. Studied biofuel markets have already reached a mature stage.

Figure 5.10: MST – Adjusted Taxonomy, Period II, Monthly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.11: HT – Adjusted Taxonomy, Period II, Monthly



Source: Author's Computation

5.2.3 Period III, 2011 – 2015

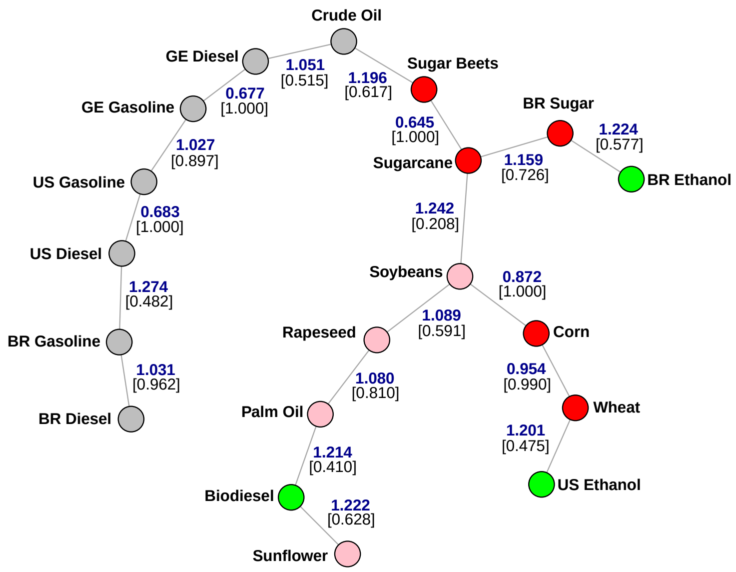
Our analysis continues by exploring the most recent period following after the second food price crisis culminating in Q1 2011. In terms of agricultural commodity prices represented by FAO Food Price Index (Figure 4.1), this period is characterized by a gradual price decline. Except for a few temporary fluctuations, the decreasing tendency ruled the agricultural commodity markets throughout the whole period until January 19, 2015. Crude oil price remained high, mostly above USD 100 per barrel, and relatively stable until a radical drop to as low as USD 47 during the last studied months. This period covers a recent development of biofuel-related prices and reflects thus an established biofuel producing industry in Brazil, the US, and in Europe.

The MST in Figure 5.12 shows a short term (weekly) taxonomy perspective. Looking from fuels towards its food part we notice that crude oil bridges the fossil fuel and food clusters—a feature we have observed in every MST generated from weekly data. According to our expectation, the food part is again structured in a way corresponding with the biofuels production logic. As usual, the BR ethanol–BR sugar pair is present. Moreover, it is comparable to that observed in the second period. The pair is a little more stable but a little less correlated, too. It further connects to the sugarcane–sugar beets pair. The resulting short term sugar cluster is relatively isolated since it is far from both fossil fuels ($d_{ij} = 1.196$) and the rest of the food items ($d_{ij} = 1.242$).

During the second period, US ethanol was attached to corn, the ethanol production volumes and its price were rising quickly. On both weekly and monthly frequencies, this link used to be very stable. Now US ethanol gets connected to wheat with relatively high mutual distance ($d_{ij} = 1.201$) and considerably lower stability ($b_{ij} = 0.475$). Ethanol's short run behavior becomes less tied to the crop price, its price is volatile and follows a decreasing trend. In 2014, US ethanol adds only a 7.8% increase to its 2010 production levels. Compared to the previous weekly MST, biodiesel adds sunflower to its earlier palm oil connection. The geographically pairwise closely interlinked fossil fuel cluster contains now Brazilian fuels as well. However, this unstable ($b_{ij} = 0.482$) link is the longest ($d_{ij} = 1.274$) in the MST and thus not of a high importance.

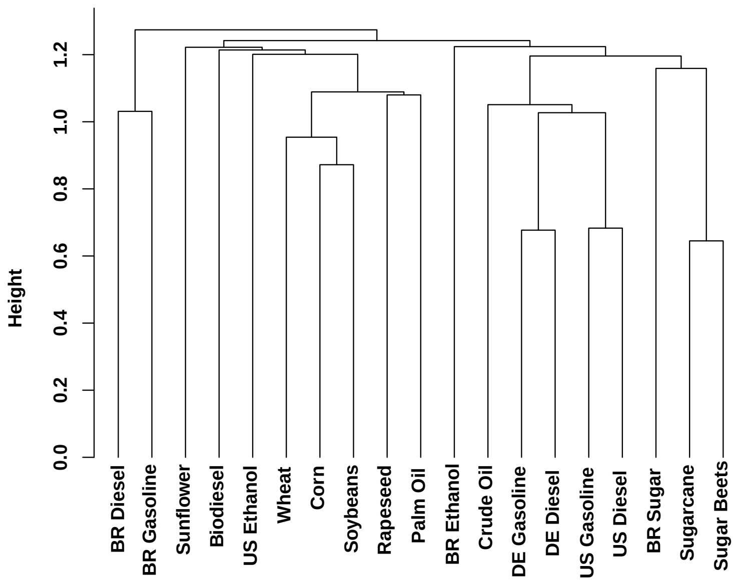
The HT (Figure 5.13) clearly reveals what was already visible from the MST–the edges become longer. Not only get the biofuels less tied to feedstock but the network becomes less interconnected as a whole. In an environment of bearish commodity prices, some of the previously close ties slacken in the short term.

Figure 5.12: lMST – Adjusted Taxonomy, Period III, Weekly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.13: HT – Adjusted Taxonomy, Period III, Weekly



Source: Author's Computation

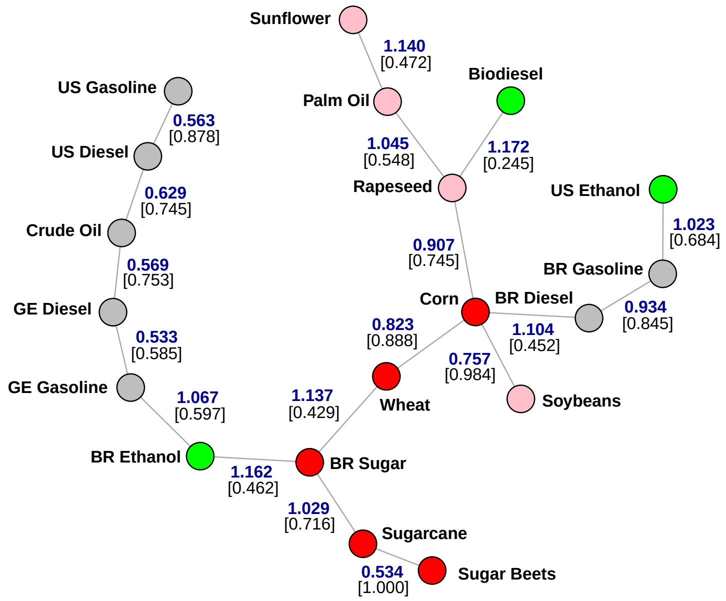
Compared to the previous period, the current monthly MST in Figure 5.14 gets little more complicated. Although a basic distinctions between individual clusters remains, we also identify some changes in the network's design. The whole tree structure gets slightly rearranged and its edges expand as individual items get less interrelated. The average length of the link extends to $d_{ij} = 0.849$ compared to $d_{ij} = 0.778$ obtained for the previous monthly MST.

As observed in each of the earlier MSTs, BR ethanol keeps its connection to BR sugar. Because we are well acquainted with the presence of this link, we notice that on monthly frequency its stability has been gradually decreasing since the first period. In a consecutive order, we obtain bootstrap values $b_{ij} = 0.705, b_{ij} = 0.625$ and $b_{ij} = 0.462$. On weekly level, the total decrease in this link's stability from $b_{ij} = 0.916$ to $b_{ij} = 0.577$ is also obvious. In the last five years, Brazilian ethanol market suffered from multiple negative effects. Due to temporary supply shortages in 2010 and 2011, ethanol price increased substantially. Brazilian government intended to reduce the demand for ethanol by introducing lower blending mandates while maintaining artificially low gasoline prices. Many flex fuel car owners switched to conventional gasoline because of ethanol's high price. As a result, a share of flex fuel cars regularly run on ethanol dropped from 66% in 2009 to as low as only 23% at year end 2013.

Biodiesel maintains its link to rapeseed. However, their connection becomes also a little weaker and even less stable ($b_{ij} = 0.245$). In fact, this is the least stable of all connection we have observed between biodiesel and its feedstock. This instability results from biodiesel's strong ties to several other items. In almost 65% of bootstrapped cases, biodiesel connects to some of the vegetable oils (mainly palm and rapeseed oils). Apart from vegetable oils, biodiesel often attaches to either corn or wheat. We repeatedly notice that the behavior of biodiesel differs from that of US and Brazilian ethanols.

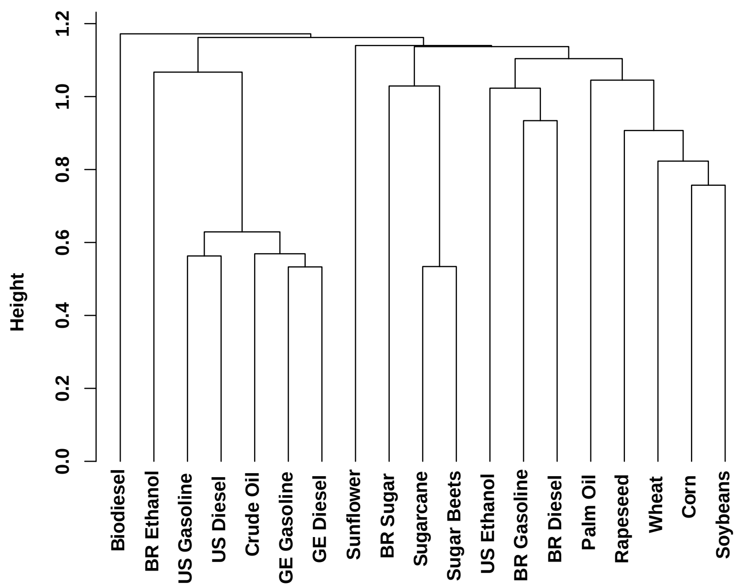
Perhaps, the most surprising change occurs in the case of US ethanol, although it still remains a part of the food branch. Already the weekly MST did not contain US ethanol's previously stable connection to corn. With decreasing frequency, it moves further away from its feedstock and becomes primarily connected to the pair of Brazilian retail fuels. Although the US ethanol's link to feedstock commodities may have loosened somehow, there is not any technical reason for the US biofuel to attach to Brazilian fossil fuels. We will focus on this anomaly in the next chapter where the ethanol's relationship to feedstock crops will be further studied. The associated HT in Figure 5.15 completes the story of the third period considered by the taxonomy analysis.

Figure 5.14: MST – Adjusted Taxonomy, Period III, Monthly



Note: [Bootstrap Value], Edge Lengths in bold, Source: Author's Computation

Figure 5.15: HT – Adjusted Taxonomy, Period III, Monthly



Source: Author's Computation

In this chapter, we modeled a system of biofuel related prices. We covered a period of the last decade focusing on world's major biofuel markets, the US, the EU, and Brazil. We started with a broadly defined experimental system. The initial price system consisted of biofuels, corresponding fossil fuels and crude oil, biofuel feedstock and food commodities along with selected financial items. Our first results then shaped a simplified version of the system that better reflects the biofuels production logic. The adjusted system includes only physical commodities; biofuels, fossil fuels, crude oil, and feedstock crops.

The taxonomy approach of MSTs and HTs was employed on the three consecutive subperiods divided by the first and second food price crises. Moreover, to distinguish between short and medium term effects, weekly and monthly data frequencies were used. Link stability was tested by the bootstrapping method. We were especially interested in describing the strength and stability of the links between biofuels and their production factors. The taxonomy perspective proved to be very useful. Resulting tree objects tell us how a given commodity system was interconnected during particular periods of time.

A key advantage of this concept is its straightforward principle and a wide applicability. In fact, the concept is based on transforming a simple correlation into a distance metric. A potential weakness stemming from the method's simplicity is that taxonomies do not tell us anything about the directions of the relationships identified among individual items.

Our taxonomy analysis results in selecting particular biofuel–feedstock connections to be subject of further study. Since the nature of these connections is of our central interest, we are going to explore them further. Thus, a pool of selected connections is followed up with the wavelet analysis in the next chapter.

Chapter 6

Wavelet Analysis

This chapter builds on the knowledge acquired through the taxonomy toolbox. Having identified several links between biofuels and their feedstock, we now approach these pair connections separately. We intend to learn mainly about their importance and their evolution in time. To accomplish this goal we employ the wavelet coherence methodology explained in Chapter 3. In short, we use continuous wavelet framework to study each biofuel's links over the 11 year time period.

Usually, statistical correlation is only studied in time dimension. Wavelet analysis basically adds the frequency dimension into the analysis. Therefore, wavelet framework allows for exploring the correlation relationship in both time and frequency domains. The output for each studied biofuel–feedstock pair is presented in form of two charts. While the horizontal axes show time (in years), there is also the frequency or period (in days) on the vertical axes. Coherence is indicated by color according to a spectrum shown at the right edge. Pale colored corner areas are not of a reliable interpretation. They resulted from artificially adding zeros to the beginning and to the end of analyzed series. A central bright colored area delivers reliable results. Furthermore, regions with statistically significant coherence are bounded with a thick black curve.

In the left panel, we preset the squared wavelet coherence between biofuel and a given feedstock commodity. Since there is no negative wavelet coherence, phase difference between the series is indicated by directed arrows. Put simply, the arrows show what the direction of the relationship is. Rightward pointing arrows mean that biofuel is positively correlated with that particular feedstock while leftward pointing arrows indicate a negative relationship. If the arrows point straight down, biofuel leads the price of feedstock by $\pi/2$. In the contrary, upward pointing arrows imply that price of biofuel is led by feedstock.

In the right panel, we preset the partial wavelet squared coherence in detail described in Chapter 3. In the food–fuel system, crude oil plays a role of an important price driver affecting both fuel and food part of the commodity system. To some extent, food, fuel, and biofuel prices can all depend on crude oil price. Therefore, our wavelet coherence output (*left*) is supplemented by partial wavelet coherence charts (*right*) where we control for the the influence of crude oil.

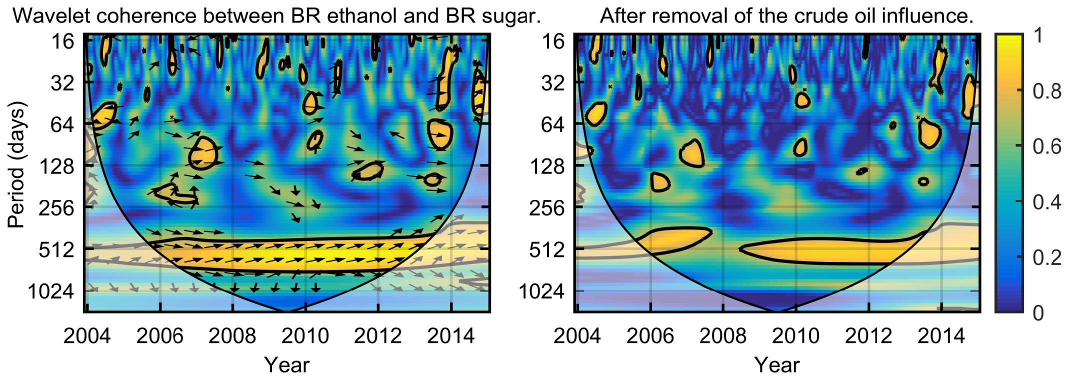
6.1 Brazilian Ethanol

In the context of global biofuel production, Brazil represents a unique scenario. Brazilian ethanol is produced from sugarcane. Intuitively, its price is supposed to be related to the price of feedstock. As documented by previous taxonomy objects, the price of ethanol is directly linked to the price of Brazilian sugar. Moreover, their connection seems to be relatively strong and stable, it appeared in every but one of the constructed MSTs. Hence, we are especially keen to explore Brazilian ethanol's correlation with local sugar price in more detail.

The left panel of Figure 6.1 shows a strong relationship placed approximately between 500 and 700 days, i.e. roughly between 1.5 and 2 years. Thus, we observe a long term relationship between BR ethanol and BR sugar that is remarkably stable in time. Apart from this main relationship, we identify only several minor coherence islands associated with relatively quick price interactions in the short term. The phase arrows in the main significant region point to the right and upwards indicating a positive correlation between the prices of ethanol and sugar with sugar leading the price of ethanol.

The right panel of Figure 6.1 delivers the output of partial wavelet coherence when the influence of crude oil has been controlled for. Apparently, crude oil consumed only a little portion of the correlation. Even if we see an interruption corresponding to the first food price crisis, our qualitative findings stay unchanged. Hence, we report a stable and positive relationship between BR ethanol and BR sugar with sugar leading the price of ethanol in the long term. We found the ethanol–sugar dependence which represents a long term systematic co-movement of biofuel and its central feedstock.

Figure 6.1: Wavelet Coherence: BR Ethanol versus Feedstock



Wavelet squared coherence (*left*) and partial wavelet coherence (*right*)
Coherence color spectrum shown at the right edge. Thick black curve marks significant regions. Phase differences are indicated by the directed arrows.
Source: Author's Computation

For decades now, Brazil represents an example of a biofuel economy. During more than forty year history, Brazilian ethanol industry went through numerous changes. Owing to its geographic and weather conditions, Brazil is well predisposed to grow sugarcane, from which ethanol fuel is produced at lower costs than from corn. Since its early days in 1970's, Brazilian biofuel industry has been primarily shaped by governmental policies. On the local retail fuel market, Brazilian sugarcane ethanol has always competed with conventional gasoline, whose regulated price has not always followed the world price. A single feedstock biofuel industry depends on annual harvests and crop yields. Unfavorable season may cause high sugarcane and ethanol prices resulting in intensified needs for foreign biofuel imports. Due to a supply shortage, Brazil imported about 1.5 billion liters of ethanol form the US during 2011-2012.

As a part of national energy security, Brazilian ethanol industry is expected to grow further. In early 2015, Brazilian government announced a new blending mandate increasing the ethanol share in gasoline from 25% to the new level of 27%. As of late 2014, Brazil was expected to generate as much as 26.9 billion liters of ethanol in 2015, a 5% increase from actual 2014 levels. Moreover, ethanol exports are projected to increase by 200 million liters reaching 1.8 billion liters in 2015. At present, flex fuel cars constitute some 55% of Brazilian fleet and the percentage is rising. In particular, more than 90% of new cars sold in Brazil are flex fuel vehicles. An 80% fleet share is expected to be reached by 2020 (Barros 2014).

6.2 US Ethanol

The biofuel–feedstock relationship of US ethanol (Figure 6.2) gets a little more complex compared to the Brazilian scenario. We explore the connections of US ethanol to corn, wheat, and sugarcane, respectively. We find that US ethanol is significantly tied to its feedstock and that the dynamics alter for individual commodities. In the taxonomy structures, US ethanol was further repeatedly linked to US and BR gasolines. For completeness, we also examine both of these dependencies with the same toolbox (Figure A.1) attributing them to the influence of crude oil.

Corn

About 90% of US ethanol is made from corn. In 2014, the US ethanol industry consumed almost 127 megatons of corn accounting for 40% of the US corn production (Conca 2014). Compared to Brazilian sugarcane industry, the US ethanol production from corn is a more technically demanding process. Put simply, corn crops first need to be converted to sugar before ethanol fuel gets produced. It implies higher production costs for US ethanol.

Our results (Figure 6.2 top) show that the relationship between US ethanol and corn consists of two strong dependencies of different kind. We find significant coherence areas associated with both short term and long term horizons. First, a long term relationship approximately at the level of 500 days (almost 1.5 years) has been steadily present since the period following the food crisis of 2008. Second, its rightwards pointing phase arrows tell us that US Ethanol has been positively correlated with corn throughout the second half of the studied time frame. Third, we learn that corn leads the price of ethanol since the arrows are also pointing slightly upwards.

The other type of dependency is a collection of short term price interactions. These time-limited episodes are associated with very high corn prices, e.g. the first food price crisis. During those events, the phase difference between the series decreases. Altogether, we claim a stable long term relationship accompanied by several short term episodes associated with very high corn prices especially between 2010 and 2013. Throughout the last decade, the relationship has always been positive with corn leading the price of ethanol. When the influence of crude oil is controlled for, we apparently loose a part of correlation. Especially the long term relationship between ethanol and corn gets somewhat reduced. However, other qualitative results do not get affected in fact.

Wheat

Our wavelet coherence output (Figure 6.2 middle) shows there exists a stable relationship between US ethanol and wheat that persists over the whole time frame. The significant coherence region begins at the period of approximately 800 days, i.e. just over two years. However, the reaction time keeps shortening up to a 500 day level. This long term dependency is accompanied by several quick interactions mostly at medium term horizon. Vast majority of phase arrows imply positive correlation between ethanol and wheat. From the long term perspective, wheat is obviously a leader of this price relationship since phase arrows are pointing upwards. Most of the short term correlation gets eliminated by removing the influence of crude oil. Our long term results get reduced especially during the first food price crisis of 2008. However, the basic dynamics remains unchanged when controlled for the effect of crude oil.

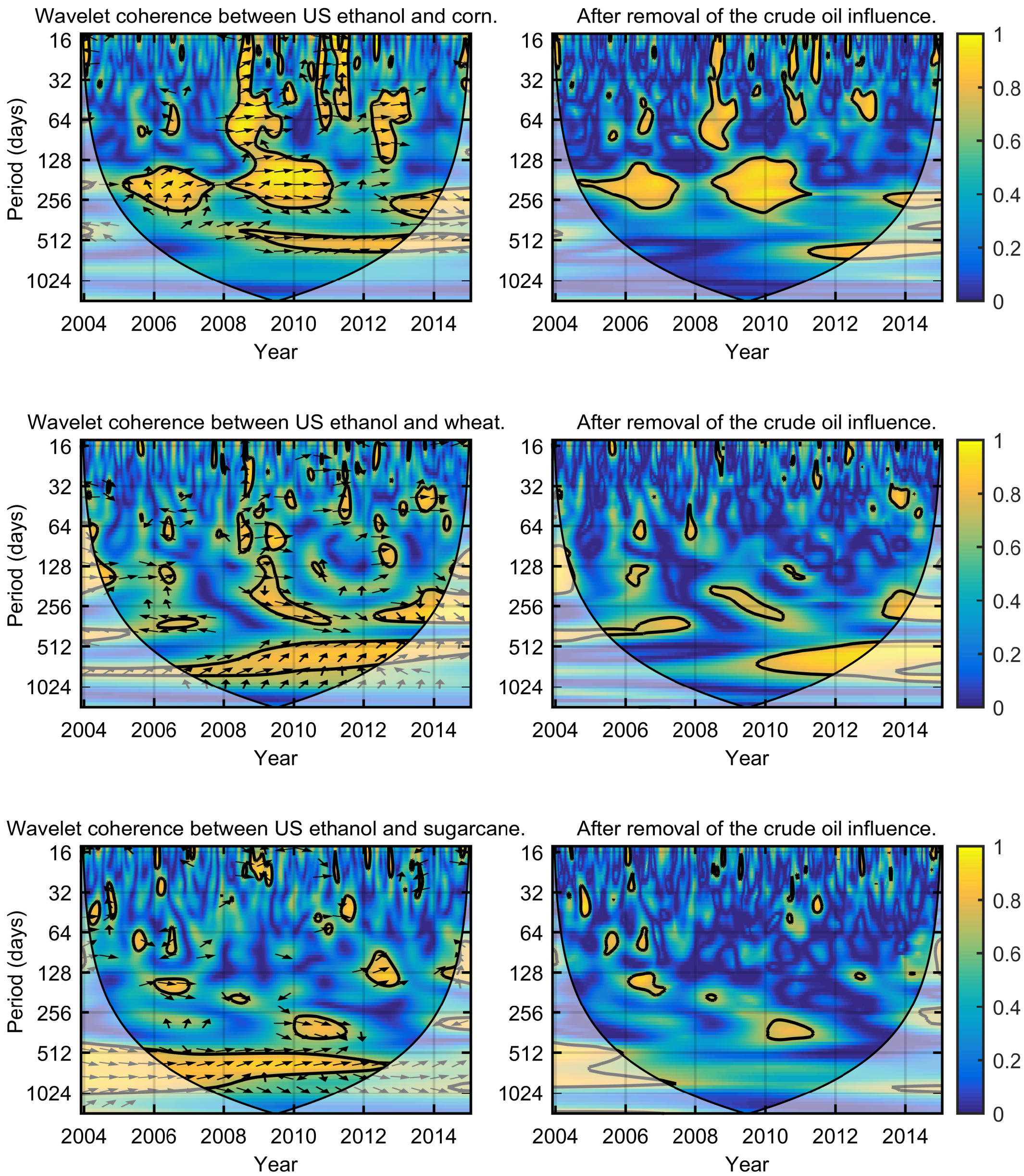
Although representing a major ethanol feedstock both in Canada and in the UK, only a marginal share of ethanol is produced from wheat in the US. As seen in Figure 4.2 and demonstrated by the taxonomy analysis, wheat price is closely linked to corn price. A major part of correlation between ethanol and wheat is therefore attributed to wheat's precise price co-movement with corn.

Sugarcane

The last panel of Figure 6.2 demonstrates the convenience of accounting for a possible influence of crude oil. It may appear from the wavelet coherence chart there exist strong and stable ties between US ethanol and sugarcane. Nonetheless, effectively all the correlation in the interpretable cone of influence disappears when the effect of crude oil is controlled for. Based on the wavelet analysis, we conclude that US ethanol has not been significantly related to sugarcane price in the last decade.

Since 2005, the US maintains a position of the world's major ethanol producing county, ahead of Brazil which used to be the previous leading ethanol producer for decades. In 2014, the US produced about 54 billion liters of ethanol. On the other hand, the US also represents the world's biggest consumer of oil. Based on EIA's statistics, the 2014 US consumption totaled 517 and 189 billion liters of gasoline and diesel, respectively. With recently expired tax credits and import tariffs, the biofuel industry has to face imports from Brazil, low crude oil prices and even attempts to reduce the blending mandate.

Figure 6.2: Wavelet Coherence: US ethanol versus Feedstock



Wavelet squared coherence (*left*) and partial wavelet coherence (*right*) Coherence color spectrum shown at the right edge. Thick black curve marks significant regions. Phase differences are indicated by the directed arrows.

Source: Author's Computation

6.3 European Biodiesel

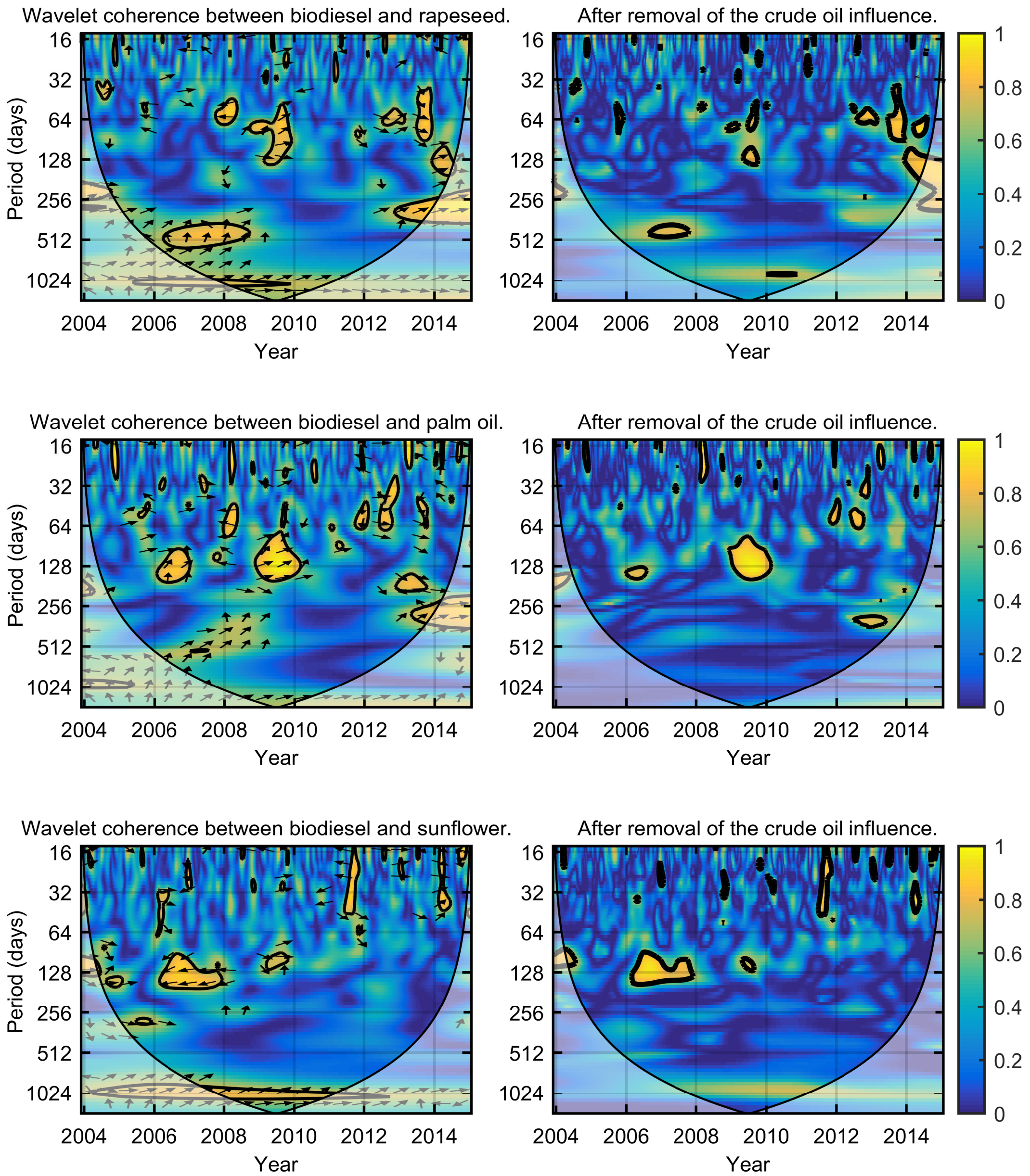
From the viewpoint of the previous taxonomy objects, biodiesel exhibited a price behavior different from both US and Brazilian ethanol. It was poorly integrated into either of the constructed networks. Now we separately analyze biodiesel's connection to its feedstock commodities; rapeseed, palm, and sunflower oil, respectively. Associated wavelet coherence output is presented in Figure 6.3. For completeness, we also examine biodiesel's relationship to selected fossil fuels and deliver the output in Figure A.2 in Appendix A.

Unlike both major ethanol producing countries, the European biofuel industry builds on biodiesel. On energy basis, biodiesel represents approximately 80% of the total transport biofuels market. It was the first EU biofuel employed in the road transport starting from 1990s. At that time, biofuel's rapid expansion was driven by increasing crude oil prices and regulations such as the Blair House Agreement and resulting provisions on the production of oilseeds under Common Agricultural Policy. Biodiesel enjoyed generous tax incentives, mainly in Germany and France. EU biofuels goals set out in Directive 2003/30/EC, subsequent Renewables Directive 2009/28/EC and Fuel Quality Directive 2009/30/EC further pushed the use of biodiesel. Today, the EU represents the world's largest producer of biodiesel. With 10.9 billion liters generated in its 266 refineries during 2014 the EU itself accounts for about 45% of the world's biodiesel production. The EU is also a primary consumer of biodiesel. In 2014, the EU consumption totaled 12.3 billion liters including 1.7 billion liters of imports (Flach *et al.* 2014).

The fact that today's EU biofuel market is dominated by biodiesel as opposed to ethanol is not random and can be attributed to several government policies which shape the European biofuel industry. As pointed out by Kristoufek *et al.* (2012a), the EU and US biofuel targets follow different settings. The US requirements are set in volumes. In this sense, a liter of ethanol is considered the same as a liter of biodiesel. On the other hand, the EU blending rules are set in energy units. According to Hofstrand (2008), the amount of energy available from one liter of biodiesel equals to 1.54 liters of ethanol[1]. Historically, the European biofuel scheme was designed so as to prefer biodiesel due to its considerably higher energy density.

[1]According to Hofstrand (2008), 1 liter of biodiesel contains 32.6 MJ compared to 21.1 MJ in 1 liter of ethanol, energy considered in terms of net heating value.

Figure 6.3: Wavelet Coherence: Biodiesel versus Feedstock



Wavelet squared coherence (*left*) and partial wavelet coherence (*right*) Coherence color spectrum shown at the right edge. Thick black curve marks significant regions. Phase differences are indicated by the directed arrows.
Source: Author's Computation

The original Diesel engine was developed by Rudolf Diesel to run on natural-fats-based fuels. Generally, biodiesel can be used in any diesel engine type. The use biodiesel causes less wear and tear on engines, increases lubricity and engine efficiency. Compared to conventional diesel, biodiesel's combustion process releases up to 60% less CO^2 emissions (Conca 2014). Kristoufek *et al.* (2012a) highlights a historically higher proportion of diesel vehicles in Europe relative to gas vehicles in the US. Considerably higher European fuel taxes provided another incentive for the use of diesel engines with lower fuel consumption.

In accordance with the previous taxonomy findings, wavelet analysis does not yield much of a reliable relationship between European biodiesel and its feedstock. However, when an occasional correlation appears, it is positive and biodiesel is being led by the price of feedstock. In case of rapeseed, we suspect certain dependency with a very low frequency. Unfortunately, this thin coherence island is mainly outside the interpretable area. Besides, we detect a few short lasting episodes of positive correlation. Specifically, we recognize a low frequency price interaction associated with the 2007–2008 food price crisis. Once controlled for the crude oil influence, effectively all the coherence between biodiesel and rapeseed oil disappears. Furthermore, there is no significant relationship between biodiesel and palm oil that would persist for longer than a year. Although we see several short lived positive price interactions, they do not allow for any strong conclusion. Perhaps, there seems to be a positive stable relationship between biodiesel and sunflower oil. However, it appears to have been caused by their mutual ties to crude oil.

In summary, there are certain signs of a time limited positive correlation between biodiesel and its feedstock. During these episodes, feedstock led the price of biodiesel. Compared to the results obtained for Brazilian and US ethanol, we argue that the price of European biodiesel is very weekly connected with prices of individual feedstock commodities. In this respect, European biodiesel market substantially differs from the analyzed ethanol markets.

The US and Brazilian biofuel industries are dominated by corn and sugarcane, respectively. A single feedstock accounts for a vast majority of local biofuel production. In case of biodiesel, rapeseed oil represents the main feedstock from which up to 58% of European biodiesel is made. Rapeseed is followed by palm oil, which became more important in recent years, especially due to large price discounts as reported by Flach *et al.* (2014). Nonetheless, during the analyzed period, no more than 2/3 of the EU biodiesel production was fed by one crop type.

In Brazil there is effectively no possibility to switch for another feedstock. In the US, there exists a very limited possibility to switch from corn to wheat. However, this process is not flexible, it requires time and additional investment in technology. In contrast, the European biofuel production facilities can switch between several feedstock types including rapeseed, palm, soybeans or sunflower oil. These vegetable oils have similar consistency and can even be mixed with each other within the same production facility. In contrast with major ethanol markets, the EU biodiesel industry enjoys therefore higher short term flexibility of production factors and better operates in an environment of fluctuating feedstock prices.

In 2015, the EU biodiesel production is expected to remain flat at about 11 billion liters. The EU domestic biodiesel consumption will likely continue in its slightly decreasing trend intensified by lower crude oil prices. In the next years, the demand for biofuels in the EU will be primarily shaped by mandates of individual member states. The most recent development of the EU bioenergy policies brought important changes in April 2015. The member states are still required to supply at least 10% of energy used in transport from renewable sources by 2020 (RED 2009/28/EC). However, the new legislation limits the share of energy coming from the first generation biofuels at a 7% level. In other words, the biofuels generated from crops grown on agricultural land cannot exceed 7% of energy used in transportation by 2020. Although a political compromise was reached, the recent decision has its loud critics. Nonetheless, this political choice restores regulatory certainty and improves investors' understanding of the EU biofuel industry and its development until 2020.

Having analyzed the EU biofuel industry in the context of its main regulatory drivers, we conclude that European biodiesel plays a different role than do US and Brazilian ethanols in their domestic markets. The difference was demonstrated by the results of the taxonomy and wavelet analyses. European biodiesel industry does not depend on a single feedstock. In the same vein, our results show that biodiesel has been very weekly tied to its production factors. This result contrasts with strong biofuel–feedstock price co-movements we found in both US and Brazilian ethanol markets.

Chapter 7

Conclusion

This thesis delivered an innovative research effort within the context of biofuel economics. Our analysis focused on the world's major biofuel markets; Brazil, the United States of America, and the European Union. Present study covers 83% of global ethanol production and about 45% of biodiesel production. We studied the relationships between ethanol, biodiesel, associated agricultural commodities, crude oil, relevant fossil fuels, and a group of financial assets. For this purpose, we compiled a unique dataset containing as many as 32 weekly price series. Compared to peer research attempts, our dataset is especially comprehensive and covers 2003–2015 time period.

We combined two methods that are still new to financial series analysis. First, we used the taxonomy method of minimum spanning trees and hierarchical trees to classify and visualize an experimental system of biofuel-related commodities and assets. Such a broad attempt has not been undertaken before. Second, we introduced an adjusted version of the initial food–fuel system based solely on physical commodities. Third, the identified biofuel–feedstock price pairs were followed up using the wavelet analysis. Please note that this paper represents the first attempt to combine the taxonomy approach with the wavelet analysis toolbox within a single research application.

To differentiate the short term effects from the medium term effects, we constructed the tree objects separately for data with weekly and monthly frequency. In our tree structures, we use several innovations that were employed for the first time in the biofuel context. First, vertices of a MST were color coded to allow for tree's better legibility. Second, the length of constructed edges reflects their weight. Finally, our non-rectangular MST arrangement allows for visualizing complicated systems with high number of items.

In accordance with our initial hypothesis, resulting commodity systems get meaningfully structured. The interconnected networks consist of a fuel and food branch. Generally, the food part includes clusters of vegetable oils, sugars, and cereals, while retail fuels and crude oil belong stably to the fuel part. Fundamental connections emerge already for weekly frequency. In the medium term then, the networks get more structured as individual links shorten, the connections become closer. Our results show several crucial patterns.

We demonstrated an important phase shift between Brazilian and the US/EU biofuel producing sectors. Brazilian mature ethanol industry was characterized by a stable link between ethanol and sugar since the beginning of the studied time frame. The US ethanol market established a similar link between ethanol and corn with certain delay. In the same vein, a rise of European biodiesel industry became visible in late 2000s. While both Brazilian and US ethanols developed stable links to their primary feedstock commodities, biodiesel reflected a different production logic. It did not become particularly tied to either of the relevant feedstock crops. Biodiesel's unstable links confirmed it was not dependent on a single feedstock. In the contrary, biodiesel lively interacted with several commodities.

Subsequent wavelet analysis reported a strong long term relationship between Brazilian ethanol and its feedstock. We showed that the price of Brazilian ethanol was positively correlated with local sugar price and their relationship was stable in time. Importantly, sugar led the price of ethanol throughout the period. The dynamics remained qualitatively unchanged when the influence of crude oil was controlled for.

Furthermore, we found a similarly strong relationship between US ethanol and its main production factor, corn. Their price co-movement consisted of two positive dependencies. A long term stable relationship was accompanied by several coherence episodes at higher frequencies. These short run events coincided with periods of very high corn prices. We found that corn led the price of ethanol across the frequencies. Moreover, the ethanol–corn relationship proved to be robust to removing the influence of crude oil.

Finally, the showed that the behavior of biodiesel contrasts with both major ethanol markets. In accordance with the previous taxonomy structures, we conclude that biodiesel and ethanol have different positions in the food-fuel system. Over the course of the last decade, we saw a few short lived price interactions between biodiesel and the analyzed feedstock crops. However, major European biofuel did not exhibit any strong co-movement with feedstock.

In summary, we succeeded in confirming our initial hypotheses. First, we described an interconnected system of biofuel-related commodities. Moreover, we commented on its evolution over the course of eleven year period. Second, we documented a phase shift that initially occurred between mature Brazilian and belated US/European biofuel industries. Third, we demonstrated a positive price co-movement of ethanol and its respective production factors. We further showed that this relationship is stable in time with feedstock leading the price of ethanol. Finally, we explained that the price of biodiesel did not depend on a single feedstock commodity. Biodiesel weakly interacted with several crops through more random price adjustments. Thus, the European biofuel industry substantially differs from both Brazilian and the US establishments.

The main contribution of this thesis lies in its innovative and comprehensive approach. Employed methods make as few ex-ante assumptions as possible. In particular, the wavelet coherence methodology represents a widely applicable model-free toolbox. Therefore, our results are not model specific. Our findings contribute to the current biofuel policy discussion. Specifically, we stress the difference between ethanol and biodiesel production processes. Eventually, we shed new light on biofuel–feedstock connections on the leading global markets.

Our results can be used as a suitable starting point for further (possibly more advanced) research. First and foremost, we recommend focusing on various other specification of sub-periods considered in the taxonomy analysis. Setting shorter periods may bring smoother resolution of a varying market environment. Perhaps, taxonomy structures shall be constructed based on lower frequency data, too. Further research attempts may focus on quarterly and yearly data to account for longer term patterns. In the realm of subsequent wavelet analysis, we suggest controlling simultaneously for multiple sources of possible spurious correlation. Lastly, it would be very interesting to rerun the wavelets also on associated demand/consumption data. Most likely, such an effort would have to cope with serious data unavailability.

Bibliography

Abbott, P. (2013): "Biofuels, Binding Constraints and Agricultural Commodity Price Volatility." *NBER Working Papers 18873*, National Bureau of Economic Research, Inc.

Algieri, B. (2014): "The influence of biofuels, economic and financial factors on daily returns of commodity futures prices." *Energy Policy* **69(C)**: pp. 227–247.

Barros, S. (2014): "Brazil biofuels annual 2014." *USDA FOREIGN AGRICULTURAL SERVICE* .

Bondy, J. & U. Murty (1976): *Graph Theory with Applications: By J.A. Bondy and U.S.R. Murty*, volume 124. Macmillan.

Carter, C., G. Rausser, & A. Smith (2012): "The Effect of the US Ethanol Mandate on Corn Prices." *Working paper,* UC Davis, Davis, CA.

Chance, W. A. (1986): "A geometric derivation of the distribution of the correlation coefficient rwhen rho = 0." *Amer. Math. Monthly* **93**: pp. 94–98.

Ciaian, P. & d. Kancs (2011): "Food, energy and environment: Is bioenergy the missing link?" *Food Policy* **36(5)**: pp. 571–580.

Conca, J. (2014): "It's final – corn ethanol is of no use." *Forbes Business* www.forbes.com/sites/jamesconca/2014/04/20/its-final-corn-ethanol-is-of-no-use/.

Cuesta, J., A. Htenas, & S. Tiwari (2014): "Monitoring global and national food price crises." *Food Policy* **49(P1)**: pp. 84–94.

Daubechies, I. (2004): *Ten Lectures on Wavelets.* SIAM.

Dickey, D. A. & W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American statistical association* **74**: pp. 427–431. Taylor & Francis.

Diestel, R. (2000): *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. http://www.math.uni-hamburg.de/home/diestel/books/graph.theory/: Springer-Verlag New York, second edition.

Drabik, D. (2011): "The theory of biofuel policy and food grain prices." *Working Papers 126615*, Cornell University, Department of Applied Economics and Management.

Drabik, D., P. Ciaian, & J. Pokrivcak (2014): "Biofuels and Vertical Price Transmission: The Case of the U.S. Corn, Ethanol, and Food Markets." *LICOS Discussion Papers 35114*, LICOS - Centre for Institutions and Economic Performance, KU Leuven.

Du, X. & D. J. Hayes (2012): "Impact of Ethanol Production on U.S. and Regional Gasoline Markets: An Update to 2012, The." *Food and Agricultural Policy Research Institute (FAPRI) Publications 12-wp528*, Food and Agricultural Policy Research Institute (FAPRI) at Iowa State University.

Flach, B., K. Bendz, & S. Lieberz (2014): "Eu-28 biofuels annual." *EU Biofuels Annual 2014* gain.fas.usda.gov/Recent

Good, P. (2010): "Robustness of pearson correlation." http://interstat.statjournals.net/YEAR/2009/articles/0906005.pdf.

de Gorter, H., D. Drabik, E. M. Kliauga, & G. R. Timilsina (2013a): "An economic model of Brazil's ethanol-sugar markets and impacts of fuel policies." *Policy Research Working Paper Series 6524*, The World Bank.

de Gorter, H., D. Drabik, & G. R. Timilsina (2013b): "The effect of biodiesel policies on world oilseed markets and developing countries." *Policy Research Working Paper Series 6453*, The World Bank.

de Gorter, H. & D. R. Just (2009a): "The economics of a blend mandate for biofuels." *American Journal of Agricultural Economics* **91(3)**: pp. 738–750.

de Gorter, H. & D. R. Just (2009b): "The welfare economics of a biofuel tax credit and the interaction effects with price contingent farm subsidies." *American Journal of Agricultural Economics* **91(2)**: pp. 477–488.

Grinsted, A., J. Moore, & S. Jevrejeva (2004): "Application of the cross wavelet transform and wavelet coherence to geophysical time series." *Nonlinear Processes in Geophysics* **11**: pp. 561–566.

Hao, N., G. Colson, B. Karali, & M. E. Wetzstein (2013): "Food before Biodiesel Fuel?" *2013 Annual Meeting, February 2-5, 2013, Orlando, Florida 143078*, Southern Agricultural Economics Association.

Hochman, G. (2014): "Biofuels at a Crossroads." *Choices* **29(1)**.

Hofstrand, D. (2008): "Liquid fuel measurements and conversions." *Iowa State University, University Extension* Www.extension.iastate.edu/agdm/wholefarm/pdf/c6-87.pdf.

Hudson, R. & A. Gregoriou (2010): "Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns." *Available at SSRN: http://ssrn.com/abstract=1549328 or http://dx.doi.org/10.2139/ssrn.1549328* .

Jarque, M. & A. K. Bera (1987): "A test for normality of observations and regression residuals." *Internat. Statist. Rev* pp. 163–172.

Kristoufek, L., K. Janda, & D. Zilberman (2012a): "Correlations between biofuels and related commodities: A taxonomy perspective." *Working Papers IES 2012/15*, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

Kristoufek, L., K. Janda, & D. Zilberman (2012b): "Mutual Responsiveness of Biofuels, Fuels and Food Prices." *CAMA Working Papers 2012-38*, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

Kristoufek, L., K. Janda, & D. Zilberman (2013): "Non-linear price transmission between biofuels, fuels and food commodities." *Working Papers IES 2013/16*, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

KRISTOUFEK, L., K. JANDA, & D. ZILBERMAN (2014): "Price transmission between biofuels, fuels, and food commodities." *Biofuels, Bioproducts and Biorefining* **8(3)**: pp. 362–373.

KRISTOUFEK, L., K. JANDA, & D. ZILBERMAN (2015): "Co-movements of ethanol related prices: Evidence from Brazil and the USA." *GCB Bioenergy* .

KRUSKAL, J. B. (1956): "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem." *Proceedings of the American Mathematical Society* **7(1)**: pp. 48–50.

KWIATKOWSKI, D., P. C. B. PHILLIPS, P. SCHMIDT, & Y. SHIN (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics* **54(1-3)**: pp. 159–178.

MANTEGNA, R. (1999): "Hierarchical structure in financial markets." *The European Physical Journal B - Condensed Matter and Complex Systems* **11(1)**: pp. 193–197.

MANTEGNA, R. N. & H. E. STANLEY (2000): *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge Univ. Press, Cambridge UK.

MATOUSEK, J. & J. NESETRIL (2008): *Invitation to Discrete Mathematics*. Oxford: Oxford Univ. Press, UK.

MCPHAIL, L. L. & B. A. BABCOCK (2012): "Impact of US Biofuel Policy on US Corn and Gasoline Price Variability." *Staff General Research Papers 34892*, Iowa State University, Department of Economics.

MIHANOVIC, H., M. ORLIC, & Z. PASRIC (2009): "Diurnal thermocline oscillations driven by tidal flow around an island in the middle adriatic." *Journal of Marine Systems* **78, Supplement(0)**: pp. S157 – S168. Coastal Processes: Challenges for Monitoring and Prediction.

MILLER, E., M. L. MALLORY, K. R. BAYLIS, & C. E. HART (2012): "Basis Effects Of Ethanol Plants In The U.S. Corn Belt." *2012 Annual Meeting, August 12-14, 2012, Seattle, Washington 124984*, Agricultural and Applied Economics Association.

Peri, M. & L. Baldi (2013): "The effect of biofuel policies on feedstock market: Empirical evidence for rapeseed oil prices in EU." *Resource and Energy Economics* **35(1)**: pp. 18–37.

Piroli, G., M. Rajcaniova, P. Ciaian, & d'Artis Kancs (2014): "From a rise in B to a fall in C? Environmental impact of biofuels." *EERI Research Paper Series EERI RP 2014/01*, Economics and Econometrics Research Institute (EERI), Brussels.

Rajcaniova, M., d'Artis Kancs, & P. Ciaian (2014): "Bioenergy and global land-use change." *Applied Economics* **46(26)**: pp. 3163–3179.

Rausser, G. C. & H. de Gorter (2013): "US policy contributions to agricultural commodity price fluctuations, 2006.12." *Working Paper Series UNU-WIDER Research Paper*, World Institute for Development Economic Research (UNU-WIDER).

Savaşsçin, Ö. (2011): "The dynamics of commodity prices: A clustering approach." *Univesity of North Carolina, Chapel Hill* .

Serra, T. & D. Zilberman (2013): "Biofuel-related price transmission literature: A review." *Energy Economics* **37(0)**: pp. 141 – 151.

Serra, T., D. Zilberman, J. M.Gil, & B. K. Goodwin (2010): "Price transmission in the US ethanol market. In M. Khanna, J. Scheffran, and D. Zilberman (Eds.)." *Handbook of Bioenergy Economics and Policy, Natural Resource Management and Policy, Natural Resource Management and Policy, Chapter 5,* pp. 55–72. Springer.

Shapiro, S. S. & M. B. Wilk (1965): "An analysis of variance test for normality (complete samples)." *Biometrika* **52(3-4)**: pp. 591–611. Http://biomet.oxfordjournals.org/content/52/3-4/591.short.

Torrence, C. & G. P. Compo (1998): "A practical guide to wavelet analysis." *Bulletin of the American Meteorological Society* **79**: pp. 61–78.

Torrence, C. & P. Webster (1998): "The annual cycle of persistence in the el nino-southern oscillation." *Quarterly Journal of the Royal Meteorological Society* **124(550)**: pp. 1985–2004.

TUMMINELLO, M., F. LILLO, & R. N. MANTEGNA (2007): "Shrinkage and Spectral Filtering of Correlation Matrices: a Comparison Via the Kullback–Leibler Distance." *Acta Physica Polonica B* **38**: p. 4079.

VACHA, L. & J. BARUNIK (2012): "Co-movement of energy commodities revisited: Evidence from wavelet coherence analysis." *ArXiv e-prints* .

VACHA, L., K. JANDA, L. KRISTOUFEK, & D. ZILBERMAN (2013): "Time–frequency dynamics of biofuel–fuel–food system." *Energy Economics* **40(C)**: pp. 233–241.

WANG, S. L. & L. L. MCPHAIL (2012): "Impacts of Energy Shocks on US Agricultural Productivity Growth and Food Prices â€"A Structural VAR Analysis." *2012 Annual Meeting, August 12-14, 2012, Seattle, Washington 124892*, Agricultural and Applied Economics Association.

WOOLDRIDGE, J. (2008): *Introductory Econometrics: A Modern Approach (with Economic Applications, Data Sets, Student Solutions Manual Printed Access Card).* South-Western College Pub, 4 edition.

ZILBERMAN, D., G. HOCHMAN, S. KAPLAN, & E. KIM (2014): "Political Economy of Biofuel." *Choices* **29(1)**.

Appendix A

Additional Results & Tables

To provide a complete analysis, we also used wavelet coherence framework to examine the relationship between biofuels and their traditional fossil fuel counterparts.

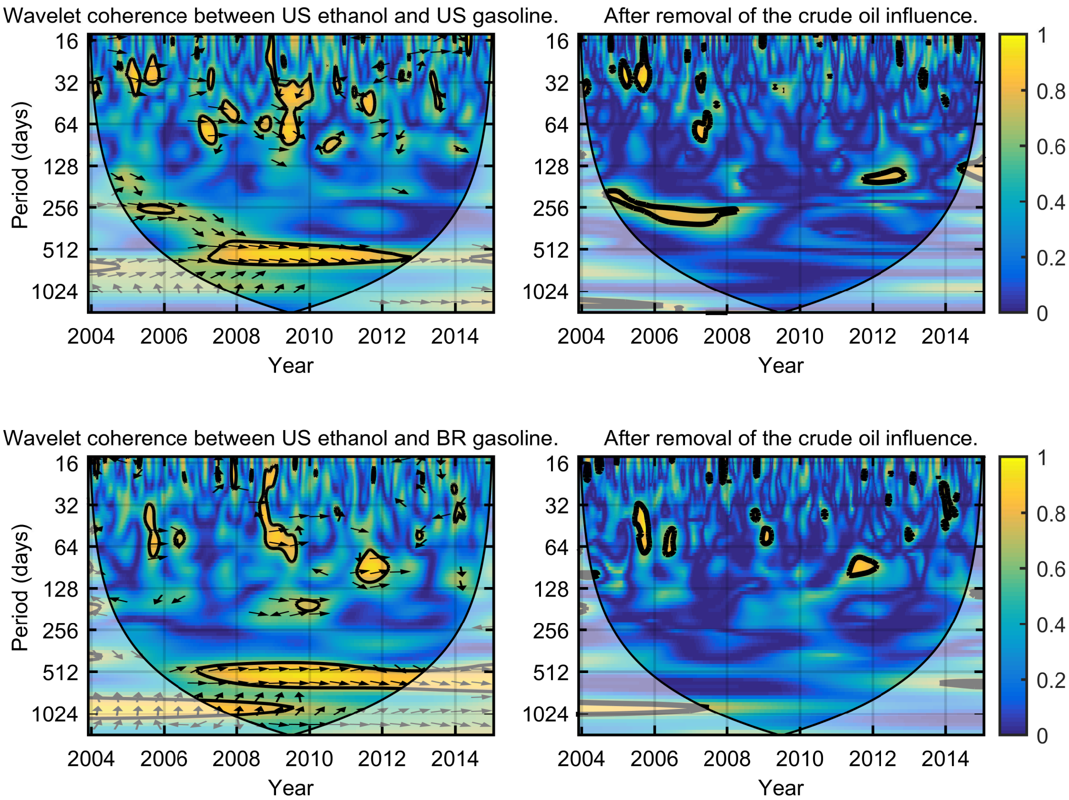
US Ethanol

In the MSTs, US ethanol was repeatedly connected to US gasoline and BR gasoline. Especially US gasoline represents its direct competitor on domestic retail fuel market. Moreover, ethanol is also blended to gasoline in both Brazil and the US. In Figure A.1 we observe that both ethanol–gasoline relationships appear due to the commodities' strong mutual ties to crude oil. Once crude oil has been controlled for, effectively all the correlation disappears.

European Biodiesel

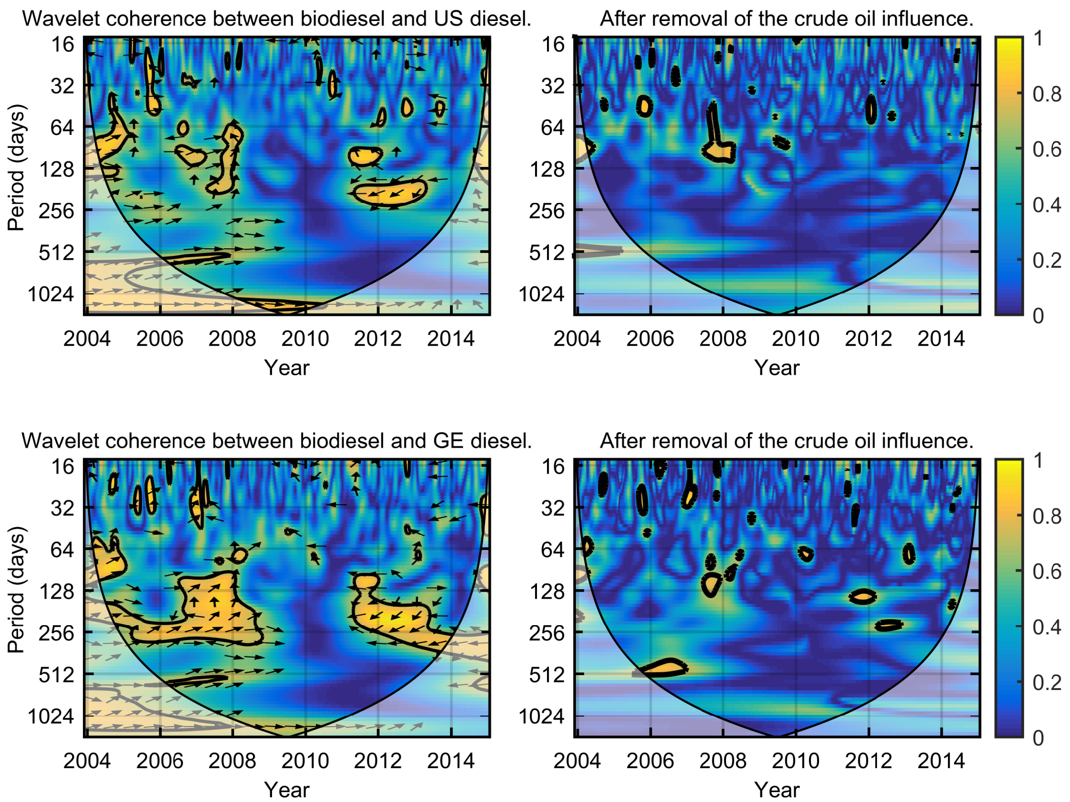
We performed a similar attempt for European biodiesel which was often tied to US diesel. Moreover, we examined biodiesel's relation to German diesel, its counterpart in the geographic market. Not surprisingly, all the coherence was found due to a strong influence of crude oil on price of biofuel and both retail fuels (Figure A.2).

Figure A.1: Wavelet Coherence: US Ethanol versus Fossil Fuels



Wavelet squared coherence (*left*) and partial wavelet coherence (*right*) Coherence color spectrum shown at the right edge. Thick black curve marks significant regions. Phase differences are indicated by the directed arrows.
Source: Author's Computation

Figure A.2: Wavelet Coherence: Biodiesel versus Fossil Fuels



Wavelet squared coherence (*left*) and partial wavelet coherence (*right*) Coherence color spectrum shown at the right edge. Thick black curve marks significant regions. Phase differences are indicated by the directed arrows.

Source: Author's Computation

Table A.1: Analyzed Data and Sources

Asset	**Ticker**	**Source**	**Type**
US Ethanol	ETHNNYPR Index	Bloomberg	Spot, FOB, anhydrous ethanol
Brazilian Ethanol	-	CEPEA	Anhydrous ethanol
Biodiesel	BIOCEUGE Index	Bloomberg	German biodiesel, spot
	FAME FOB ARA	Reuters	Spot price, ARA OTC
Corn	C 1 Comdty	Bloomberg	1st month futures, CBOT
Wheat	W 1 Comdty	Bloomberg	1st month futures, CBOT
Sugarcane	SB1 Comdty	Bloomberg	1st month futures, ICE
Sugar Beets	QW1 Comdty	Bloomberg	1st month futures, LIFFE
Brazilian Sugar	-	CEPEA	Spot USD Price
Rapeseed Oil	IJ1 Comdty	Bloomberg	1st month futures
Soybean Oil	S 1 Comdty	Bloomberg	1st month futures, CBOT
Sunflower Seeds	SU1	Bloomberg	1st month futures
Palm Oil	KO3 Comdty	Bloomberg	1st month futures
Brent Crude Oil	CO1 Comdty	Bloomberg	1st month futures, ICE
German Diesel	-	EIA	Retail Diesel Prices
German Gasoline	-	EIA	Retail Premium Gasoline
US Diesel	-	EIA	Retail Diesel Prices
US Gasoline	-	EIA	Retail Premium Gasoline
Brazilian Diesel	-	ANP Brazil	Weighted av. consumer price
Brazilian Gasoline	-	ANP Brazil	Weighted av. consumer price
Coffee	AX1 Comdty	Bloomberg	Arabica, 1st month futures
Cocoa	CC1 Comdty	Bloomberg	1st month futures, NYBOT
Rice	RR1 Comdty	Bloomberg	1st month futures, CBOT
Oranges	OR1 Comdty	Bloomberg	1st month futures
Dow Jones	DJI Index	Bloomberg	US Dow Jones Ind. Average
S&P 500	SP1 Index	Bloomberg	US S&P 500 Index
FTSE 100	UKX Index	Bloomberg	British FTSE 100 Index
DAX	DAX Index	Bloomberg	German DAX Index
BOVESPA	IBOV Index	Bloomberg	Brazilian BOVESPA
Federal Funds	-	Federal Reserve	US Fed Funds Rate
LIBOR	-	ECONSTATS	3 months USD LIBOR
USD/EUR	-	ECB	
USD/BRL	-	Federal Reserve	

Table A.2: Descriptive Statistics

	Mean	**St. Dev.**	**Min**	**Median**	**Max**
Biodiesel	−0.0013	0.0883	−0.4613	0.0000	0.3159
US Ethanol	−0.0003	0.0570	−0.2201	0.0000	0.2085
BR Ethanol	0.0015	0.0433	−0.2952	0.0026	0.2213
Crude Oil	0.0010	0.0428	−0.1595	0.0036	0.2019
US Gasoline	0.0008	0.0244	−0.1040	0.0000	0.1816
DE Gasoline	0.0008	0.0462	−0.1807	0.0000	0.2250
BR Gasoline	0.0005	0.0212	−0.1241	0.0013	0.0896
US Diesel	0.0015	0.0225	−0.1095	0.0000	0.1394
DE Diesel	0.0008	0.0400	−0.1268	0.0000	0.1435
BR Diesel	0.0012	0.0220	−0.1240	0.0015	0.1031
Dow Jones	0.0010	0.0237	−0.1251	0.0019	0.1310
S&P 500	0.0011	0.0257	−0.1491	0.0023	0.1295
FTSE 100	0.0007	0.0246	−0.1303	0.0029	0.1426
DAX	0.0017	0.0298	−0.1604	0.0041	0.1482
BOVESPA	0.0015	0.0408	−0.2926	0.0032	0.2619
Fed Funds	−0.0036	0.1545	−1.4663	0.0000	1.0415
LIBOR	−0.0026	0.0383	−0.2461	0.0000	0.2207
USD/EUR	−0.00003	0.0138	−0.0745	0.0004	0.0498
USD/BRL	0.0002	0.0200	−0.1243	0.0014	0.0869
Corn	0.0009	0.0452	−0.2546	0.0036	0.1774
Wheat	0.0006	0.0445	−0.1204	−0.0024	0.1636
Sugarcane	0.0016	0.0477	−0.2372	−0.0007	0.1493
Sugar Beets	0.0013	0.0387	−0.1713	0.0014	0.1135
Sugar Brazil	0.0017	0.0354	−0.1595	0.0027	0.1399
Soybeans	0.0005	0.0416	−0.2809	0.0013	0.1390
Sunflower	0.0012	0.0366	−0.1849	0.0032	0.2060
Rapeseed	0.0005	0.0275	−0.1200	0.0018	0.0905
Palm Oil	0.0005	0.0388	−0.1877	0.0016	0.1811
Coffee	0.0020	0.0421	−0.1450	0.0004	0.2552
Cocoa	0.0012	0.0430	−0.1867	0.0017	0.2097
Rice	0.0005	0.0377	−0.2842	0.0012	0.1269
Oranges	0.0002	0.0420	−0.2583	0.0000	0.3021

Source: Author's Computation

Table A.3: Stationarity Tests – Log Price

	ADF	**p-value**	**KPSS**	**p-value**
Biodiesel	-2.3663	0.4232	2.2572	< 0.01
US Ethanol	-3.0861	0.1186	0.7632	< 0.01
BR Ethanol	-1.0217	0.9352	2.1431	< 0.01
Crude Oil	-2.807	0.2367	2.1194	< 0.01
Corn	-1.9261	0.6096	2.0771	< 0.01
Wheat	-1.9438	0.6021	1.6961	< 0.01
Sugarcane	-2.4172	0.4017	2.0697	< 0.01
Sugar Beets	-1.8895	0.6251	2.1029	< 0.01
BR Sugar	-2.1896	0.4981	1.8264	< 0.01
Soybeans	-2.5349	0.3519	2.3310	< 0.01
Sunflower	-2.8746	0.2081	2.4920	< 0.01
Rapeseed	-1.9080	0.6172	2.0757	< 0.01
Palm Oil	-2.2951	0.4534	1.6788	< 0.01
US Gasoline	-3.3778	0.0572	2.0791	< 0.01
US Diesel	-2.6930	0.2850	2.1536	< 0.01
DE Gasoline	-2.9360	0.1821	2.3059	< 0.01
DE Diesel	-1.8853	0.6269	2.1440	< 0.01
BR Gasoline	-1.8845	0.6272	1.4796	< 0.01
BR Diesel	-1.9533	0.5981	1.7767	< 0.01
Coffee	-2.1033	0.5346	1.8896	< 0.01
Cocoa	-1.9540	0.5978	2.3167	< 0.01
Rice	-1.7621	0.6790	2.0414	< 0.01
Oranges	-1.5344	0.7754	1.3045	< 0.01
Dow Jones	-1.5401	0.7730	1.4927	< 0.01
S&P500	-1.4610	0.8065	1.1970	< 0.01
FTSE 100	-2.1243	0.5257	1.0243	< 0.01
DAX	-1.9883	0.5833	1.9936	< 0.01
Bovespa	-1.5339	0.7756	2.0589	< 0.01
Fed Funds	-1.9916	0.5819	2.6192	< 0.01
Libor	-2.4946	0.3689	2.5366	< 0.01
USD/EUR	-2.2157	0.4870	0.5765	0.0247
USD/BRL	-1.4212	0.8233	1.0486	< 0.01

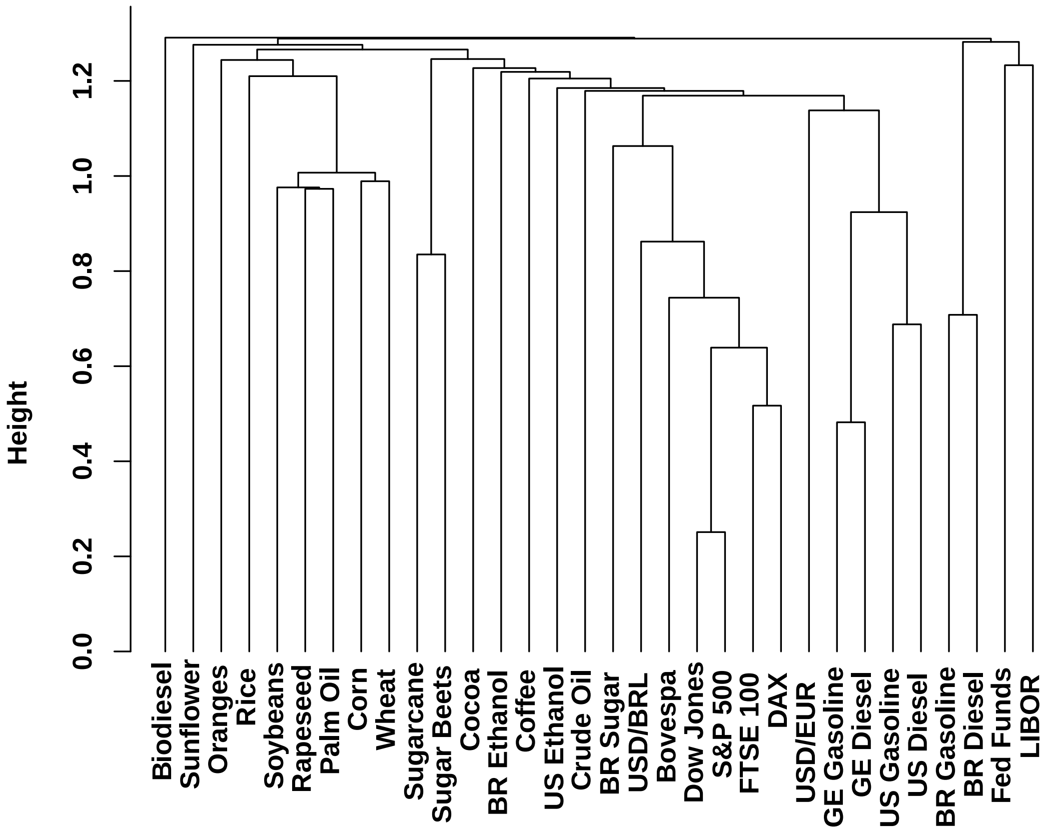
Source: Author's Computation

Table A.4: Normality Tests – Log Returns

	JB Test		SW Test	
	χ^2	p-value	W	p-value
Biodiesel	45,049.5	< 0.01	0.6331	< 0.01
US Ethanol	725.9	< 0.01	0.9308	< 0.01
BR Ethanol	2,372.8	< 0.01	0.8722	< 0.01
Crude Oil	116.9	< 0.01	0.9706	< 0.01
Corn	211.5	< 0.01	0.9702	< 0.01
Wheat	23.2	< 0.01	0.9884	< 0.01
Sugarcane	56.9	< 0.01	0.9843	< 0.01
Sugar Beets	49.2	< 0.01	0.9852	< 0.01
BR Sugar	171.7	< 0.01	0.9489	< 0.01
Soybeans	658.1	< 0.01	0.9708	< 0.01
Sunflower	494.1	< 0.01	0.9405	< 0.01
Rapeseed	125.5	< 0.01	0.9708	< 0.01
Palm Oil	188.1	< 0.01	0.9707	< 0.01
US Gasoline	1,177.6	< 0.01	0.9672	< 0.01
US Diesel	941.2	< 0.01	0.9201	< 0.01
DE Gasoline	157.3	< 0.01	0.9672	< 0.01
DE Diesel	21.7	< 0.01	0.9853	< 0.01
BR Gasoline	248.9	< 0.01	0.9661	< 0.01
BR Diesel	549.2	< 0.01	0.9544	< 0.01
Coffee	194.0	< 0.01	0.9774	< 0.01
Cocoa	42.8	< 0.01	0.9892	< 0.01
Rice	797.1	< 0.01	0.9607	< 0.01
Oranges	2,464.2	< 0.01	0.8867	< 0.01
Dow Jones	549.2	< 0.01	0.9409	< 0.01
S&P500	863.7	< 0.01	0.9212	< 0.01
FTSE 100	544.5	< 0.01	0.9427	< 0.01
DAX	344.9	< 0.01	0.9531	< 0.01
Bovespa	1,451.5	< 0.01	0.9373	< 0.01
Fed Funds	25,358.9	< 0.01	0.6493	< 0.01
Libor	4,813.6	< 0.01	0.7156	< 0.01
USD/EUR	105.7	< 0.01	0.9821	< 0.01
USD/BRL	324.1	< 0.01	0.9626	< 0.01

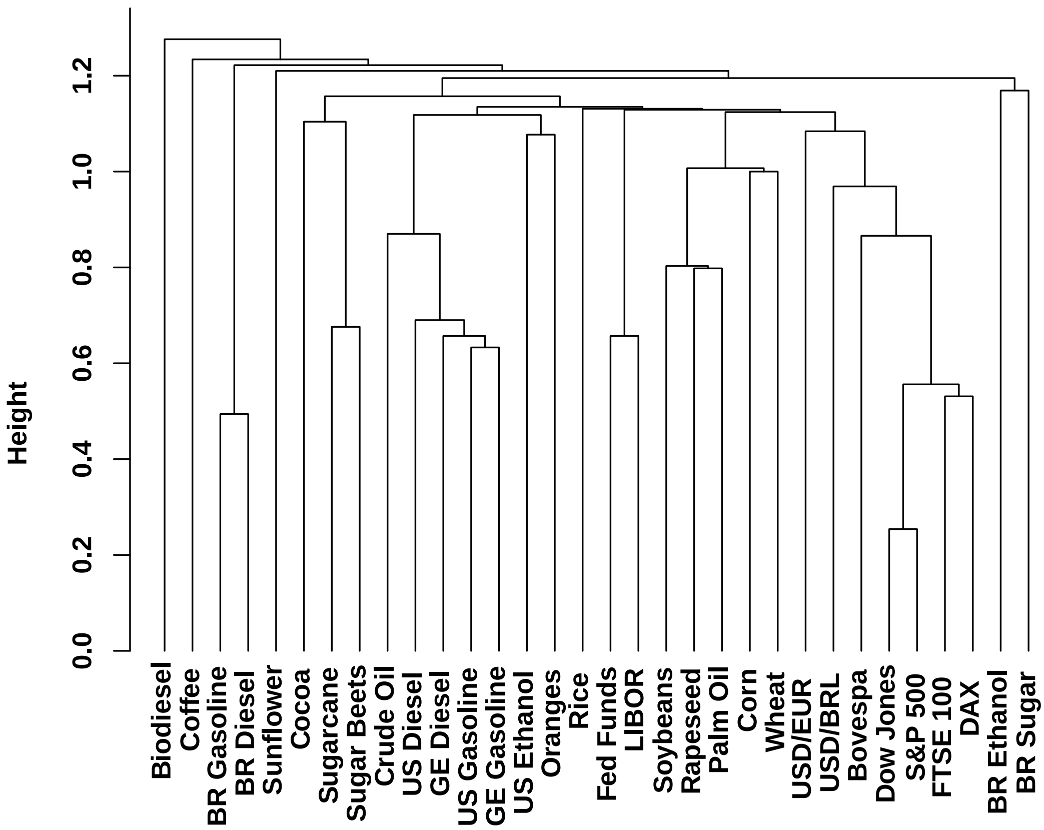
Source: Author's Computation

Figure A.3: HT – Experimental Taxonomy, Period I, Weekly



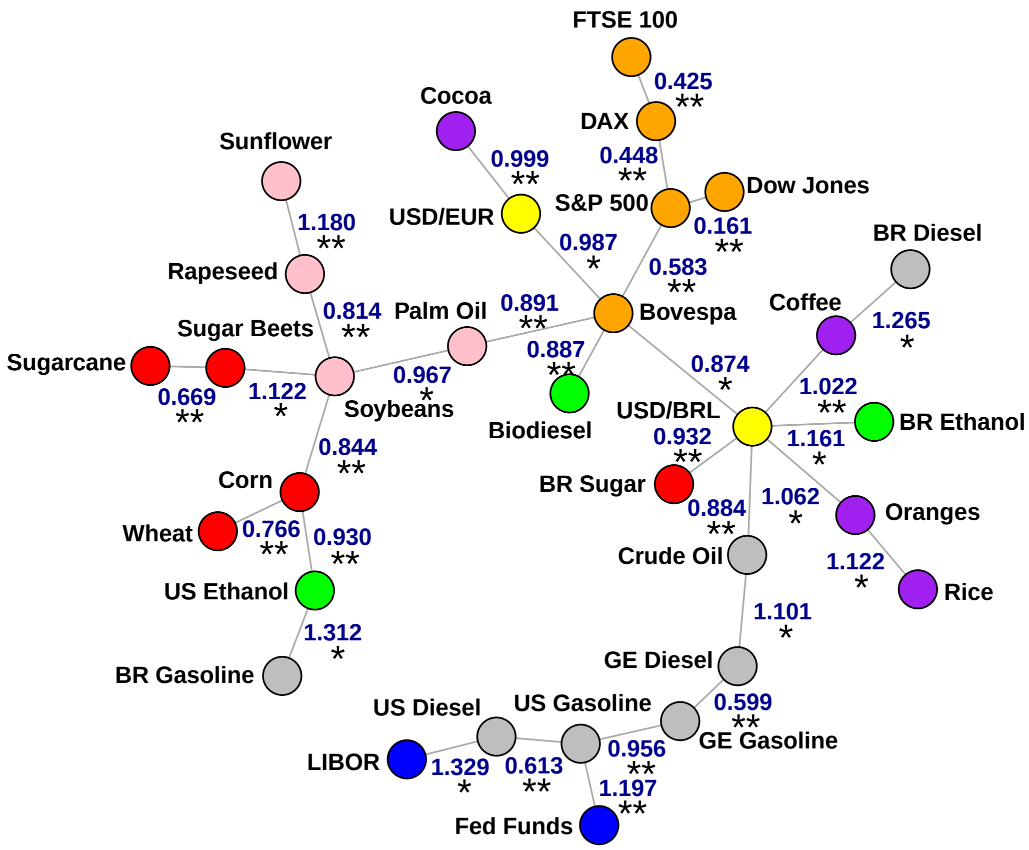
Source: Author's Computation

Figure A.4: HT – Experimental Taxonomy, Period I, Monthly



Source: Author's Computation

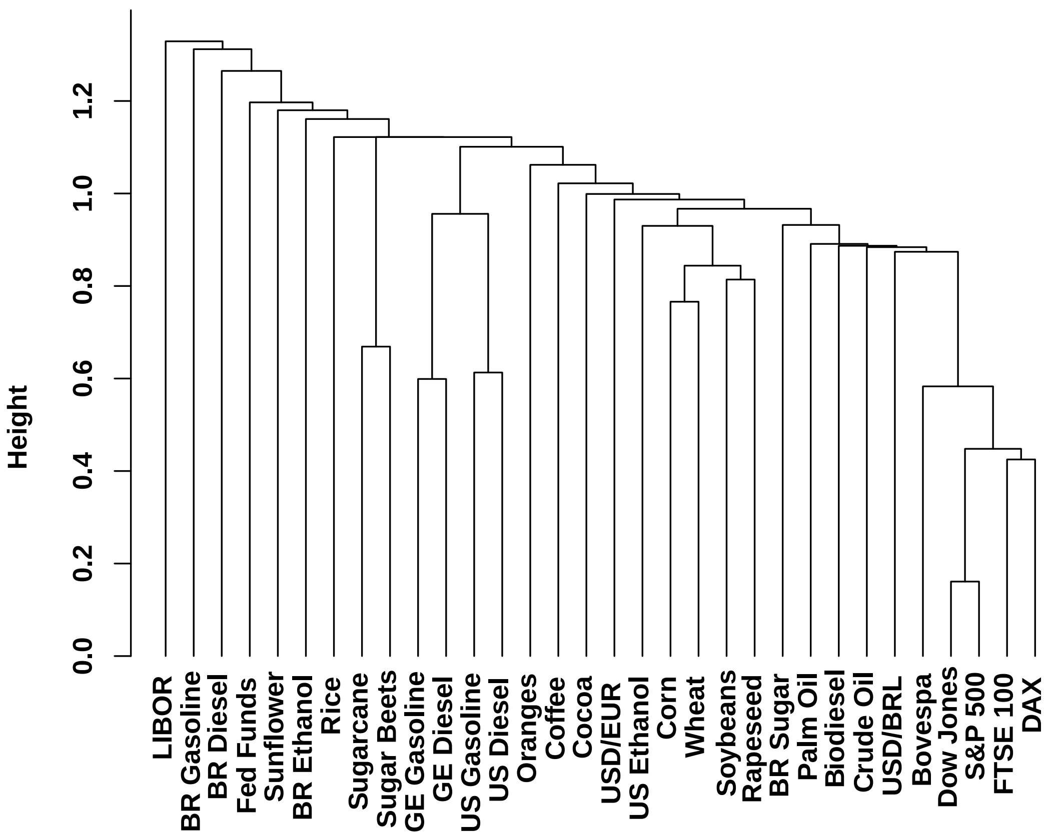
Figure A.5: MST – Experimental Taxonomy, Period II, Weekly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

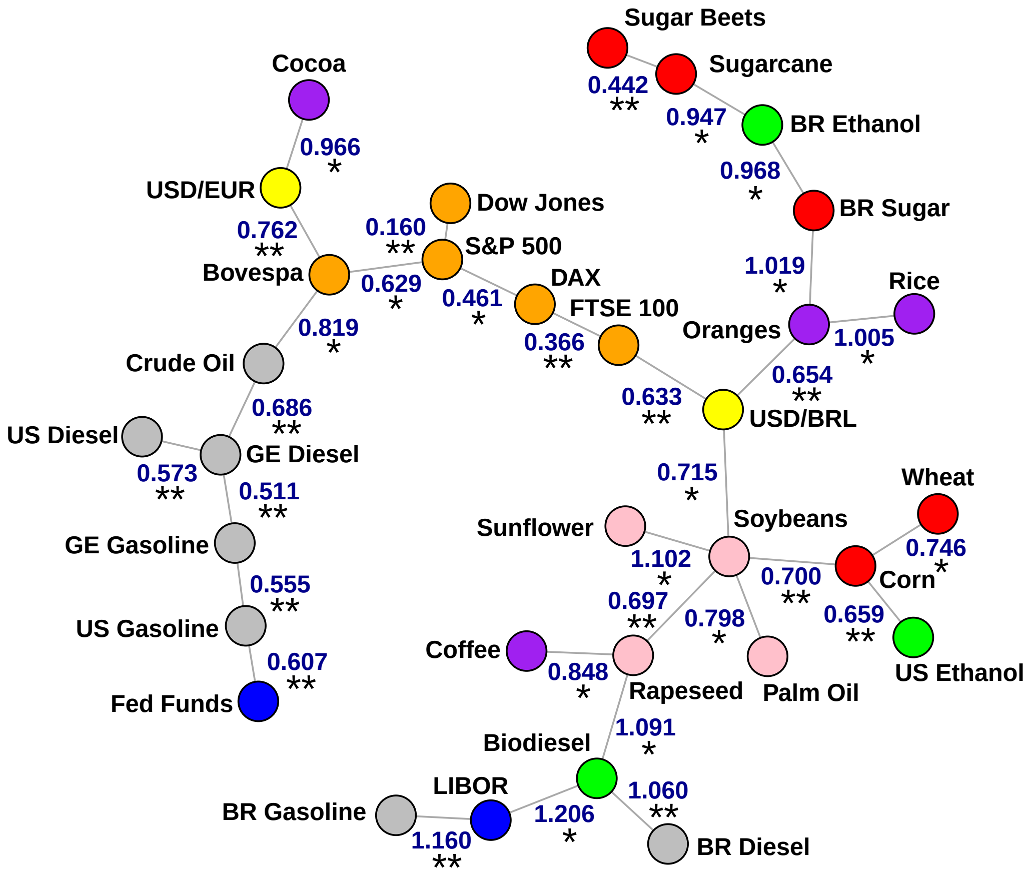
Source: Author's Computation

Figure A.6: HT – Experimental Taxonomy, Period II, Weekly



Source: Author's Computation

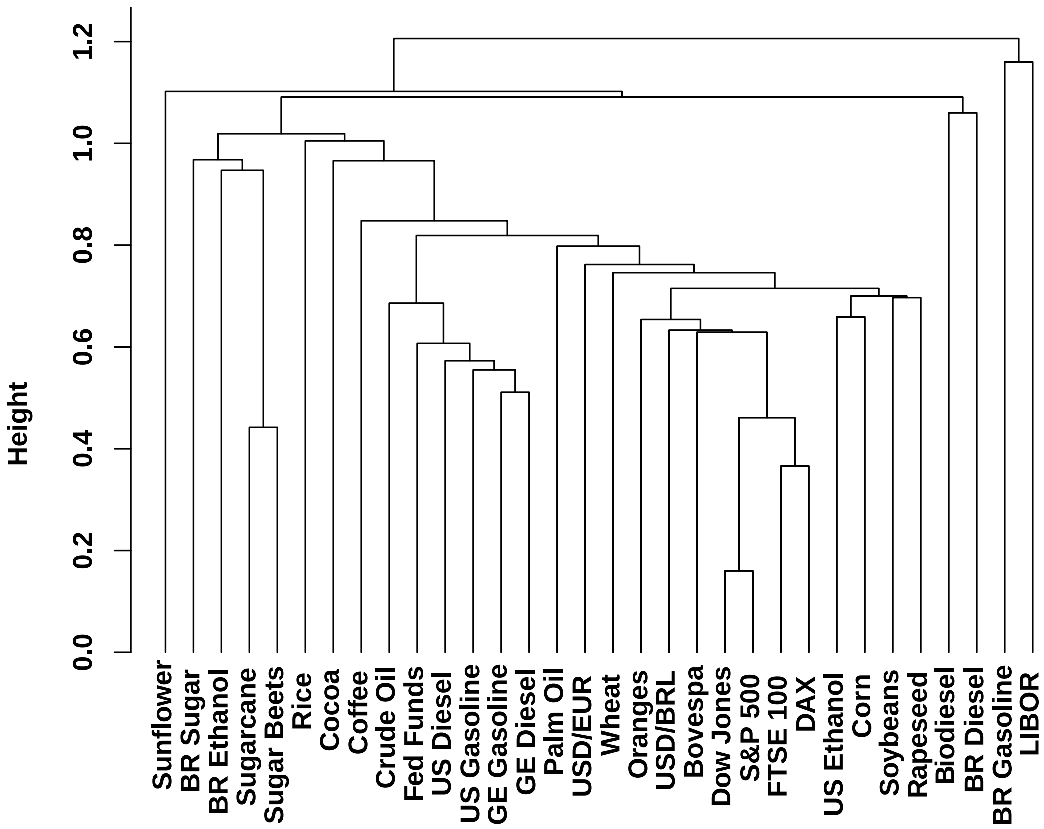
Figure A.7: MST–Experimental Taxonomy, Period II, Monthly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

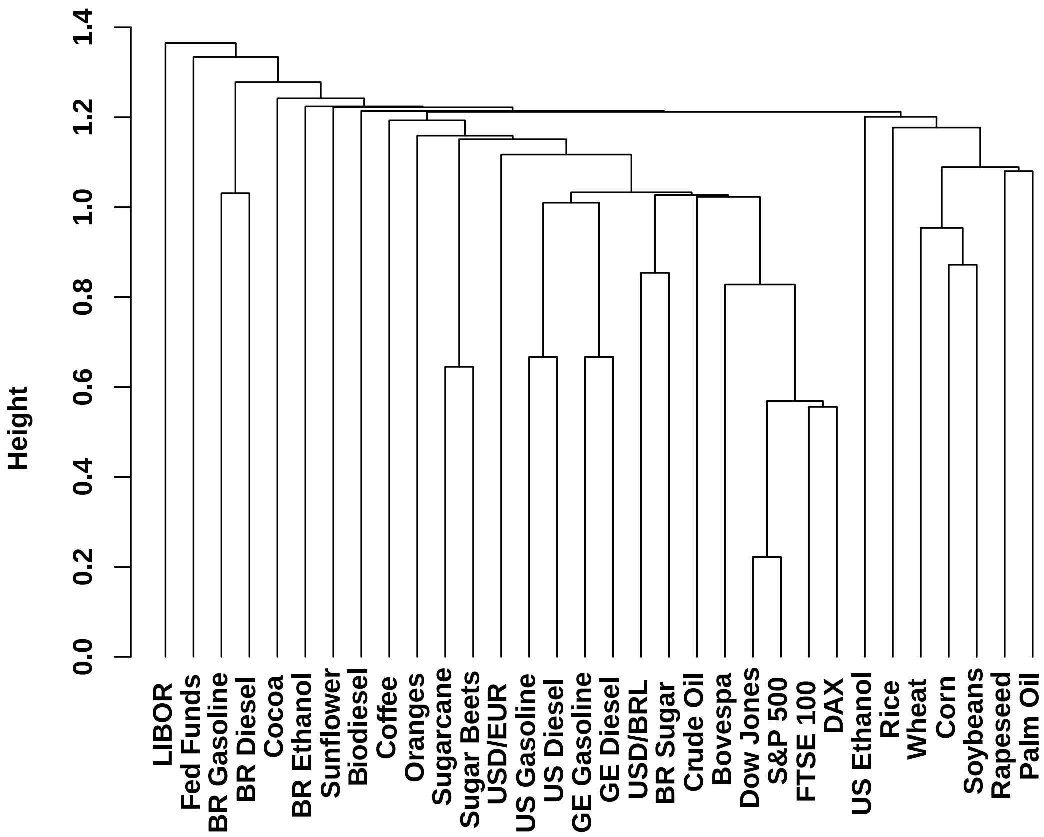
Source: Author's Computation

Figure A.8: HT – Experimental Taxonomy, Period II, Monthly



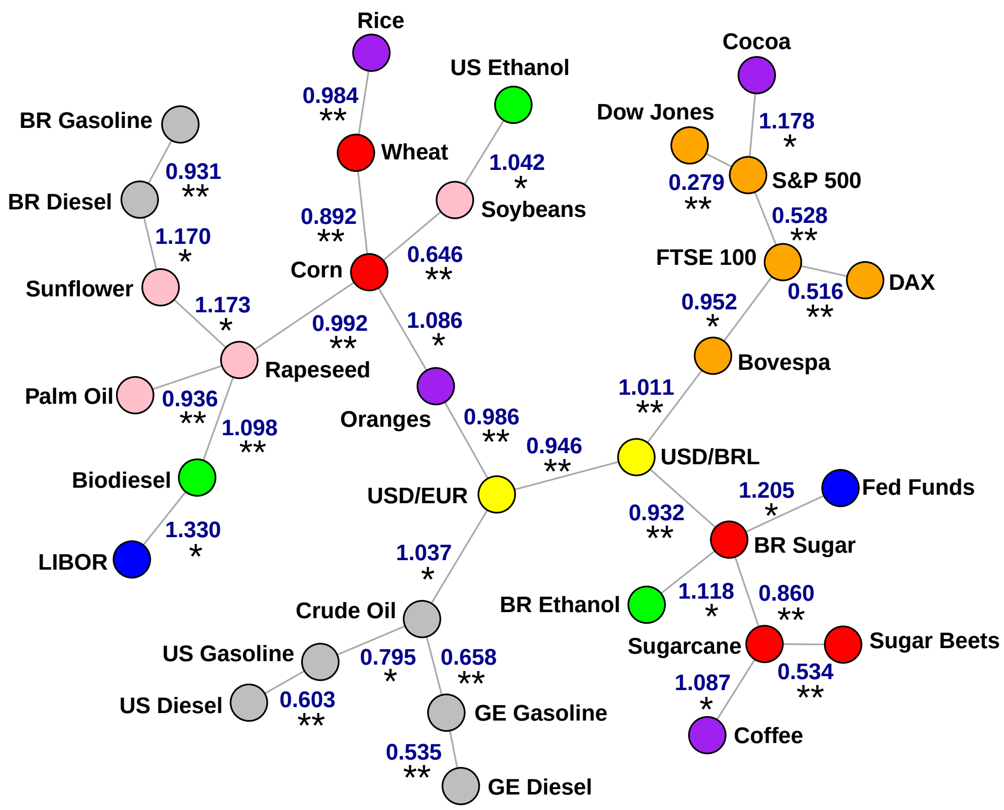
Source: Author's Computation

Figure A.9: HT – Experimental Taxonomy, Period III, Weekly



Source: Author's Computation

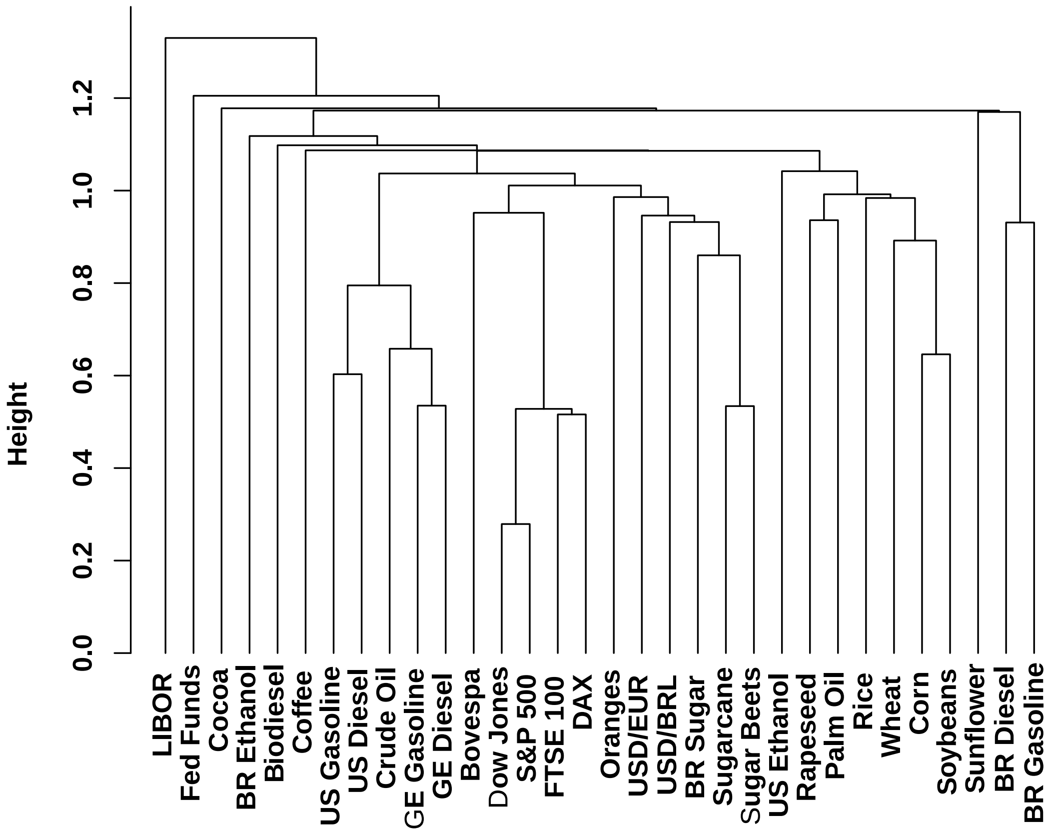
Figure A.10: MST – Experimental Taxonomy, Period III, Monthly



Note: * if Bootstrap Value < 0.5, ** otherwise, Edge Lengths in bold

Source: Author's Computation

Figure A.11: HT – Experimental Taxonomy, Period III, Monthly



Source: Author's Computation