University of Stavanger

Faculty of Science and Technology

MASTER'S THESIS

Study program/ Specialization: Industrial Economics/ Investments and Finance	Fall ~~Spring~~ semester, 20..16.. Open / ~~Restricted~~ access
Writer: Henrik Langdalen	…………………………………………… (Writer's signature)

Faculty supervisor:
Atle Øglend
External supervisor(s):

Thesis title:
The Global Vegetable Oil Market: A Multivariate Cointegration Study

Credits (ECTS):
30 sp.

Key words: Cointegration, Market integration, LOP, Time series	Pages: ..101…………… + enclosure: ..17…….. Stavanger, ..15.12.2016…….. Date/year

Front page for master thesis
Faculty of Science and Technology
Decision made by the Dean October 30th 2009

Preface

This master thesis was carried out at the University of Stavanger (UiS), at the Department of Industrial Economics, Risk Management and Planning, fall semester 2016.

This thesis was written with Professor Atle Øglend as my supervisor. The area of interest is econometric, and statistical modeling of economical values. The reader is assumed to have basic knowledge of statistics.

Stavanger, 2016-12-15

Henrik Langdalen

Acknowledgment

I would like to thank Professor Atle Øglend for giving me this interesting assignment. He has been helpful and constructive in his feedback, which have helped me in my work with this thesis. And I have really enjoyed writing my master thesis at the department of Industrial Economics, Risk Management and Planning, UiS. Every professor I have inquired with, always responded positive and gave me further insight in the fish oil market and the theory of market integration.

H.L.

Summary and Conclusions

This thesis investigates the existence of one global vegetable oil market. We analyze the historic price developments in the period 2000 until the end of 2015 of fish oil, soybean oil, groundnut oil, sunflower oil, rapeseed oil, corn oil, palm oil, palm kernel oil, coconut oil and linseed oil. The existence of one market can be investigated by the means of Johansen cointegration tests. There is statistical evidence which suggests that there is a strong relationship among the oils in the long-run. The quality of the oils and which end-users the oils have, are the main factors which cause the degree of integration. The general oils (soybean, sunflower, rapeseed and palm oil) which are close substitutes in the cooking and margarine manufacturing are cointegrated over the full period. In addition, the law of one price is accepted for these oils. Tests of all ten oils as in one system rejects the law of one price, but suggests cointegration in the long-run. The lauric oils, coconut and palm kernel oil, are considered as cointegrated and forms a separate sub-market within the global vegetable oil market. Hence, the degree of market integration is strongly linked to the qualities of the oils.

Fish oil has risen to become a premium oil compared to the other vegetable oils. During the commodity boom era, fish oil formed a stationary linear relationship with rapeseed oil. The observed integration between the two can be explained by the degree of substitution of fish oil by rapeseed oil as input in salmon feed – fish oil and rapeseed oil was substitutes in this period, despite fish oil's premium qualities as a source of Omega-3. There is a break around June 2011 which caused rapeseed oil (and the other vegetable oils) to diverge from fish oil. The high price of fish oil in the recent years is deemed as a result of high demand of fish oil in a new Omega-3 market for human consumption. Thus, fish oil is no longer considered as integrated with the vegetable oils. The high prices of fish oil has caused the salmon feed costs to increase, as fish oil is an important input in the feed. The demand of fish oil is believed to be high in the future, both from the aquaculture industry and the Omega-3 market. Therefore, we believe that the price of fish oil will be traded at premium compared to the vegetable oils in the future.

Contents

List of Figures

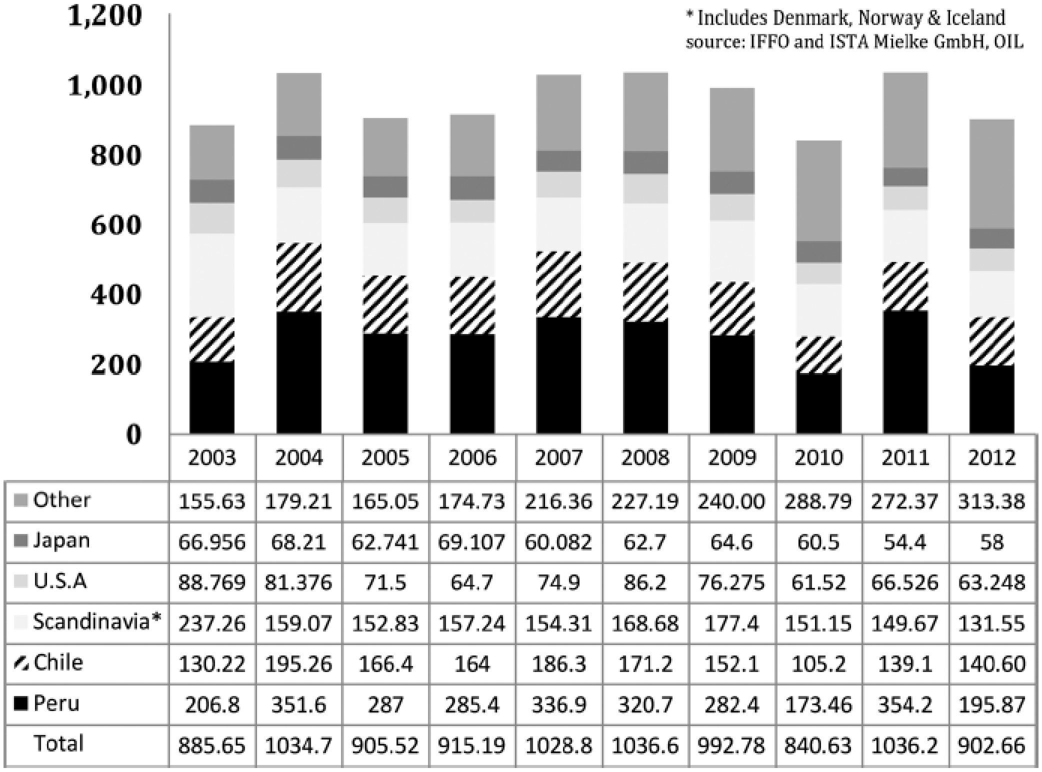
List of Tables

Chapter 1

Introduction

The world's population is increasing by 80 millions p.a. and is expected to reach 9 billions in 2050 (Ytrestøl et al., 2015). Hence, the continual increase in world's population will equalize a greater demand of food in the future. Competition of water, area, feed and energy are all limiting factors for further growth in land-based animalistic protein. Aquaculture, which is defined as the farming of aquatic animals and plants (Naylor et al., 2009), is therefore believed to play an important role in feeding the world. However, Ytrestøl et al. (2015) states that the fast growth in aquaculture production has resulted in concerns among the environmental impact and sustainability of fish farming. Especially for the marine commodities fish meal and fish oil, which have a dominating role in fish feed. In economic theory, a commodity is a basic good that is produced to satisfy needs or wants (Tomek and Kaiser, 2014). The *need* for fish oil has increased tremendously the past decade. In recent years, consumption of fish oil indicates that the awareness of health benefits related to omega-3 fatty acids, of which fish oil is a primary source, has augmented the global demand for fish oil (Shepherd and Bachis, 2014). In addition, the continual growth in farmed fish has further expanded the global demand for fish oil – fish oil has risen from an inferior to a premium oil product. Hence, there is a question of whether fish oil can be considered as a part of a bigger global oil market.

Throughout the last decade or so, global supply of fish oil has remained stable at approximately 1 million tonnes p.a. (Shepherd and Bachis, 2014). Figure 1.1 shows the production of fish oil from the major producing regions. Most of the fish oil come from industrial fisheries in high producing countries such as Peru, Chile, the United States and the Scandinavian countries.



	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
■ Other	155.63	179.21	165.05	174.73	216.36	227.19	240.00	288.79	272.37	313.38
■ Japan	66.956	68.21	62.741	69.107	60.082	62.7	64.6	60.5	54.4	58
■ U.S.A	88.769	81.376	71.5	64.7	74.9	86.2	76.275	61.52	66.526	63.248
■ Scandinavia*	237.26	159.07	152.83	157.24	154.31	168.68	177.4	151.15	149.67	131.55
▨ Chile	130.22	195.26	166.4	164	186.3	171.2	152.1	105.2	139.1	140.60
■ Peru	206.8	351.6	287	285.4	336.9	320.7	282.4	173.46	354.2	195.87
Total	885.65	1034.7	905.52	915.19	1028.8	1036.6	992.78	840.63	1036.2	902.66

Figure 1.1: Global supply of fish oil (in 1000 tonnes) from the five largest producing regions during 2003 to 2012 (Shepherd and Bachis, 2014).

Consumption of fish oil includes human consumption, animal and aquaculture feed and industrial processing. It is the unique content of nutritional properties in fish oil that causes its popularity. While the supply has remained stable, the demand has been questioned to be increasing. Until the commodity boom in the beginning of the 21st century, fish oil has been traded slightly below oils such as soybean oil and rapeseed oil. In the period up to the financial crisis, fish oil was traded at par with those oils. As the prices kicked off in the aftermath of the financial crisis, fish oil diverged from the vegetable oils in June 2011; fish oil was still traded at high prices, whereas the prices on most vegetable oils declined. This price development questions fish oil as a member of a *global vegetable oil market*. In this thesis we test the hypothesis of whether all vegetable oils and fish oil constitute one market. Further we try to shed light on the the relationship between different vegetable oils and fish oil through the means of market integration by the *cointegration* approach.

1.1 The History of Cointegration

Historicaly, the interest for cointegration can be traced back to 1926 through the work by Yule (1926). Yule discovered that correlation between time series sometimes yielded nonsense, or *spurious*, results. Granger and Newbold (1974) found it curious how the contemporary literature often reported high degree of fit in time series regressions when the Durbin-Watson statistics was very low (Durbin-Watson is a test for autocorrelation (Durbin and Watson, 1950)). Until the 1980's, linear regression was commonly accepted as the method to study time series – independent of the time series being stationary or non-stationary. Granger (1981) introduced the concept of cointegration in his paper about time series and econometric models, and proved that linear regression on non-stationary data was harmful. Simply stated, cointegration is a method to study the relationship between time series. Granger (1983) progressed in his study of spurious regression, and in 1983 he stated that two or more non-stationary time series can form a linear relationship which is stationary. This was the root for the first cointegration test developed by Engle and Granger (1987), which could handle non-stationary time series. Prior to the research by Engle and Granger, Dickey and Fuller (1979, 1981) had developed tests for the presence of unit roots which were necessary for the feasibility of the Engle-Granger cointegration test. Later on, Johansen (1988) developed a *multivariate* cointegration framework. These models made it possible to investigate systems of more than two non-stationary variables with cointegration.

Clive W. J. Granger received the Nobel Prize in economics for his research related to non-stationary time series and cointegration (for your information, Robert F. Engle also received the Nobel Prize in economics for his contribution in analyzing economic time series with time dependent volatility in 2003). Today the Engle-Granger method and Johansen framework are the two most common methods for analyzing non-stationary time series. These methods have shown great result in terms of analyzing economical theories such as the *purchasing power parity* (PPP), the *law of one price* (LOP) and *market integration*.

The literature regarding market integration and cointegration has become excessive. Different types of commodities have been studied with the approach of cointegration: Ardeni (1989) performed a test of the long-run relationship of a group of commodities (wheat, tea, sugar, etc.)

with cointegration; Godwin and Schroeder (1991) showed how the distance between markets, market volumes and market types can influence the cointegration test, by an empirical analysis of a regional cattle markets; Gordon and Hannesson (1996) applied both the Engle-Granger and Johansen procedure to test for price relationships between the European and the U.S. cod markets; Asche et al. (2000) applied the cointegration approach to examine market integration in French gas imports.

Fish meal is probably, second to fish oil, the most important ingredient in fish feed. With respect to cointegration, fish meal has received more attention than fish oil in the literature, e.g. Tveterås (2000) and Asche et al. (2013). Despite fish oil's significant role in e.g. salmon feed, the literature is deficient with respect to fish oil in terms of market integration and cointegration with the global edible oil market; thus, this thesis will hopefully help to shed further light on the subject. With respect to the vegetable oil market, there is reported a few interesting studies which are highly relevant for this report. In and Inder (1997) studied eight edible oil prices by cointegration of data from October 1976 until March 1990, and concluded that there was a long-run relationship between the oils. This report can be used for comparison with our results, and therefore an indications of how the vegetable oil market has developed in the last 20 years or so. Williams and Thompson (1984) examined how the global vegetable oil market affects edible oil exporting countries, by a study of soybean oil from Brazil. Goddard and Glance (1989) studied the economical effects supply and demand of vegetable oils had on Canada, the U.S. and Japan.

1.2 Objectives

Market integration is the base for the approach of this thesis. The extent of a market is studied by the existence of cointegration relationships between the oils, and related to the law of one price. In the last decades, the development of efficient cointegration methods has made it possible to execute empirical testing of market integration. This thesis explores the long-run relationship between the fish oil market and different vegetable oil markets in a dynamic framework by employing multivariate cointegration analysis on historic price data. Possible dominant market(s) that drives the prices of other markets is detected by the mean of Granger-causality test and weak exogeneity test.

The purpose of this thesis is two-folded. Firstly, the historic developments of fish oil, coconut oil, corn oil, groundnut oil, linseed oil, palm oil, palm kernel oil, rapeseed oil, soybean oil and sunflower seed oil prices in the time period 2000 to 2016 are evaluated empirically. We will try to answer the following questions:

- Is there any statistical evidence of *one* global vegetable oil market?
- How is fish oil related to the vegetable oils?
- Which oil prices, if any, are cointegrated?
- If the oils are integrated, how are they related to each other with respect to exogeneity and the Law of One Price?

Secondly, fish oil and the vegetable oils prices are evaluated with respect to the share of usage in fish feed. The thesis has the following configuration:

- **Chapter 2** provides a description of the different vegetable oils and fish oil in terms of historic price development and attributes. In addition, a brief description of a typical salmon feed is included to shade some light on its important features, and how the increased fish oil price has affected the feed costs.
- **Chapter 3** introduces the economic theory of the law of one price and market integration, and implications of these terms in a study of commodity prices.
- **Chapter 4** describes the econometric methods applied to investigate the relationship between the different oils. Especially, the Dickey and Fuller (1979, 1981) test, Johansen (1988) cointegration test and Toda and Yamamoto (1995) causality test are elucidated.
- **Chapter 5** includes the empirical models applied and the results of the analysis. Following each result a complimenting discussion is provided. Both bivariate and multivariate cointegration tests are performed on the oil prices, including specific restrictions on the relationship among the oils.
- **Chapter 6** summarizes the results with a discussion and conclusions.

Chapter 2

The Global Oil Market

Fish oil is a necessary ingredient in fish feed. The global demand for fish oil has pushed the prices to all time high in the recent years, whereas the prices on most vegetable oils have diverged. This chapter has the purpose of giving the reader a thoroughly understanding of why fish oil is a crucial ingredient in fish feed, followed by a description of fish oils primary markets throughout history. In addition, fish oil and the vegetable oils applied in the study of the *global* oil market, are presented in terms of historic price developments and its most common applications. The effect of high fish oil price on fish feed costs will be elucidated throughout this chapter.

2.1 Salmon Feed

Farmed salmon diet is dry feed shaped as pellets. All salmon feeds must contain a certain amount of fish oil (fats), vegetable ingredients, fish meal (proteins) and "other" (Laksefakta, 2016). Without going into details about the specific composition in fish feed, a short description of the main ingredients is presented below:

- **Fish oil** is fat from forage fish (also known as industrial fish), not intended for human consumption, which is rich of the omega-3 fatty acids eicosapentaenoic (EPA) and docosahexaenoic (DHA). Ytrestøl et al. (2015) states that from one kilogram forage fish around 50 to 100 grams of fish oil can be produced.

- **Fish meal** is another by-product from the fish, e.g. fish heads and other parts not consumed by humans. Fish meal is a premium source of proteins and minerals. Ytrestøl et al. (2015) have estimated 1 kilogram industrial fish to result in around 230 gram fish meal.

- **Vegetable ingredients** in fish feed are produced from plants such as soy, sunflowers, rapeseed and wheat (Laksefakta, 2016). They are sources of proteins, carbohydrates and fats.

- "**Others**" include vitamins, pigments, minerals and amino acids (Ytrestøl et al., 2015).

Salmon feed composition will differ with respect to feed types and availability and price on the different ingredients. Ytrestøl et al. (2015) states that in 2013 approximately 70 % of the feed was composed by vegetable ingredients, whereas the remaining 30 % was of marine origins, with fish oil at 11 % and fish meal at around 19 %, Figure 2.1. The oil content in the feed amounted to 30 %, with vegetable oils at approximately 19 % and fish oil around 11 %. It must be emphesized that these shares are not exact, but they are in compliance with other sources such as Tacon and Metian (2008). Currently fish oil is the only economically feasible source of the fatty omega-3 rich acids EPA and DHA, hence its dominating position in aqua feed.

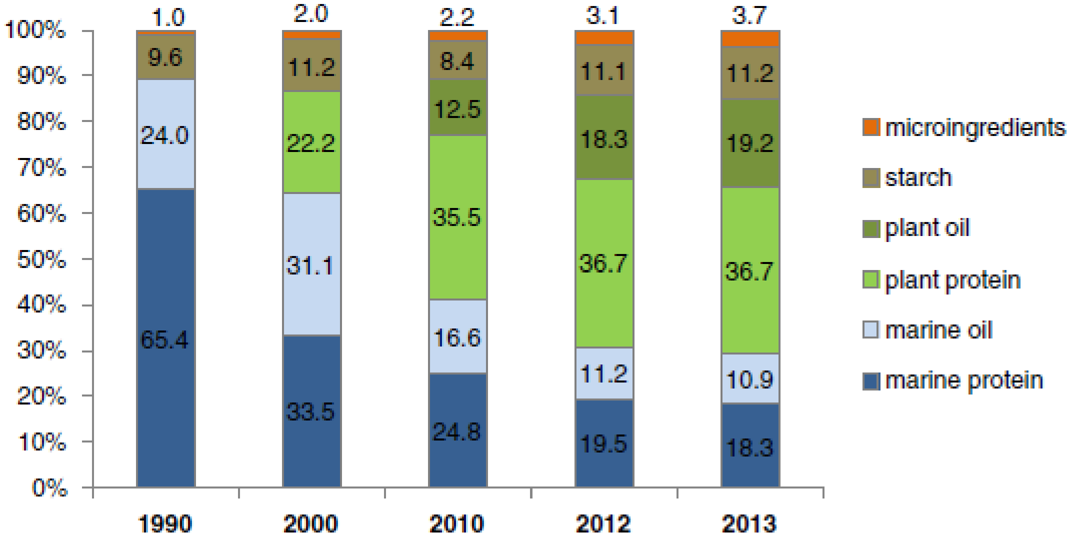


Figure 2.1: Nutrient content in Norwegian salmon feed from 1990 to 2013 (Ytrestøl et al., 2015).

One of the main cost drivers in aquaculture is salmon *feed*. In Table 2.1 we see that feed was the most significant operational cost in Marine Harvest for 2014 and 2015. It counted for almost 50 % of the operational costs in 2014, and around 47 % in 2015 (Marine Harvest ASA, 2015, 2016). Increased prices of marine ingredients have affected the feed cost negatively for the salmon producers, such as Marine Harvest. The transition from fish oil to vegetable oil as the primary oil input in salmon feed has, among other reasons, made it possible to increase salmon production despite the limited supply of fish oil. From the oil shares and total feed volumes provided by Ytrestøl et al. (2015) (see Figure 2.1) and Norwegian salmon production volumes reported by Statistisk Sentralbyrå (2015), we can illustrate this trend with an overview of the salmon production, fish feed usage and fish oil usage in Norway between 2010 and 2013: In Figure 2.2, the total salmon feed usage (red area) in Norway between 2010 and 2013 is presented in the same figure as the total fish oil consumption (green area) and salmon production (blue area). Both salmon production and total feed usage in Norway increased from 2010 to 2012, in contrast with the total usage of fish oil which had a slow decreasing trend. Hence, the vegetable oil share has increased to make up the necessary feed volume. For completeness, total feed and fish oil consumption and salmon production decreased between 2012 and 2013.

Table 2.1: The feed variable is the main cost driver in the Norwegian operational costs in Marine Harvest, 2015 and 2014 (Marine Harvest ASA, 2015, 2016). Numbers are for the gutted weight equivalent (GWE).

	2015		**2014**	
Variable	**NOK**	**%**	**NOK**	**%**
Feed	13.34	46.7	12.35	48.1
Primary processing	2.67	9.4	2.62	10.2
Smolt	2.67	9.4	2.28	8.9
Salary	1.62	5.7	1.49	5.8
Maintenance	0.94	3.3	0.89	3.5
Well boat	0.95	3.3	0.98	3.8
Depreciation	0.78	2.7	0.76	3.0
Sales & Marketing	0.62	2.2	0.62	2.4
Mortality	0.44	1.5	0.34	1.3
Other	4.47	15.7	3.34	13.0
Total	28.54	100.0	25.69	100.0

To illustrate the monetary effect of the shift from marine oil to vegetable oils, we can calculate the cost of oils in salmon feed. Lets say that there was no increase in vegetable oil share between 2010 and 2013, and that vegetable oils are represented purely by rapeseed oil (which is the main vegetable oil used in fish feed). The prices of fish oil and rapeseed oil, provided by ISTA Mielke GmbH (2016), in this example are taken as the average prices for 2010 and 2013. Fish oil

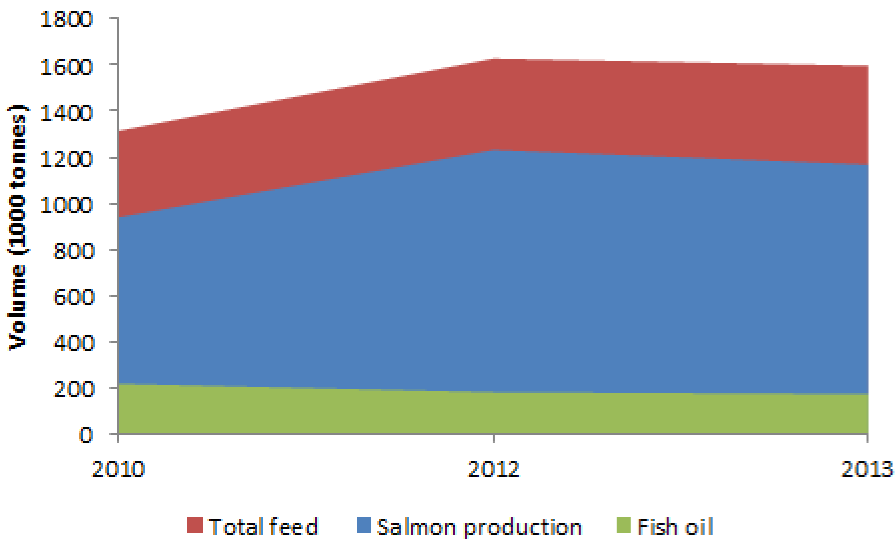


Figure 2.2: Salmon production in Norway compared to total fish feed and fish oil used in 2010, 2012 and 2013 (Ytrestøl et al., 2015; Statistisk Sentralbyrå, 2015).

was traded at 1121 USD/tonnes in 2010 compared to 2042 USD/tonnes in 2013. Rapeseed oil costed 1013 USD/tonnes in 2010 and 1081 USD/tonnes in 2013, at average. If the share of vegetable oils (rapeseed oil) remained constant at 12.5 % (2010-level) throughout 2013, we assume that content of fish oil increased from 16.6 % to 17.6 % to cover the reduction in vegetable oils, instead of the actual share of fish oil in 2013 of 10.9 %. Simultaneously, the feed consumption increased from approximately 1.3 mill. tonnes to 1.6 mill. tonnes. Without an increased vegetable oil share in fish feed, there would have been approximately an additional 100 mill. USD in expenses for the Norwegian salmon farmers in 2013 due to increasing fish oil prices, as illustrated in Figure 2.3.

2.2 Fish Oil

Marine oils are used in different markets such as industrial processing, animal feed, aquaculture, food and hydrogenation and the omega-3 market (Bimbo, 2013). An overview of the fish oil market from 1990 to 2015 is given in Figure 2.4. Over the last 20 years the marine oil market has progressed. Bimbo (2013) described the changes in the marine oils market with three main phases: (1) The primary market in 1990 was in hydrogenated form for use in margarine and baking fats; (2) the second phase was ignited by an awareness of the disadvantage with *trans* fatty

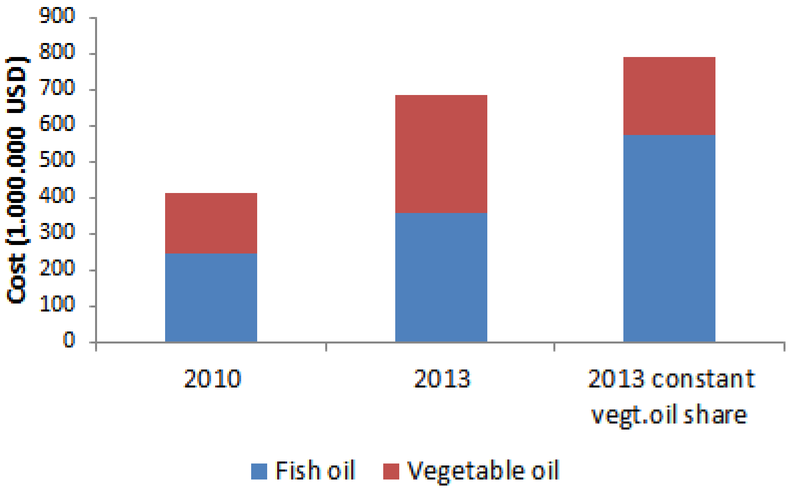


Figure 2.3: An increased fish oil share and constant vegetable oil share in fish feed will result in increased feed costs, due to the increasing fish oil prices.

acids, and the increased aquaculture production; (3) In recent years the fish oil price has increased, and aquafeed producers have looked for substitutes, and there are indications that the omega-3 market could be an important *new* primary market for fish oil. Bimbo (2013) estimates that the Omega-3 market represented around 10-12 % of the global consume of the marine oils. Between 2005 and 2010, the Omega-3 market gained approximately 1-2 % p.a. Hence, if this trend continue in the future, the Omega-3 market can become a primary consumer of marine oils in the future.

Fish oil is a primary source of the omega-3 fatty acids eicosapentaenoic (EPA) and docosahexaenoic (DHA), and its use in feed ensures a healthy product for the end consumer. The limited supply of oil has changed the feed composition the last two decades. In Figure 2.1 a typical composition of Norwegian salmon feed was presented. There has been a shift from marine ingredients towards plant (vegetable) oils and proteins. In 1990 salmon feed was composed with approximately 90 % marine ingredients (fish meal and fish oil), while approximately 30 % marine ingredients were found in the diet in 2013. The reduction in marine ingredients have resulted in an equally large increase in use of plant proteins and plant oils. Marine Harvest reported in their yearly report "Salmon Farming Handbook 2016", that their average Norwegian salmon feed consisted of 9 % fish oil and 21 % vegetable oil in 2015 (Marine Harvest ASA, 2016). Thus, the trend outlined by (Ytrestøl et al., 2015), of reducing fish oil shares by increasing veg-

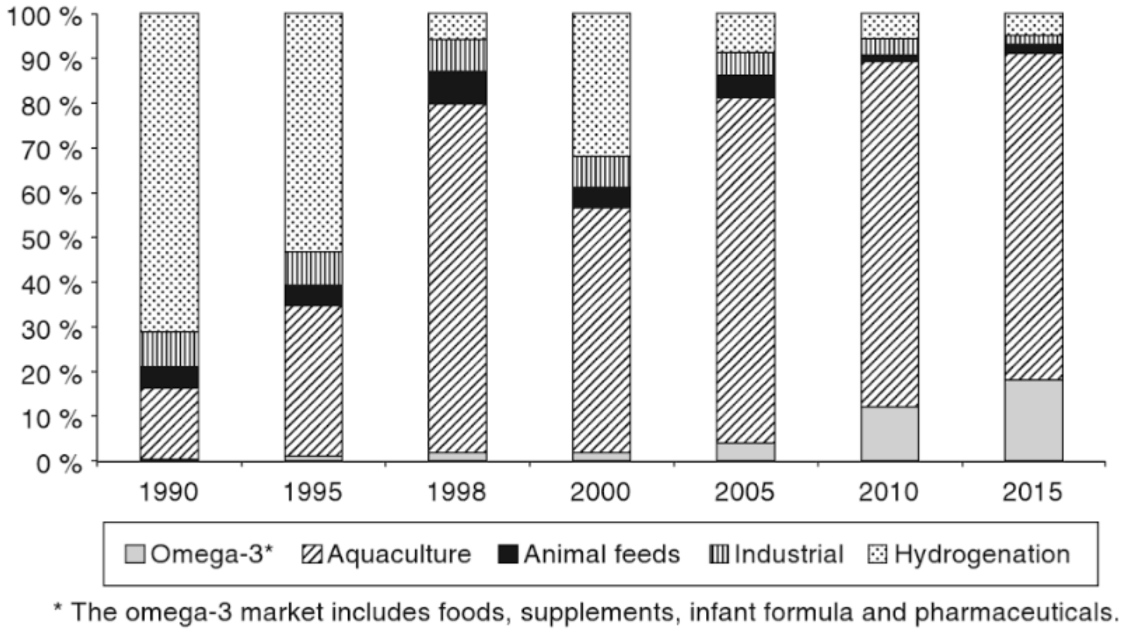


Figure 2.4: An overview of the fish oil market structure (Bimbo, 2013).

etable oil shares, is supported by the industry.

The main source of fish oil is reduction fishery and fish by-products turned into fish oil (and fish meal) (Shepherd and Bachis, 2014). The global supply, reported in Figure 1.1, has remained relatively stable at below 1 million tonnes p.a. between 2003 and 2012. And the major producing regions are Peru and Chile, followed by the Scandinavian countries, the U.S. and Japan (Shepherd and Bachis, 2014). In addition to demand of fish oil in the aquaculture and omega-3 market, another important parameter has great influence on the fish oil supply and price - the El Niño phenomena. These events can be explained by increased temperature in the sea surface water in the Southeast Pacific, which results in lower forage fishery in these areas. We study historic prices of fish oil between 2000 until 2015, therefore can the El Niño events in 2006-07 and 2009-10 reported by Asche et al. (2013), be of interest in the upcoming analysis.

The market of share of salmon feed production in Norway has changed since the 1998. Market shares of feed producers in Norway, reported by Marine Harvest ASA (2016), are presented in Figure 2.5. The Herfindahl-index, which is a measure of competition in a market, can be calculated as:

$$H = \sum_{i=1}^{N} s_i^2 \tag{2.1}$$

where N is the number of firms, and s_i are the firms respective market share. From the shares in Figure 2.5, the salmon feed market in Norway is considered as highly concentrated (oligopoly-tendencies in the market), with a Herfindahl-index of 2742 and 2664 for 1998 and 2015, respectively. It is beyond the scope of this thesis to investigate the market shares, but it must be emphasized that high concentration of market power could effect the feed price in the salmon industry.

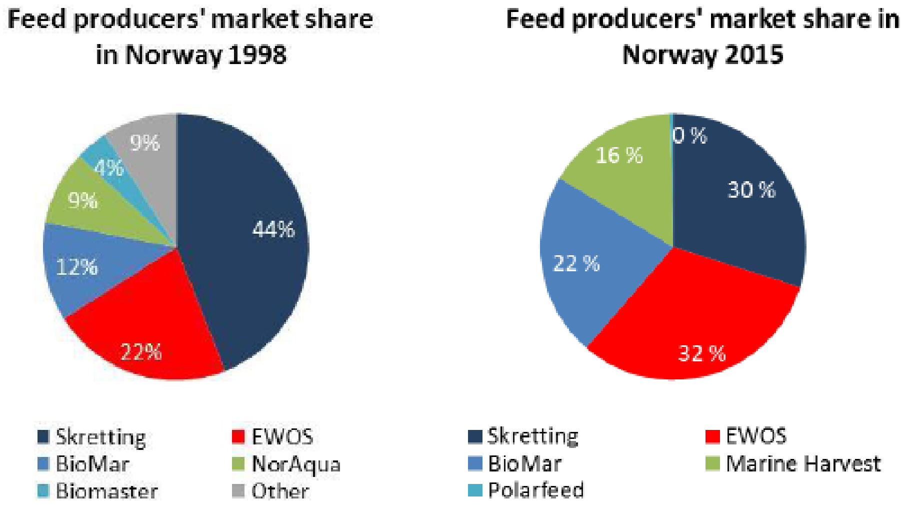


Figure 2.5: The salmon feed producers market share in Norway (Marine Harvest ASA, 2016)

2.2.1 Introducing The Fish Oil Price

All commodity prices in this thesis have been extracted from ISTA Mielke GmbH (2016) and are given as monthly average prices in USD/tonnes. In Figure 2.6 the fish oil price over the time period January 2000 to December 2015 is plotted. Historically, fish oil has been traded at around the prices in year 2000 at approximately 300 USD/tonnes up to the start of the commodity boom (Bimbo, 2013). In the early 2000s there was a slight shift in the prices. This period is known as the *commodity boom* (2003-2008) (World Bank, 2008), which ended just before the financial crisis in 2007-08. In September 2007 the global financial crisis exploded, the price went to a contemporary all-time high, before the price "collapsed" at the end of 2008. In the aftermath of the crisis, the fish oil price continued its increasing trend from the commodity boom era. A monthly all-time high was reached in April 2013 at 2400 USD/tonnes. Since then, fish oil prices have been fluctuating around 2000 USD/tonnes, and stabilized at around 1700 USD/tonnes in the last months of 2015. The period from 2009 until 2015 will be referred to as the *post-financial crisis* period. In the next section the historical price development will be further discussed with respect to other vegetable oils.

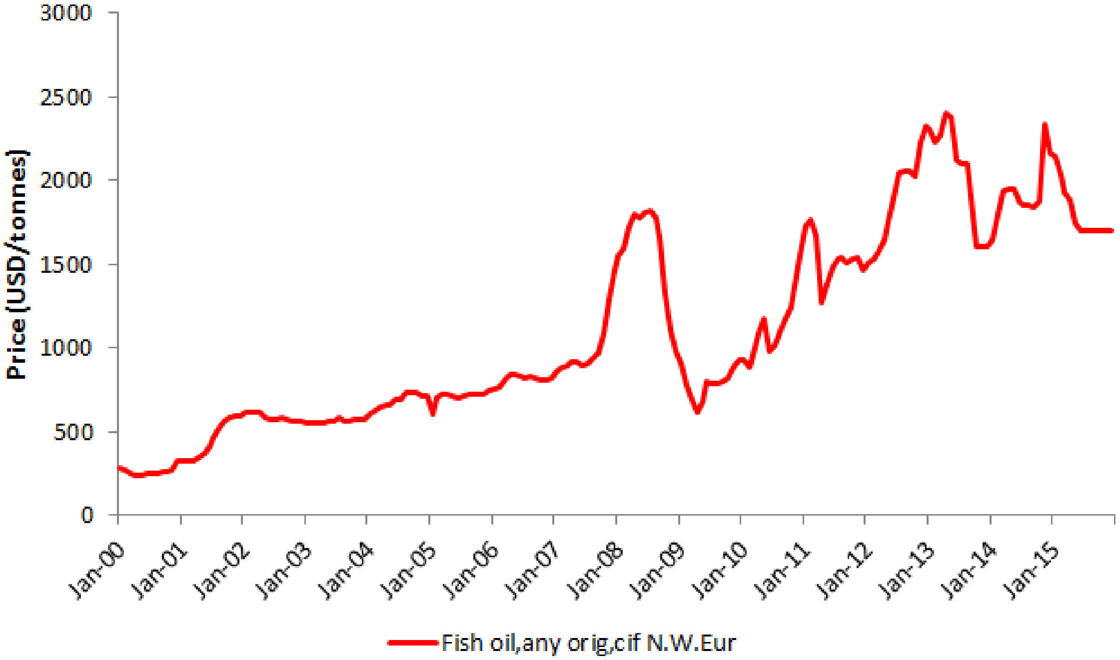


Figure 2.6: Historic overview of the fish oil price from 2000 to 2016.

2.3 The Oil Market

With time the fish oil share in fish feed has decreased. In 2000 and 2013 fish oil represented 31 % and 11 % of the total feed, respectively (Ytrestøl et al., 2015). The total oil (marine and vegetable) share in the feed has remained relatively stable around 30 %. Rapeseed oil is the primary substitute to fish oil in fish feed. Thus, among the vegetable oils, rapeseed oil is of greatest interest related to the salmon farming industry. However, in this thesis we have included eight vegetable oils in addition to fish oil and rapeseed oil to get a better understanding of the *global oil market*. In the report by USDA (2016), the following oils are assumed as "major" vegetable oils with respect to production and consumption: coconut, cottonseed, olive, palm, palm kernel, peanut (groundnut), rapeseed, soybean and sunflower. Table 2.2 lists the ten oils studied in this thesis.

Hence, seven of the nine major vegetable oils are included. Olive oil and cottonseed oil are left out, and replaced with linseed oil and corn oil. Corn and linseed oils are included since they are produced in smaller quantum than the others, and have the potential to reflect any relationship between oils of different supply volumes. We will look at the price development of all the ten oils between January 2000 until December 2015.

The oils have different origins and (slightly) different primary markets. In Table 2.2, *price type* indicates whether the oil price is given as a *free on board* (FOB) or a *cost, insurance and freight* (CIF) price. CIF-contract is more costly than a FOB-contract; CIF includes the cost of transportation and insurance of the goods, which is excluded in a FOB contract. The difference in *price type* is argued to weaken the comparison. Still, the prices can be assumed as a good representation of the global market through the historic development in price and by the fact that they are the *major* vegetable oils in a global perspective. In addition, the quality of the oils will influence the prices, and they will never be identical. Third column of Table 2.2 gives the geographical location which the oil is shipped from. Note that a CIF-contract has two locations; where the first being the origin location and the second is where the contract ends. Meaning, if a Norwegian producer needs groundnut oil he trades groundnut oil on a CIF-contract which ships the oil to Rotterdam. Additional transportation from Rotterdam upto Norway will be an extra cost of transportation for the producer. Moreover, the prices of palm kernel oil and coconut oil includes a shipment from Malaysia and Philippines to Rotterdam, respectively. Obviously, this boosts the price and complicates the comparison of the oil prices.

Table 2.2: An overview of the different oils used in this study, in terms of contract/price type and origin/destination (ISTA Mielke GmbH, 2016).

Oil Type	**Price Type**	**Description**
Fish oil	CIF	Any origin, North West Europe
Rapeseed oil	FOB	Rotterdam
Soybean oil	FOB	Decature, USA
Groundnut oil	CIF	Any origin, Rotterdam
Sunflower oil	FOB	North West Europe
Corn oil	FOB	Midwest, USA
Palm oil crude	CIF	Nort West Europe
Palm kernel oil	CIF	Malaysia, Rotterdam
Coconut oil	CIF	Philippines, Rotterdam
Linseed oil	Ex-tank	Any origin, Rotterdam

Before an introductory analysis of the oil prices, it can be of value to characterize the vegetable oils in brief terms (Wikipedia, 2016):

- *Rapeseed oil* is an edible vegetable oil, often used in animal feed and as an input in biodiesel.
- *Soybean oil* is one of the most produced edible oils, and primarily used in baking and frying.
- *Groundnut oil*, also known as peanut oil, common in cooking as frying oil.
- *Sunflower oil* is an edible oil, common in cooking, and is considered as healthy.
- *Corn oil* is used in cooking and other industrial products, e.g. soap and salve.
- *Palm oil* is second to soybean oil the most common edible oil, often used as a substitute for butter in recipes and as replacement for trans fats.
- *Coconut oil* comes from coconut kernels, and has a high degree of saturated fats (hence, not recommended to consume too much of it), and is widely used in snacks such as popcorn. In addition it is used in soaps (In and Inder, 1997).
- *Palm kernel oil* is comparable to coconut oil in properties, and often used in commercial food processes.
- *Linseed oil* is suitable for human consumption dependent on the production process. Linseed oil is high in nutrition, and is considered to have health benefits.

While the oils can be described as close substitutes by their applications, there is likely to be a difference depending on the end-users preference. And it is naturally to divide the oil into different groups depending on their quality. We define rapeseed, soybean, sunflower and palm oil as *Group A* or the general oils as In and Inder (1997) denoted them. These oils are considered as close substitutes within cooking oils, margarine and compounded fats (In and Inder, 1997). Hence, these oils are believed to co-move over time. The second group is the unhealthy oils, denoted *Group B*, coconut oil and palm kernel oil. The two latter oils are also known as *lauric* oils, which have different qualities than the other edible oils, and are often suitable for making soaps (In and Inder, 1997). The third group, will contain groundnut, linseed and corn oil - denoted

Group C. These three oils are a mixture of the two prior groups, and have multiple end-users and are therefore hard to categorize in one of the previous groups.

Production volumes must be considered in relationship of demand. The yearly production of some of the vegetable oils between 2010-2013 are given in Table 2.3 (FAO, 2016). Palm oil and soybean oil have the largest production volumes in the group of oils. Hence, there is a larger demand for these oils globally. Rapeseed oil, palm kernel oil and sunflower oil can be considered as medium production quantities. Fish oil, given the supply volumes in Fig. 1.1, must be considered as an inferior oil in terms of quantity compared with most of the vegetable oils in our group.

Table 2.3: Global Production of the vegetable oils in Table 2.2 in million tonnes (FAO, 2016).

Oil Type	**2010**	**2011**	**2012**	**2013**
Rapeseed oil	22.84	23.13	24.23	24.69
Soybean oil	40.69	41.92	42.00	42.66
Groundnut oil	5.62	5.41	5.30	5.18
Sunflower oil	12.64	13.37	14.84	12.59
Palm oil crude	45.77	49.42	52.46	54.38
Palm kernel oil	22.84	23.13	24.23	24.69
Coconut oil (Copra)	3.86	3.18	3.32	3.22
Linseed oil	0.55	0.55	0.59	0.56

Figure 2.7 shows the historic prices for soybean oil, rapeseed oil, palm oil and fish oil. All four oils have somewhat similar price-trend until 2011-12 where fish oil became a premium oil product compared to the other three, with respect to price. Comparing Figure 2.8 with Figure 2.7, the historic price developments of palmkernel oil, corn oil, sunflower oil and coconut oil are similar to the other vegetable oils in Figure 2.7. Groundnut oil and linseed oil which have some major deviation from the common trend, differs the most compared to the vegetable oils. Overall, the vegetable oils seem to follow the same path in the long-run, but with different contemporary fluctuations. Is this long-run path an indication of an integrated market?

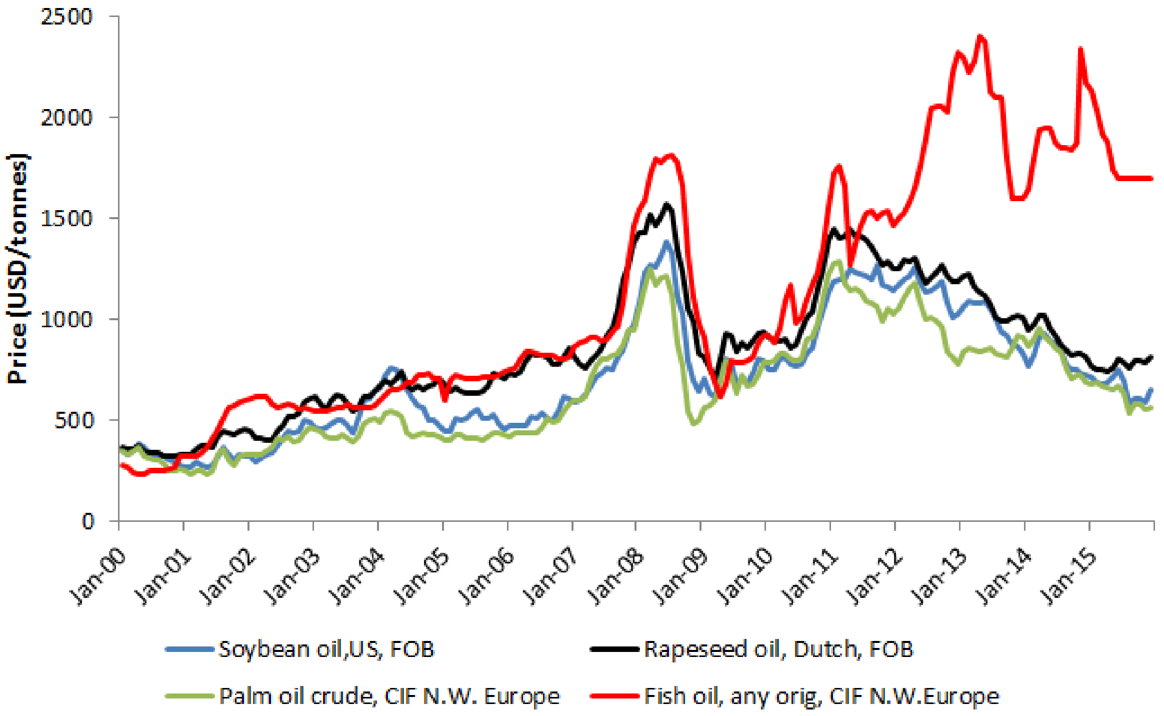


Figure 2.7: Historic price developments of soybean oil, rapeseed oil, palm oil and fish oil (ISTA Mielke GmbH, 2016).

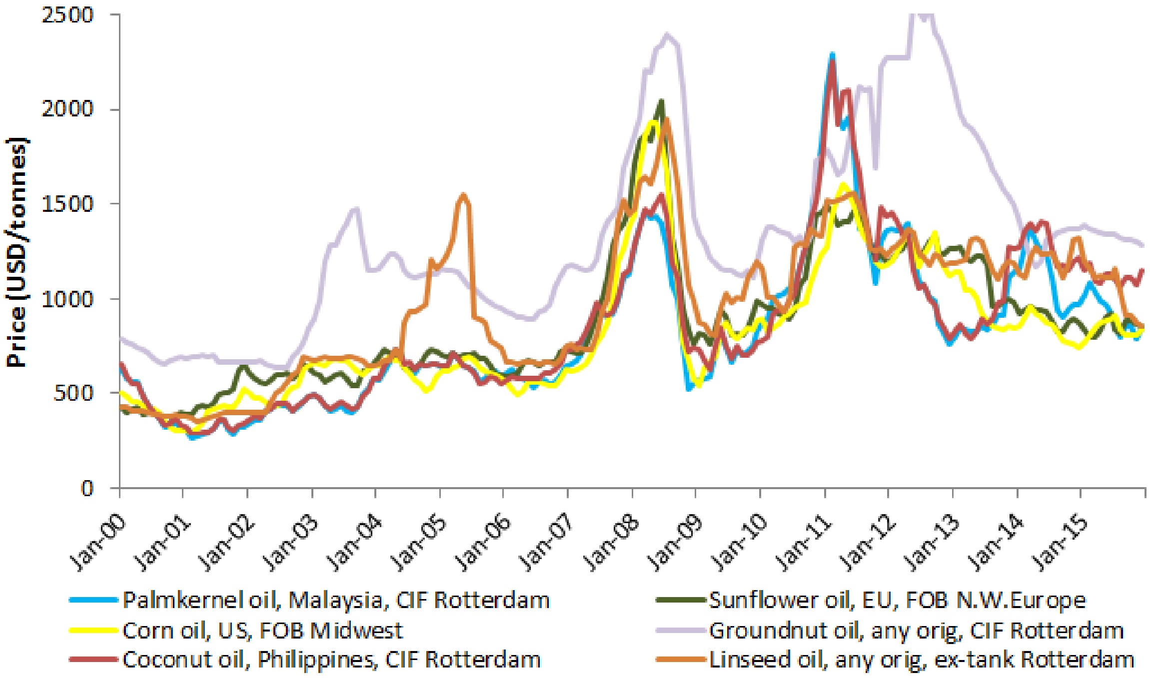


Figure 2.8: Historic price developments of palm kernel oil, sunflower oil, corn oil, groundnut oil, coconut oil and linseed oil (ISTA Mielke GmbH, 2016).

Chapter 3

Market Integration and the Law of One Price

The big question in this thesis is whether or not fish oil should be considered as a part of the global oil market. In order to evaluate this hypothesis we need to understand the underlying economics. First in this chapter there will be a delineation of the market concept, followed by an empirical description of market integration. Finally, the concepts of product aggregation and composite commodities are outlined.

3.1 Definition of a Market

Supply and demand is the root in most definitions of a market. The laws of supply and demand dictates the relationship between supply and demand in a market, and how they are related to the price of goods or services. If there is high demand, prices tend to rise. If there is a surplus of a good in the market prices tend to fall. The relationship between a quantity demanded and supplied of a commodity (ceteris paribus) yields a price which represents the market equilibrium (Tveterås, 2000). A change in the price of one product yields a change in demand for another. In what direction the change will be is related to if the products are considered as complements or substitutes (Tomek and Kaiser, 2014). If the commodities compete in the same market, they are considered as substitutes (Asche et al., 2003). Degree of substitutability can be measured by the cross-price elasticity. The cross-price elasticity between good i and j can be formulated as:

$$\epsilon_{ij} = \frac{\partial y_i^D}{\partial p_j} \frac{p_j}{y_i^D} \tag{3.1}$$

where y_i^D is the demand function for good *i*. The cross-price elasticity yields the percentage change in demand for good *i* as a result of a 1 % change in the price of good *j*. Based on the sign of the cross-price elasticity, the relationship between good *i* and *j* can have three outcomes:

1. $\epsilon_{ij} > 0 \rightarrow$ goods *i* and *j* are substitutes.
2. $\epsilon_{ij} < 0 \rightarrow$ goods *i* and *j* are complements.
3. $\epsilon_{ij} = 0 \rightarrow$ goods *i* and *j* have unrelated demands.

In addition to cross-price elasticities, the extent to which commodities compete in the same market can be investigated by price changes over time. What constitutes a market has been defined by many economists, e.g. Stigler and Sherwin (1985); Cournot (1971); Marshall (1947). Stigler's definition of the extent of a market is well known:

> "*the area within which the price of a good tends to uniformity, allowance being made for transportation*" (Stigler, 1969)

Basically, Stigler (1969) says that if two goods are in the same market their prices can differ in the short-run, but opportunity of arbitrage will force the prices back to equilibrium; hence, there is a long-term relationship between the commodity prices. An arbitrage opportunity opens with the existence of a price difference within or between markets, such that an agent can buy at low price and sell at high price. The agent will then have a risk-free certain profit opportunity. Historically, prices in the same market have a tendency to co-move. Formally, the *co-movement* is related to the *law of one price* (LOP), which states that there is a long-run relationship between the prices. Stigler's definition is in compliance with Antoine A. Cournot's definition of a market from 1838:

> "*It is evident that an article capable of transportation must flow from the market where its value is less to the market where its value is greater, until the difference in value, from one market to the other, represents no more than the cost of transportation*" (Cournot, 1971)

The many definitions of a market has motivated a waste literature regarding the test of market integration, e.g. Asche et al. (1999); Tveterås (2000); Asche et al. (2003, 2013); Lloyd (2008).

From a theoretical perspective the concept of a market is simple to understand. However, from an empirical stand point it can be difficult to define. The effect of supply and demand on the market price was described by Asche et al. (2003) in a simple manner. We will describe the effect of supply and demand by reproducing the work of Asche et al. (2003): In Figure 3.1 we have two products traded in two different markets, with a normalized price *p*. Initially, assume that the supply curve shifts to the right in *Market 1* from *S1* towards *S1'*, as a result of e.g. lower production costs. The initial price, *p*, will decrease to *p*' and the quantity will increase from *q1* to *q1'*. What impact this supply-shift has on *Market 2* is determined by the cross-price elasticity, given by Equation (3.1) – or, the degree of substitutability between the goods (Asche et al., 2003). First, if the commodities are perfect substitutes a positive shift in supply (from *S1* to *S1*') of commodity 1 will shift the demand of commodity 2 to the left (from *D2* towards *D2*'). The consumer is assumed indifferent, hence he will substitute commodity 1 for commodity 2 as the price of commodity 1 is cheaper. Here, the LOP will apply, and the price in *Market 2* will reach a new equilibrium at the same price as in *Market 1.* The relative price relationship is unchanged (Tveterås, 2000). A price change in either one of the markets will have a response in the other market, which causes the prices to be equal – hence, the two markets are perfectly integrated (Tveterås, 2000).

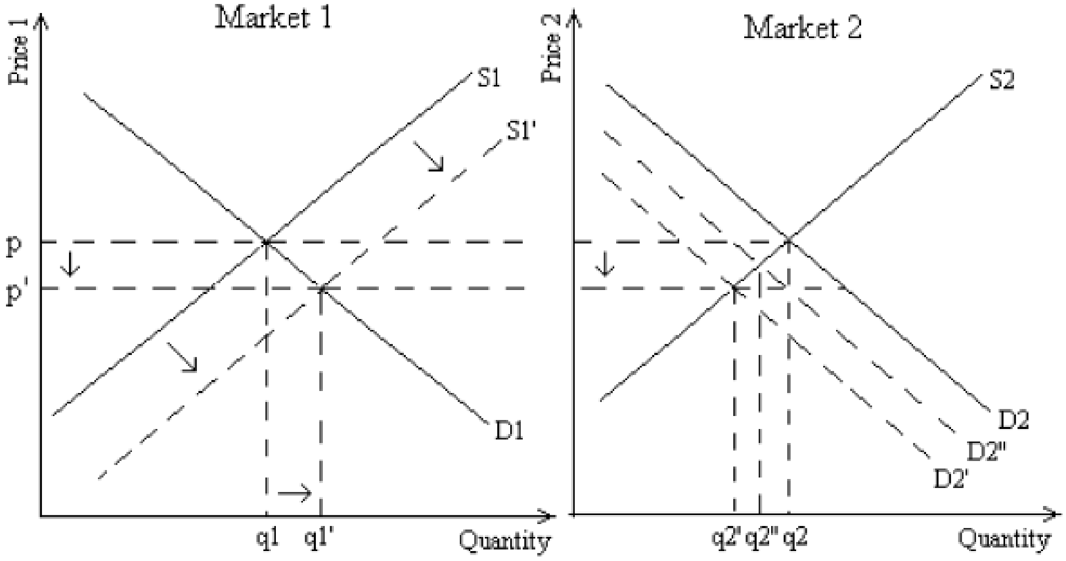


Figure 3.1: Supply and demand curves for two goods competing in two markets (Asche et al., 2003).

Second, if the commodities are imperfect substitutes a positive supply shift in *Market 1* will lower demand in *Market 2*, but not to the same extent as with perfect substitutes. Therefore, the demand schedule in *Market 2* will settle at around *D2”* before the "perfect substitutes demand schedule", *D2’*. The relative price relationship is changed.

Third, if the cross-price elasticity is zero, the price in *Market 2* will be indifferent of a supply shift in *Market 1*. The increased supply in *Market 1* results in reduced price and increased demand. Since the two commodities have a zero cross-price elasticity, there is no possibility of arbitrage. The price of commodity 2 will remain at the initial price *p*.

Finally, if the demand schedule in *Market 2* would shift upwards due to the positive supply shift in *Market 1*, the two products are complements (Asche et al., 2003). Hence, the structure of a market yields information about the relationship between commodities. This is actually the fundamentals behind the hypothesis tested in this thesis.

3.2 Empirical Approach to Market Integration

The previous chapter outlined the fundamentals of a market and market integration. This subsection introduces market integration by an empirical approach. Market integration indicates that the commodity prices are affected by each other. With an econometric approach, a test of market integration is equivalent to testing for existence of relationship among the commodities. Lets say we have two products, and their associated price time series are P_{1t} and P_{2t}. Then, market integration can be tested with the following equation (Tveterås, 2000):

$$P_{1t} = \alpha P_{2t}^{\beta} \tag{3.2}$$

where α is a scaling parameter which can represent the quality difference or transportation costs of the products (Asche et al., 2003). β represents the degree of market integration; a $\beta = 1$ indicates integrated markets, whereas a $\beta = 0$ is a sign of zero integration. If $0 < \beta < 1$, the two commodities are imperfect integrated. Working with daily prices (or returns), the price is often given as the logarithmic value. Hence, we take the natural logarithm of Equation (3.2) and get a linear equation:

$$p_{1t} = \alpha_0 + \beta p_{2t} \tag{3.3}$$

where β has the same characteristics as before; a $\beta = 1$ implies the LOP, and a $\beta = 0$ implies no integration. The intercept, $\alpha_0 = \ln \alpha$, accounts for price differentials among the products (it allows for non-homogeneous products to be integrated (Tveterås, 2000)). Equation (3.3) suffers when price adjustments is not instantaneous. Slade (1986) developed a more dynamic model which included lags of the two prices to incorporate the actual adjustment time.

$$p_{1t} = a + \sum_{j=1}^{m} b_j p_{1t-j} + \sum_{i=0}^{n} c_i p_{2t-i} \tag{3.4}$$

In the latter equation, causality can be tested on the parameter c_i, with a null of:

$$H_0 : c_1 = c_2 = \ldots = c_{n-1} = c_n = 0 \tag{3.5}$$

If the null is rejected, commodity *2* is causing commodity *1*; hence, p_2 will significantly affect p_1. Note that when switching the dependent and independent variable in Equation (3.4), it is possible to test if p_1 causes p_2. In the framework proposed by Slade (1986), the long-run LOP hypothesis is formulated as:

$$\sum_{j=1}^{m} b_j + \sum_{i=0}^{n} c_i = 1 \tag{3.6}$$

Equation (3.3) is just a special case of Equation (3.4) when $c_0 = 1$ and all other coefficients $\left(b_j,\ c_i \text{ for } i, j > 0\right)$ are set to zero.

It must be emphasized that testing for causality is not equivalent with testing for integration (Tveterås, 2000). Granger (1969) states that if a variable Y_t is causing X_t it implies that we better can predict X_t by including all information about Y_t than without Y_t. Hence, causality indicates that one market is affected by another (Tveterås, 2000). Whereas market integration implies that prices of related commodities or markets follow a long-run relationship. On the contrary, causality does not preclude the LOP hypothesis.

At the end of this section we introduce the term *price leader*. The causality between two markets can be either uni- or bidirectional (Asche et al., 2003) (or neither). If there is evidence of exclusively unidirectinal causality, this may indicate that there is one price leader (Lloyd, 2008;

Asche et al., 2003). A price leader dictates prices in other markets.

3.3 Aggregation and the Composite Commodity Theorem

In addition to market integration, *product aggregation* is of great interest when analyzing commodity prices. Hicks (1936) and Leontief (1936) are known as the founders of the *composite commodity theorem*, which states that if a group of individual product prices progress together over time, this group can be represented by a *composite price index*. In other words, the theorem questions if it is possible to to study one commodity independent of the demand in another (Tveterås, 2000). The theorem can be illustrated with the approach of Deaton and Muellbauer (1980) with a simple two-commodities model:

$$P_{1t} = \theta_t P_1^0 \quad \text{and} \quad P_{2t} = \theta_t P_2^0 \tag{3.7}$$

where the superscript "0" indicates a certain base period (which is independent of the time t). θ_t is a time dependent scaling parameter, which causes the ratio between the prices P_1 and P_2 to remain constant. A deviation from this relationship indicates a violation of the composite commodity theorem (Tveterås, 2000). Solving either one of the two equations in Equation (3.7) with respect to θ_t gives one of the prices as a function of the other:

$$P_{1t} = \left(\frac{P_{2t}}{P_2^0}\right) P_1^0 \tag{3.8}$$

The ratio of the base periods is a constant, $a = P_1^0 / P_2^0$, which implies that the latter equation can be rewritten into:

$$P_{1t} = aP_{2t} \tag{3.9}$$

Thus, the two prices are in a relationship. With a logarithmic representation of Equation (3.9), we get a linear equation such as in Equation (3.3):

$$\ln P_{1t} = \ln a + \ln P_{2t} \tag{3.10}$$

Imposing a $\beta = 1$ in Equation (3.3) it is transformed into Equation (3.10). Hence, if the strict LOP holds ($\beta = 1$) the composite commodity theorem is applicable (Asche et al., 1999); there is perfect proportionality between the prices. In theory the theorem is simple, although there is one limitation: the theorem requires an exact proportionality between the prices (Tveterås, 2000). A generalized composite commodity theorem was developed by Lewbel (1996) to permit for small deviations from proportionality (Tveterås, 2000). In summary, there is a strong link between market integration and product aggregation. Asche et al. (1999) proved that if the LOP is true, there is a possibility of product aggregation. Thus, a test of the composite commodity theorem can be executed with a test of the LOP (Asche et al., 1999).

Chapter 4

Time Series Methodology

Econometric analysis of time series data is fundamental in investigation of commodity prices and their co-movement in time. Time series data have special characteristics, which are none existing in the traditional cross-section data. The definition of a stationary time series will be described in Section 4.1, including some techniques to check for stationarity. In Section 4.2 the unit root tests applied in this thesis will be explained. Finally, the concept of cointegration is described in mathematical terms. In this chapter the methods applied in the quantitative analysis will be described. A common assumption of "the future will be like the past" is important in time series regression, this assumption is the base for the term *stationarity*.

4.1 Stationary or Non-Stationary Time Series?

To determine whether a time series is stationary or non-stationary is essential in time series analysis. A *stationary process* is one with a time independent probability distribution (Woolridge, 2009). In other words, the joint probability distribution is not affected by a time shift. Woolridge (2009) provides a definition of a stationary process as the following:

Stationary Stochastic Process. Any stochastic process $\{x_t : t = 1, 2, 3, \ldots\}$ is stationary if $\left(x_{t_1}, x_{t_2}, \ldots, x_{t_m}\right)$ and $\left(x_{t_1+h}, x_{t_2+h}, \ldots, x_{t_m+h}\right)$ for any integer $h \geq 1$ has the same joint probability distribution (Woolridge, 2009).

A time series, x_t, is therefore stationary if the three criteria in Equations (4.1) to (4.3) are satisfied for all values and time periods.

$$E(x_t) = \mu \tag{4.1}$$

$$Var(x_t) = \sigma^2 \tag{4.2}$$

$$Cov(x_t,\ x_{t+h}) = Cov(x_t,\ x_{t-h}) = \gamma_h \tag{4.3}$$

where $E(x_t)$ and $Var(x_t)$ is the mean and variance, respectively, to time series x_t. The variable, *x*, will therefore oscillate about its mean with a constant variation – hence, the variable feature mean reversion (Fabozzi et al., 2014). A stationary time series, as opposed to a non-stationary time series, has a constant mean and variance. $Cov(x_t,\ x_{t+h})$ is the covariance between two values, x_t and x_{t+h}, which is exclusively dependent on the lag length h and not the actual time t. In addition, Equation (4.3) implies that the correlation between x_t and x_{t+h} depends only on h (Woolridge, 2009). A *non-stationary process* is a time series which violate at least one of the criteria in Equations (4.1) to (4.3).

There are two groups of non-stationary variables, *difference-stationary processes* (DSP) and *trend-stationary processes* (TSP) (Tveterås, 2000). The latter is generally formulated as

$$x_t = \mu + \beta t + u_t \tag{4.4}$$

where u_t are independent residuals at time t with *white noise* properties. Pfaff (2008) defines a white noise as a process which have the following properties:

$$E(u_t) = 0 \tag{4.5}$$

$$E\left(u_t^2\right) = \sigma^2 \tag{4.6}$$

$$E(u_t u_\tau) = 0 \quad \text{for} \quad t \neq \tau \tag{4.7}$$

A difference-stationary process is formulated as:

$$x_t = x_{t-1} + u_t \tag{4.8}$$

In this thesis a non-stationary time series is equivalent with a DSP. *Difference* refers to that the time series must be *differenced* to become stationary, since a non-stationary process is integrated of a higher *order* than zero. Expressed the other way around, the *order* refers to the number of times the series must be differenced to become stationary – hence the name *differenced*-stationary process. The first order difference of time series, Δx_t, is the current value, x_t, subtracted with the previous value, x_{t-1}, i.e. $\Delta x = x_t - x_{t-1}$. An integrated series of order one is denoted $I(1)$. The simplest case of a non-stationary time series is given in Equation (4.9), which is known as a random walk which is integrated of order one (Woolridge, 2009).

$$x_t = x_{t-1} + u_t \tag{4.9}$$

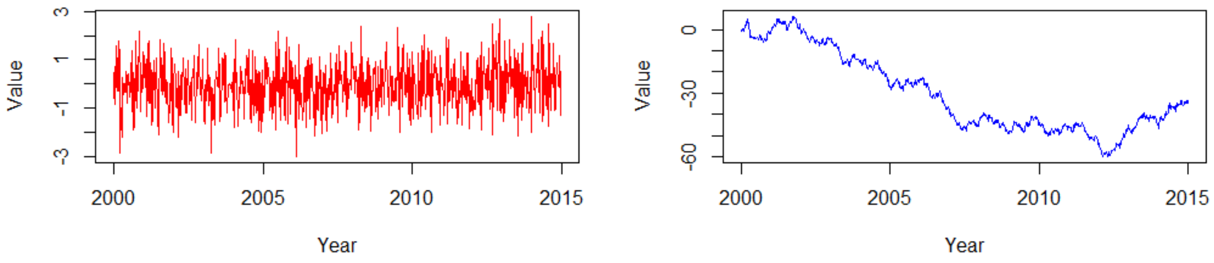
Equation (4.9) is a special case of the more popular process known as the *autoregressive process of order one*, denoted as AR (1), which is formulated as:

$$y_t = \rho y_{t-1} + u_t \tag{4.10}$$

where u_t is a stochastic error term with white noise properties. If ρ equals one in Equation (4.10) it becomes the random walk process given in Equation (4.9). With the stability condition $|\rho| < 1$, the time series y_t is a stationary process - also known as a *stable AR(1) process* (Woolridge, 2009).

Failure to recognize a non-stationary time series can lead to spurious regression results (Woolridge, 2009). Hence, it is important to correctly determine if the time series is stationary or non-stationary prior to any further analysis such as cointegration. The existence of stationarity in a time series can be motivated by visual examination and autocorrelation plots. In Figure 4.1a a stationary time series, with a zero-mean and standard deviation of one, is randomly generated in RStudio. Whereas in Figure 4.1b a non-stationary time series is generated by calculating the inverse function of the first order differences for the time series in Figure 4.1a. In this case, it is easy to distinguish the stationary process from the non-stationary by visual examination; the stationary process is fluctuating around its mean.

An autocorrelation plot, also known as a correlogram, can be of further assistance to determine which "type" of process the underlying data is. Simply stated, the autocorrelation, ρ_h is



(a) A stationary time series.

(b) A non-stationary time series.

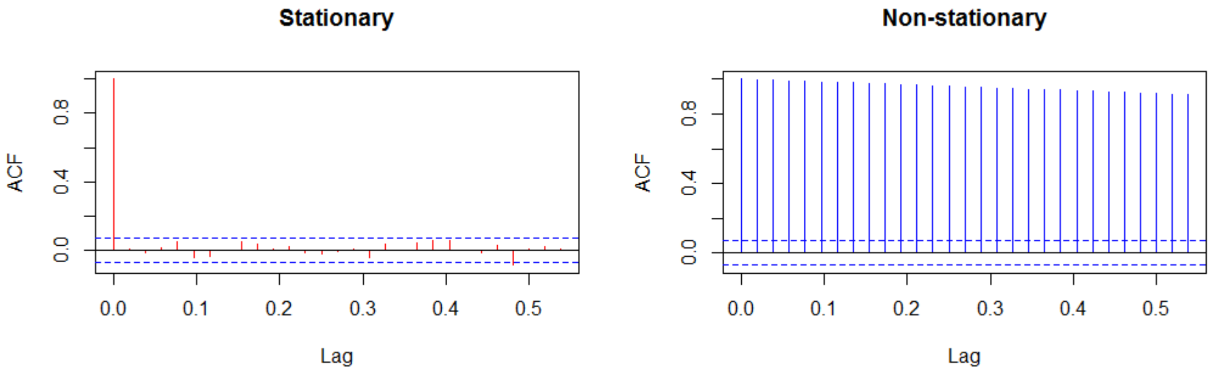
Figure 4.1: A stationary and non-stationary random generated time series with a zero mean and standard deviation of 1.

given by the covariance in Equation (4.3) divided by the variance of the time series, $Var(x_t)$.

$$\rho_h = \frac{Cov(x_t, x_{t+h})}{\sqrt{Var(x_t)\,Var(x_{t+h})}} = \frac{Cov(x_t, x_{t+h})}{Var(x_t)} \tag{4.11}$$

where we apply Equation (4.2) to assume that $Var(x_t) = Var(x_{t+h})$. The autocorrelation describes the relationship between values of a time series at different points in time – the variable is correlated with itself. If the correlation between x_t and x_{t+h} goes to zero quickly as the lag length $h \to \infty$ the process is said to be asymptotically uncorrelated (Woolridge, 2009). Figure 4.2 gives the autocorrelation plots for the two time series in Figure 4.1. The blue horizontal lines in Figure 4.2a and 4.2b give the 95 % confidence intervals. In the stationary process, the autocorrelation function remains within the 95 % confidence lines and the correlation quickly decreases with increasing lags, hence no sign of autocorrelation. Whereas the non-stationary process experiences high degree of autocorrelation, as expected. The two methods described are easy to use, and can be sufficient for a simple time series.

To summarize, we can say that weakly dependent processes are synonymous to stationary processes, which are said to be integrated of order zero (Woolridge, 2009). In other words, there is no need for any transformation of the data before applying regression analysis. However, commodity price time series often behave as they are stationary when integrated by a higher order. Which order of integration a series have is not intuitively. Hence, it is necessary with an empirical method to distinguish between a stationary and non-stationary time series.



(a) A stationary time series.

(b) A non-stationary time series.

Figure 4.2: Autocorrelation plots of the time series in Figure 4.1

4.2 Unit Roots Test

A statistical test that investigates if a time series is stationary or non-stationary is called a *unit root test* (Alexander, 2008). Basically, the test evaluates the order of integration. It exists several unit roots tests: the Phillips-Perron test, the Elliott-Rothenberg-Stock test, the Schmidt-Phillips test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test (Pfaff, 2008). We will apply the most popular unit root test, which is the Dickey-Fuller (DF) test (Dickey and Fuller, 1979). The Dickey-Fuller test is based on a regression of the form:

$$y_t = \rho y_{t-1} + u_t \tag{4.12}$$

with the associated null and alternative hypotheses, respectively:

$$\begin{aligned} H_0 &: \rho = 1 \\ H_1 &: \rho < 1 \end{aligned} \tag{4.13}$$

The test evaluates a null hypothesis of non-stationarity versus the alternative of stationarity. The latter regression equation can be formulated by the first difference of y_t. Hence, equation (4.12) can be formulated as:

$$\begin{aligned}\Delta y_t &= y_t - y_{t-1} \\ &= \left(\rho - 1\right) y_{t-1} + u_t \\ &= \gamma y_{t-1} + u_t\end{aligned} \tag{4.14}$$

where $\gamma = \rho - 1$ with the following null and alternative hypotheses:

$$\begin{aligned}H_0 &: \gamma = 0 \\ H_1 &: \gamma < 0\end{aligned} \tag{4.15}$$

The Dickey-Fuller test is a one-tailed test which applies the t-statistics on $\hat{\rho}$ and $\hat{\gamma}$ (where the hat "^" indicates estimates of ρ and γ, respectively) (Alexander, 2008). Again, the null hypothesis is that the time series is *non-stationary* ($H_0 : \gamma = 0$). In other words, if the null is rejected the time series will be considered as stationary. Equation (4.12) is not necessarily the best description of the true data generating process. Deterministic components such as a constant or a time trend could be included in Equation (4.12) (Woolridge, 2009). In this thesis three different cases of the autoregressive process were evaluated. The first case is a Dickey-Fuller test with no constant nor trend component, which was formulated in Equation (4.12). The second is the Dickey-Fuller test including a constant, α. Finally, both a constant and a trend component, t, are included. A trend can be defined as a continual long-run movement of the variable, and the variable will vibrate around its trend. All three equations are summarized in Equations (4.16) to (4.18).

(i) *The Dickey-Fuller test with no constant and no trend*:

$$\Delta y_t = \gamma y_{t-1} + u_t \tag{4.16}$$

(ii) *The Dickey-Fuller test with a constant, α but no trend*:

$$\Delta y_t = \alpha + \gamma y_{t-1} + u_t \tag{4.17}$$

(iii) *The Dickey-Fuller test with both constant, α, and trend, t*:

$$\Delta y_t = \alpha + \rho t + \gamma y_{t-1} + u_t \tag{4.18}$$

The null and alternative hypothesis for all three cases are given in Equation (4.19); the null is that the time series is non-stationary.

$$H_0 : \gamma = 0 \quad \text{vs.} \quad H_1 : \gamma < 0 \tag{4.19}$$

If the null hypothesis can not be rejected, meaning that y_t is non-stationary, the standard normal distribution of t statistics can not be applied (Woolridge, 2009). Dickey and Fuller solved this problem by developing new critical values with Monte Carlo simulations of a random walk process such as Equation (4.8) (Tveterås, 2000). Another major limitation with the DF-test is presence of autocorrelation. The critical values become biased if there is autocorrelation in the regression residuals (Alexander, 2008). Dickey and Fuller (1981) investigated such higher order autoregressive processes and the associated test statistics. Motivated by the bias test-statistics modified models have been developed, which augmented the regression with the lagged differences, Δy_{t-h}, to correct for autocorrelation – the *Augmented Dickey-Fuller* (ADF) test. The ADF-test can be expressed in three forms, such as the DF-test:

(i) *The Dickey-Fuller test with no constant and no trend*:

$$\Delta y_t = \theta y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + u_t \tag{4.20}$$

$$\tau \rightarrow H_0 : \ \theta = 0 \quad vs. \quad H_1 : \ \theta \neq 0$$

(ii) *The Dickey-Fuller test with a constant, α but no trend*:

$$\Delta y_t = \alpha + \theta y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + u_t \tag{4.21}$$

$$\tau_2 \rightarrow\ H_0:\ \theta = 0 \quad vs. \quad H_1:\ \theta \neq 0$$

$$\phi_1 \rightarrow\ H_0:\ \theta = \alpha = 0 \quad vs. \quad H_1: \theta \neq 0 \text{ and/or } \alpha \neq 0$$

(iii) *The Dickey-Fuller test with both constant, α, and trend, t*:

$$\Delta y_t = \alpha + \rho t + \theta y_{t-1} + \sum_{i=1}^{p-1} \beta_i \Delta y_{t-i} + u_t \tag{4.22}$$

$$\tau_3 \rightarrow\ H_0:\ \theta = 0 \quad vs. \quad H_1:\ \theta \neq 0$$

$$\phi_2 \rightarrow\ H_0:\ \theta = \alpha = 0 \quad vs. \quad H_1: \theta \neq 0 \text{ and/or } \alpha \neq 0$$

$$\phi_3 \rightarrow\ H_0:\ \theta = \rho = 0 \quad vs. \quad H_1: \theta \neq 0 \text{ and/or } \rho \neq 0$$

where β_i are the estimated lag coefficients and p is the order of the autoregressive process.

Phillips and Perron (1988) argues that it is beneficial to start with a richly specified regression, and continue with a stricter model until the null hypothesis is rejected. Thus, we will first run regression on Equation (4.22). Output from the ADF-test of Equation (4.22), yields three test-statistics τ_3, ϕ_2 and ϕ_3; failing to reject ϕ_3, indicates that there is a unit root and no trend. Hence we check ϕ_2 for the presence of drift. Failing to reject this null, implies that there is a unit root and no drift term. Finally, we check the τ_3 hypothesis' test statistics for unit root. The joint hypotheses are tested with an F-ratio test. Applying this *step-by-step*-approach on the output from Equations (4.20) to (4.22), we can better understand which one of the equations is most likely to represent the real case.

The differences, Δy_{t-i}, are included to remove any serial correlation in Δy_t (Woolridge, 2009). The number of lags included must be sufficient to clean up any serial autocorrelation in the data. Hence, it comes at a price. The more lags included in the process, the more initial observations are lost. Consequently, the small sample power will suffer (Woolridge, 2009). If too few lags are including, Blough (1992) states that the probability of rejecting a true null can be very high. Hence, there is a trade-off between power and size when selecting number of lags. Lag selection rules have been developed and studied excessively, such as the Akaike Information Criteria (AIC) and the Schwarz or Bayesian Information Criteria (BIC) (Ng and Perron, 2001). Woolridge (2009) argues that the lag length is a result of the frequency of the data, and recommends that 12 lags are sufficient for monthly data. Hall (1992) came up with two sequential approaches to determine the lag size; the *specific to general* and the *general to specific* rules. The prior rule recommend to start with a small lag k and increase by one until a *non-significant* lag is detected. The latter starts with a large lag k_{max} and successively reduces the lag until a *significant* lag is encountered. The drawback with the general to specific rule is that the null will be falsly rejected 5 % of the times, just by chance. In this thesis the general to specific rule was applied to find the correct number of lags, starting from $k_{max} = 12$, as suggested by Woolridge (2009).

4.2.1 Structural Breaks

It is proved that presence of *structural breaks* will contaminate a time series (Perron, 1989). Structural breaks can be a result of any occurrences which affect the variable of interest, e.g. increased taxes on import of vegetable oils or an El Niño event could both potentially lead to a structural break. Basically, there are two possible breaks: (1) a break can occur at a time τ and last until the end of the period, or (2) it can only occur at a certain time period. Structural breaks can be modeled by including a dummy in the regression. Lets say y_t is a data generating process with a structural break at time τ:

$$y_t = \alpha + \delta D_t + y_{t-1} + u_t \tag{4.23}$$

where we define a dummy variable to encounter the structural break as:

$$D_t = \begin{cases} 1, & \text{if } t = \tau \\ 0, & \text{otherwise} \end{cases} \tag{4.24}$$

Most statistical software have built-in functions to check for structural breaks. It is beyond the scope of this thesis to derive the equations, but it must be emphasized that occurrence of structural breaks can influence the series. Hence, we will apply the test proposed by Zivot and Andrews (1992) to check for structural breaks.

4.3 Cointegration

In Section 1.1, cointegration was described in light of significant contributors and motivation to eliminate spurious regression. Here, the mathematical models will be derived. First, the concept of cointegration is reviewed. Second, the Johansen procedure is derived. Followed by a description of diagnostic tests, applied to control for the regression outcomes in the models.

4.3.1 The Concept of Cointegration

Granger (1981) introduced the idea of *cointegration* into the literature in 1981, which was extended with estimation procedures and tests by Engle and Granger (1987). Cointegration can be explained as a procedure to find a linear combination between two *I(d)* time series that results in a variable with a lower order of integration than *d*. Formally, if two variables x_t and y_t are integrated of order d ($I(d)$), then Engle and Granger (1987) shows that a linear combination:

$$z_t = x_t - a y_t$$

could also be an $I(d)$-process. Furthermore, according to Engle and Granger (1987) it can exist a parameter $b > 0$ such that $z_t \sim I(d-b)$. Lets say that both x_t and y_t are integrated of order $d = 1$, and that $b = 1$, then z_t will be an $I(0)$-process. Engle and Granger (1987) formalized this concept with a definition of cointegration:

"The components of a vector $\mathbf{x}_t$ *are said to be cointegrated of order d, b, denoted* $\mathbf{x}_t \sim CI(d, b)$ *if (a) all components of* $\mathbf{x}_t$ *are* $I(d)$ *and (b) there exists a vector* $\alpha \neq 0$ *so that* $z_t = \alpha' \mathbf{x}_t \sim I(d-b)$, $b >$ *0. Where* α *is the so called cointegration vector" (Engle and Granger, 1987).*

This definition by Engle and Granger (1987) made it possible to study long-run relationships between non-stationary time series. If the variables of interest are cointegrated they have a tendency to not wander too far away from each other (Tveterås, 2000). This is in compliance with the law of one price given in Chapter 3. Market integration tests in econometric are usually executed with either the Engle and Granger bivariate test (Engle and Granger, 1987) or the Johansen procedure (Johansen, 1988; Johansen and Juselius, 1990; Johansen, 1991).

The fundamental difference between the Engle-Granger and Johansen procedure, is that the former is a uni-equation model, whereas the Johansen procedure can handle more than one equation. Arguably the only strength of the Engle-Granger test, which is also its weakness, is its simplicity. The Johansen frame-work has some advantages over the Engle-Granger test. First, it can handle multivariate systems. The Engle-Granger methodology has to be normalized upon one of the variables of interest, without any guidance of which variable should be the dependent variable. Secondly, the Johansen procedure allows hypothesis testing on the cointegration outcomes, which is not valid in the Engle-Granger test. This is important, as we must remember that it is not sufficient to detect a long-run relationship between the variables. Working with market integration, additional restrictions must be imposed on the cointegration results (Tveterås, 2000). Therefore, the Johansen procedure is deemed as more attractive for our multivariate system with ten oils.

4.3.2 The Johansen Procedure

A *vector autoregressive* (VAR) system have n regression equation, in which the independent variables are lagged values of all the n time series. If all the regressors have the same lag order, p, it is called a VAR(p) system. Lets say we have n variables which all have p lags, then the VAR system is formulated as as:

$$\begin{aligned}
y_{1,t} &= \alpha_1 + \beta_{1,1} y_{1,t-1} + \cdots + \beta_{1,p} y_{1,t-p} + \gamma_{1,1} y_{2,t-1} + \cdots + \gamma_{1,p} y_{2,t-p} + \cdots + \omega_{1,1} y_{n,t-1} \\
&\quad + \cdots + \omega_{1,p} y_{n,t-p} + u_{1,t} \\
y_{2,t} &= \alpha_2 + \beta_{2,1} y_{1,t-1} + \cdots + \beta_{2,p} y_{1,t-p} + \gamma_{2,1} y_{2,t-1} + \cdots + \gamma_{2,p} y_{2,t-p} + \cdots + \omega_{2,1} y_{n,t-1} \\
&\quad + \cdots + \omega_{2,p} y_{n,t-p} + u_{2,t} \\
&\ \ \vdots \\
y_{n,t} &= \alpha_n + \beta_{n,1} y_{1,t-1} + \cdots + \beta_{n,p} y_{1,t-p} + \gamma_{n,1} y_{2,t-1} + \cdots + \gamma_{n,p} y_{2,t-p} + \cdots + \omega_{n,1} y_{n,t-1} \\
&\quad + \cdots + \omega_{n,p} y_{n,t-p} + u_{n,t}
\end{aligned} \tag{4.25}$$

where $y_{i,t}$ for $i = 1, \ldots, n$ are the variables at time t. Intuitively, this can be rewritten into a more compact form by matrix notations:

$$\mathbf{Y_t} = \mu + \Phi D_t + \Pi_1 \mathbf{Y}_{t-1} + \Pi_2 \mathbf{Y}_{t-2} + \cdots + \Pi_p \mathbf{Y}_{t-p} + \mathbf{u}_t \tag{4.26}$$

where the $n \times 1$ vector $\mathbf{u}_t$ is the residuals, where each component has white noise properties. $\mathbf{Y}_t$ is an $n \times 1$ vector containing all the endogenous variables. The intercepts are collected in the intercept vector μ, and D_t is a deterministic dummy vector. Π_i, for $i = 1, \ldots, p$, are the $n \times n$ coefficient matrices for the endogenous variables. The latter equation is rewritten by taking the first difference of all the endogenous variables, except the $\mathbf{Y}_{t-p}$:

$$\Delta\mathbf{Y_t} = \mu + \Phi D_t + \Gamma_1 \Delta\mathbf{Y}_{t-1} + \Gamma_2 \Delta\mathbf{Y}_{t-2} + \cdots + \Gamma_{p-1} \Delta\mathbf{Y}_{t-p+1} + \Pi\mathbf{Y}_{t-p} + \mathbf{u}_t \tag{4.27}$$

$$= \mu + \Phi D_t + \Pi\mathbf{Y}_{t-p} + \sum_{i=1}^{k-1} \Gamma_i \Delta\mathbf{Y}_{t-i} + \mathbf{u}_t \tag{4.28}$$

which is often called a *vector error correction model* (VECM). Comparing Equation (4.27) with (4.26), we see that the new parameters Γ_i and Π must be combinations of the "old" parameters Π_i:

$$\Gamma_i = -\mathbf{I} + \Pi_1 + \cdots + \Pi_i, \quad \text{for } i = 1, \ldots, p-1 \tag{4.29}$$

$$\Pi = -\mathbf{I} + \Pi_1 + \ldots + \Pi_p \tag{4.30}$$

where $\mathbf{I}$ is the identity matrix, and Γ_i accumulates the long term impacts; hence Equation (4.27) is known as the *long-run form* (Pfaff, 2008). We assume that each variable is integrated of order one, such that the left hand side in Equation (4.27) is stationary. Hence, the right hand side must be stationary too. Since the lagged differences are the results of previous periods, they are assumed to be stationary. Whats left in the equation for the residuals to behave as white noise is to achieve stationarity in the long-run term $\Pi\mathbf{Y}_{t-p}$. There are three possible cases where this is true:

(i) $rank(\Pi) = N$. It exists n linearly independent combinations, which all are stationary. In other words, all the variables are stationary, and there is no interrelationship among the variables. Obviously, this is of no further interest.

(ii) $rank(\Pi) = 0$. Zero rank indicates no cointegration vectors, as the coefficients are zero, which is an uninteresting trivial solution.

(iii) $0<rank(\Pi) = r < N$. This is the interesting case, where we can decompose the long-run matrix into two $(N \times r)$ matrices, α and β, which causes $\Pi\mathbf{Y}_{t-p} = \alpha\beta'\mathbf{Y}_{t-p}$. α is considered as the *adjustment* matrix and β is the *long-run coefficients* matrix (Pfaff, 2008).

Hence, if $\alpha\beta'\mathbf{Y}_{t-p}$ is stationary, then the $\beta'\mathbf{Y}_{t-p}$ must be stationary. The rank of Π, denoted r, determines the number of linear independent columns of β; hence, the number of cointegration vectors. That is, each column in the long-run coefficient matrix β represents a long-run relationship of the variables in $\mathbf{Y}_t$ (Pfaff, 2008). Johansen (1988, 1991) and Johansen and Juselius (1990) have developed two *Likelihood Ratio* tests to estimate the number of cointegration vectors: the *trace* test and *maximum eigenvalue* the test.

The Trace Test

The trace statistic is formulated as:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{N} \ln(1-\lambda_i) \tag{4.31}$$

where r is increasing from zero to $N-1$, T is the number of observations, and the estimated eigenvalues are ordered successively as $\lambda_1 > \lambda_2 > \ldots > \lambda_N$. The latter equation tests the following hypothesis:

$$H_0:\ rank(\Pi) \leq r \quad \text{vs.} \quad H_1:\ rank(\Pi) > r \tag{4.32}$$

The Maximum Eigenvalue Test

The maximum eigenvalue test is defined as:

$$\lambda_{max}(r,\ r+1) = -T \ln(1-\lambda_{r+1}) \tag{4.33}$$

which test the following hypothesis:

$$H_0: \ rank(\Pi) \leq r \quad \text{vs.} \quad H_1: \ rank(\Pi) = r+1 \tag{4.34}$$

Critical values for both tests have been calculated, and are found in e.g. the Appendix of Johansen and Juselius (1990). The null hypothesis for both tests are equal, that the rank is less than *r*. If the null is rejected, the trace test accepts that the rank is greater than *r*. The maximum-eigenvalue test, when rejection of the null hypothesis, accepts the alternative of the rank being $r+1$. In theory, *r* can take on any value between 0 and N (number of variables), and the tests are evaluated sequentially by starting at r=0 until the null hypothesis can not be rejected, and the number of cointegration vectors (if any) have been disclosed.

4.3.3 The Relationship Between Johansen Cointegration Test and the LOP

Since α and β are undefined, the factorization of $\Pi = \alpha\beta'$ is not unique, it could in theory exist an arbitrary non-singular matrix **A** which causes:

$$\alpha\beta' = \alpha\left(\mathbf{A}\mathbf{A}^{-1}\right)\beta' = \alpha\mathbf{A}\left(\beta\mathbf{A}^{-1}\right)' = \alpha^{*}\beta'^{*} = \Pi$$

Thus, further restrictions on α and β' are necessary to obtain unique values. This possibility is one of the main advantages of the Johansen cointegration test. Lets illustrate this point with a bivariate VAR(1) model (which corresponds well to the bivariate test of cointegration and the LOP for our data). Let $\mathbf{Y}_t = \left(y_{1,t}, \ y_{2,t}\right)$, which can be formulated as the following VECM:

$$\Delta\mathbf{Y}_t = \Pi\mathbf{Y}_{t-1} + \mathbf{u}_t \tag{4.35}$$

If $y_{1,t}$ and $y_{2,t}$ are cointegrated, there exists a vector $\beta = \left(\beta_1, \ \beta_2\right)'$ such that the linear combination $\beta'\mathbf{Y}_t = \beta_1 y_{1,t} + \beta_2 y_{2,t} \ \sim \ I(0)$. Next, a normalization of $\beta_1 = 1$ with $\beta_2 = -\beta$ implies that $\beta'\mathbf{Y}_t = y_{1,t} - \beta y_{2,t}$. The long-run relationship between $y_{1,t}$ and $y_{2,t}$ is then formulated as:

$$y_{1,t} = \beta y_{2,t} + u_t \tag{4.36}$$

where the residuals are an I(0) process, which corrects for deviations from a perfect long-run equilibrium given as $y_{1,t} = \beta y_{2,t}$ – which is the case if the strict LOP holds. Hence, we can impose restrictions on the cointegration vectors of being $\beta = (1, -1)'$, where failing to reject the null implies that the LOP holds for the two prices $y_{1,t}$ and $y_{2,t}$. In a multivariate case, with n variables and $r = (n-1)$ cointegration vectors (which is the case if the LOP holds for the bundle of variables) we can normalize upon the first variable to get a similar expresion as in the bivariate case. Asche et al. (1999) express this cointegration matrix, as cointegration vectors of pairs which sum to zero, hence $\beta_i = (1, -1)'$ for $i = 1, \dots, r$. Thus, the LOP for a multivariate system can be tested with the following restrictions on the $n \times r$ long-run cointegration matrix:

$$\beta = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{bmatrix} \tag{4.37}$$

The α matrix can be defined as the speed of adjustment. Lets elucidate this matrix, by rewriting the bivariate VECM Equation (4.35) into two equations:

$$\Delta y_{1,t} = \alpha_1 \left(y_{1,t-1} - \beta y_{2,t-1}\right) + u_{1,t} \tag{4.38}$$

$$\Delta y_{2,t} = \alpha_2 \left(y_{1,t-1} - \beta y_{2,t-1}\right) + u_{2,t} \tag{4.39}$$

where both $\Delta y_{1,t}$ and $\Delta y_{2,t}$ are represented by $\beta' \mathbf{Y}_{t-1}$, which is known as the *lagged disequilibrium error* (Pfaff, 2008). The rate at which the the first difference of $y_{1,t}$ and $y_{2,t}$ reacts to the lagged disequilibrium error term is controlled by the adjustment coefficients α_1 and α_2, respectively. Another aspect of the adjustment coefficients are observed by pre-multiplication of Equation (4.35) by β':

$$\beta' \Delta \mathbf{Y}_t = \beta' \alpha \beta' \mathbf{Y}_{t-1} + \beta' \mathbf{u}_t$$

Which can be rewritten into:

$$\beta'\mathbf{Y}_t = \beta'\left(1+\alpha\beta'\right)\mathbf{Y}_{t-1} + \beta'\mathbf{u}_t$$

Lets say that $y_t = \beta'\mathbf{Y}_t$, $\rho = \left(1+\alpha\beta'\right) = 1 + \left(\alpha_1 - \beta\alpha_2\right)$ and $v_t = \beta'\mathbf{u}_t = u_{1,t} - \beta u_{2,t}$. Hence, the latter equation has the same form as the AR(1) model in Equation (4.10):

$$y_t = \rho y_{t-1} + v_t$$

Thus, the stability conditions is given as $|\rho| = |1 + \left(\alpha_1 - \beta\alpha_2\right)| < 1$. In the case of the LOP, β was said to be 1 (or rather, -1, but the negative sign is already incorporated in Equations (4.38) and (4.39)), which gives the following stability condition:

$$|\rho| = |1 + (\alpha_1 - \alpha_2)| < 1 \tag{4.40}$$

Hence, the stability condition is satisfied for $\alpha_1 - \alpha_2 < 0$ and $\alpha_1 - \alpha_2 > -2$. If $\alpha_2 = 0$, then $-2 < \alpha_1 < 0$, which can be considered as if variable $y_{1,t}$ in Equation (4.38) shows evidence of *mean reversion*.

4.3.4 Weak Exogeneity and Causality

Until now all variables in $\mathbf{Y}_t$ have been treated as endogenous variables. A *weakly exogenous* variable, $\Delta y_{i,t}$, is defined as a variable where its determinants are left out of the regression will cause no loss of information (Woolridge, 2009). One approach to test for weak exogeneity of variable $\Delta y_{i,t}$ is to test whether we can reject the hypothesis of the associated row-elements in the adjustment coefficient, α, being different from zero or not (Liddle, 2009; Tveterås, 2000). Hence, if the row contains only zero elements, the other endogenous variables will have no influence on the variable $\Delta y_{i,t}$.

Causality is the link between two variables, or time series, and causality tests can be utilized to determine if one time series is useful to predict another. The most common method is the Granger causality test (Granger, 1969). A variable y_1 is said to *Granger-cause* variable y_2 if we better can predict variable y_2 by its own lagged value *and* the lagged values of y_1, than by only past values of y_2. In this thesis the Toda and Yamamoto (TY) causality test is implemented (Toda and Yamamoto, 1995). The reason for checking causality by this approach over the popular Granger-causality test is the fear of having two variables of different order (e.g. I(1) and I(0)). Hence, cointegration results are not valid. The process relies on a VAR(p), which has white noise residuals. The null hypothesis is that variable y_1 does not Granger-cause y_2.

Lets say we have a VAR(2) for a bivariate system where it is believed that the previous two time periods or less may influence the variable today. This example may be formulated as:

$$y_{1,t} = c_1 + \alpha_{1,1} y_{1,t-1} + \alpha_{1,2} y_{1,t-1} + \beta_{1,1} y_{2,t-1} + \beta_{1,2} y_{2,t-2} + \epsilon_{1,t}$$
$$y_{2,t} = c_2 + \alpha_{2,1} y_{2,t-1} + \alpha_{2,2} y_{2,t-1} + \beta_{2,1} y_{1,t-1} + \beta_{2,2} y_{1,t-2} + \epsilon_{2,t}$$

where the $\alpha_{i,j}$ and $\beta_{i,j}$ are the coefficients for the lagged values of the dependent and independent variables, respectively. Whether y_2 does Granger-cause y_1 or not, is tested by imposing the following restrictions, $\beta_{1,1} = \beta_{1,2} = 0$. Similarly we can test if y_1 does Granger-cause y_2 by testing if $\beta_{2,1} = \beta_{2,2} = 0$. Toda and Yamamoto (1995) propose that if the variables are cointegrated, there must be Granger causality in at least one direction.

4.4 Diagnostic Tests

Diagnostic tests implies to validate the assumptions made for the VAR system by inspecting the residuals. In this section we will give a brief overview of the different methods applied to our data.

The existence of serial correlation (autcorrelation) in the residuals can be checked with the *Breusch-Godfrey test* (Breusch, 1978; Godfrey, 1978). Where the null states that there is no serial correlation among the residuals. Presence of serial correlation can be removed by increasing the lags. The Breusch-Godfrey test is a *Lagrange Multiplier* (LM) test.

Heteroskedasticity is checked with an *autoregressive conditional heteroskedasticity* (ARCH) test (Engle, 1982). This test evaluates if the observations have different variances. A rejection of the null indicates that ARCH-effects are present in the sample.

The Johansen cointegration assumes normality, hence the importance of checking whether the series are normally distributed. The Jarque and Bera (1980) test checks for kurtosis and skewness in the observations. The null is that the data are normal distributed.

In addition, the residuals from the ADF-tests have been tested for normality and independence by the Shapiro-Wilks test and Ljung-Box test, respectively. These results are unreported, but important for the strenght of the upcoming analysis and validation of the final results.

Chapter 5

Empirical Results

In this chapter all test results performed to elucidate integration in the global vegetable oil market are presented. Section 5.1 builds on the prelimenary analysis of the oil prices in Chapter 2. The data will be further described in terms of descriptive statistics to better understand the underlying data. In addition, the sub-samples of interest will be presented. In the second section, the results from the ADF-tests are presented. In the third part of this chapter, Johansen cointegration results are presented. We apply cointegration on pairs and full systems. At the end of this chapter, we try to detect any price leader(s) by the means of weak exogeneity test, and further investigate the causality between the oils by the Toda and Yamamoto causality test. Finally, we question if there is cointegration among the different groupings of oils, as mentioned in section 2.3.

All econometric models have been executed in `RStudio` version `3.2.5` developed by the RStudio Team (2016).

5.1 The Data and Descriptive Statistics

In Chapter 2 the characteristics of each oil were presented. The oil prices are given as average monthly prices from January 2000 upto December 2015, hence we have 192 observations in the full sample. All prices are provided by ISTA Mielke GmbH (2016), better known as the *Oil World*. Some of the observations had a zero price, probably due to lack of trade in that month. An arithmetic average of the adjacent prices was applied to fill the gaps. Besides from the descriptive statistics presented in this section, the econometric modeling where performed with the natural logarithm of the price series.

Firstly, a discussion of price development in the vegetable oil market can shed further light on the market when splitting the full sample into shorter sub-samples. In Table 5.1 and Figure 5.1 the three sub-samples investigated in this thesis are presented. The *Full sample* includes all observations, and is motivated by a potential long-run relationship between the oils. The *Sub-sample 0*, or the pre-commodity boom period, is just a check for consistency in the data. Oil prices have been traded at stable ratios upto the commodity boom, and this period will not be given much consideration in this study. *Sub-sample 1*, or the commodity boom period, focus on the price development in a period of continual increase in the vegetable oil prices. The *Sub-sample 2* questions the price development in the aftermath of the financial crisis until the end of our sample (December 2015). By visual examination of the price developments in Figure 5.1, fish oil seems to be in premium compared to the others from June 2011 until the end of the sample. Hence, we study the relationship between the oils after they started to diverge with *Sub-sample 3*.

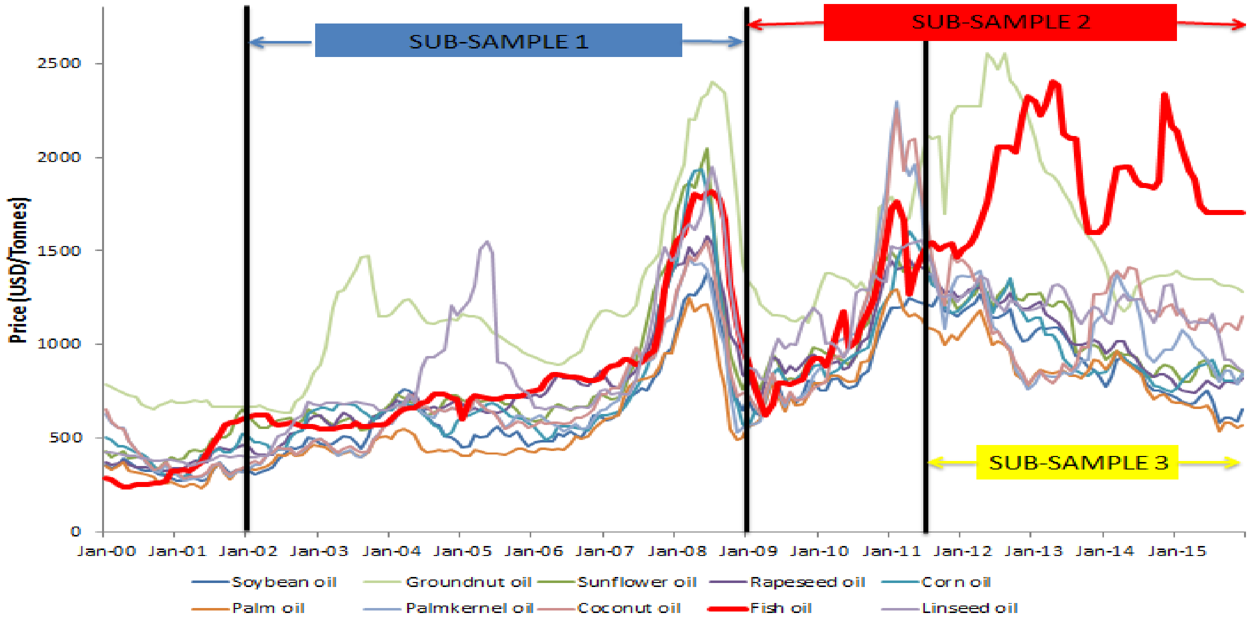


Figure 5.1: Historic development of all oils, and the sub-sample periods (ISTA Mielke GmbH, 2016).

Secondly, when working with large data some explanatory analysis is advised to get a better

Table 5.1: The full sample and three sub-samples are investigated with associated curiosity.

Name	**Synonym**	**Time Period**	**Hypothesis**
Full sample	-	Jan. 2000 - Dec. 2015	One global market?
Sub-sample 0	Pre-commodity boom period	Jan. 2000 to Dec. 2002	Consistency check
Sub-sample 1	Commodity boom period	Jan. 2003 - Dec. 2008	Did the oils rise together?
Sub-sample 2	Post financial crisis	Jan. 2009 - Dec. 2015	What happened after the boom?
Sub-sample 3	The fish oil era	Jun. 2011 - Dec. 2015	Are the oils integrated with a shift?

understanding of your data. Figure 5.2 plots the mean price for each oil for the associated sample periods (given in Table 5.1). Exact values and the associated standard deviations are given in in Tables B.1 to B.3. This gives a good impression of the price development since January 2000. The pre-commodity boom period clearly has the lowest mean value for all the oils, followed by the commodity boom era. Sub-sample 2 includes the increase in price in the aftermath of the financial crises, hence its lower values for groundnut and fish oil than the Sub-sample 3 period. The peaks in fish oil price in Sub-sample 2 and 3, is higher than all the other oil (except for the groundnut oil price from January 2009 until December 2015). Whereas in the pre- and commodity boom periods, several oils showed approximately the same mean prices as fish oil. Hence, the prices are deemed as similar in the pre-commodity boom and commodity boom periods.

In addition, the standard deviations, given in Tables B.1 to B.3, indicates that the fish and groundnut oil prices are deemed as more volatile than the other oils. This indicates a less stable price development than for the other oils; greater variation in the price through a given time period. The variation is mainly driven by increased prices, imposed by e.g. higher demand. Especially is the standard deviation of fish oil price in the post-financial crisis period significantly larger than the others in this period. However, in the period from middle of 2011 to the end of the sample, the price variation among the oils seem to be fairly stable. For now, we can state that fish oil was traded approximately at par with the other oils until the aftermath of the financial crisis.

Another approach to compare the prices is to calculate the ratio between the prices to get the relative price between the oils. Here, relative price is defined as the price ratio of one oil

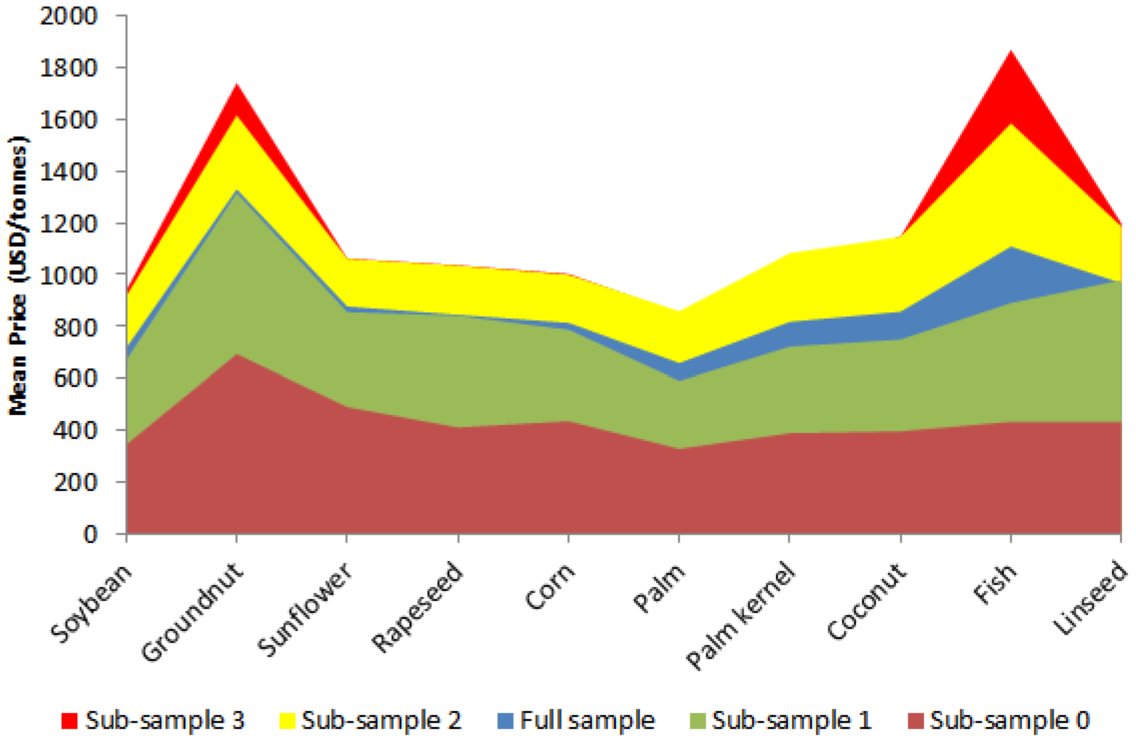


Figure 5.2: Mean prices for the ten oils at respective sample periods given in Table 5.1.

divided by another, e.g. the relative price of fish oil to soybean oil is given as:

$$Fish\ oil\ to\ soybean\ oil\ ratio = \frac{Fish\ oil\ price}{Soybean\ oil\ price}$$

In Figure 5.3 the relative prices of fish oil to rapeseed oil, soybean oil and palm oil are plotted (dashed lines). The relative prices are volatile, and seem to fluctuate above 1.0 until 2011-12 where it progress upwards. In the commodity-boom period, there is a weak tendency of mean-reversion of the fish oil-ratios. In the fifth column of Tables B.5 and B.6, fish oil relative ratios are given for Sub-sample 1 and Sum-sample 2, respectively. Clearly, fish oil has become premium to the other oils, with an average relative price of 1.5 or higher to the majority of the other oils. By considering the period from 2011 until the end of our sample, this ratio is substantially higher as observed in Figure 5.3.

Contrary to the relative prices of fish oil, the ratios of one vegetable oil divided by another, which are plotted as solid lines in Figure 5.3, are more stable around just below 1.0. For now, stable relative prices, such as soybean to rapeseed oil (red line), are interpreted as potential constant long-run relationship. Tables B.4 to B.6 report the average relative prices for soybean oil, rapeseed oil, palm oil and fish oil to the other oils, over the Full sample, Sub-sample 1 and 2, respectively. Unreported calculations of the other relative prices shows similar results as the reported. In the commodity boom period soybean oil is traded at a cheaper price than all oils, except for palm oil, with relative prices below 1.0. Soybean oil seemed to lose its relationship with the others as fish oil price increased around 2002/03. The soybean oil price recovered around the financial crisis, and increased its relative price ratios by approximately 10 %, but still a ratio below 1.0.

Rapeseed oil was traded approximately at par with fish oil between 2003 until 2009 with an average ratio of 1.04 (fish oil/rapeseed oil), except for a small deviation around 2008-09. In the same period, the rapeseed to soybean oil prices were traded at a ratio of 1.28; the rapeseed oil was on average 28 % more costly than soybean oil. However, as fish oil and rapeseed oil diverged in June 2011, the relationship between rapeseed oil and soybean oil reduced to a ratio of 1.14 (on average). The relative price of fish oil to rapeseed oil in the period 2009 to 2015 was at 1.57. Thus, the relative stable relationship of rapeseed oil and fish oil in the commodity boom period had disappeared. Fish oil had become a premium oil product.

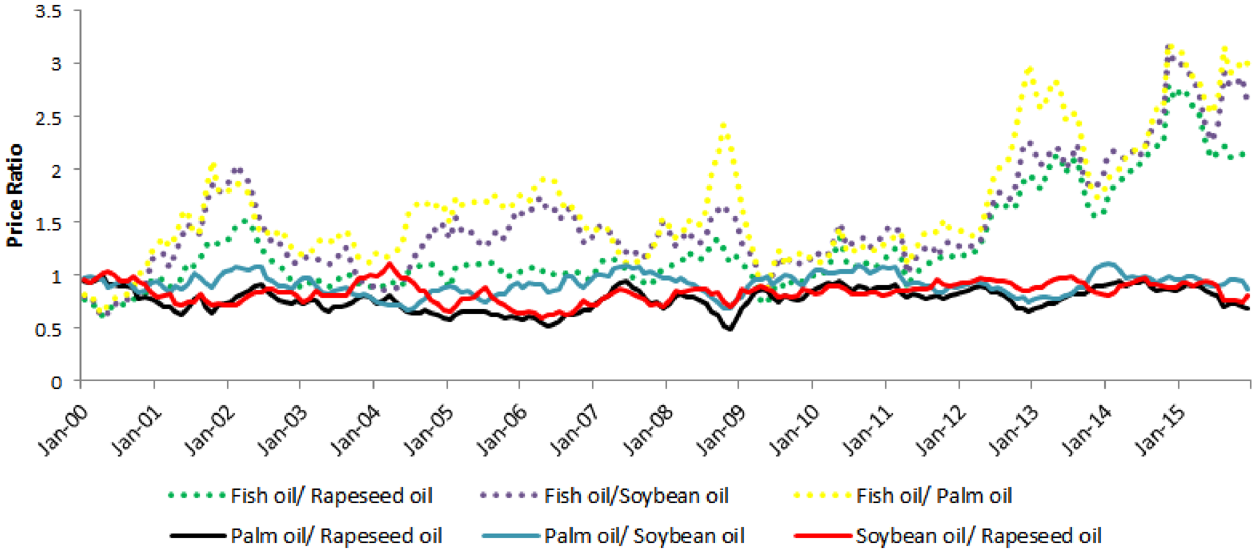
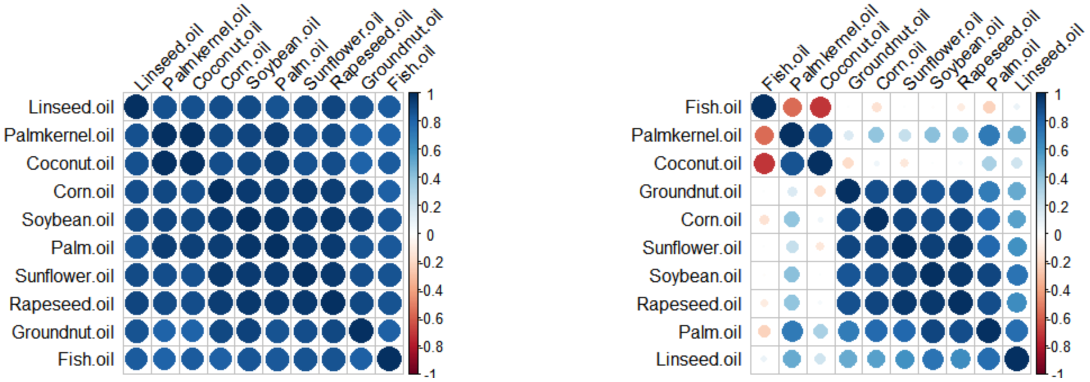


Figure 5.3: The relative price for fish oil to any vegetable oils (dashed lines) is significantly different from one. Whereas the relative prices for the vegetable oils (solid lines) are close to one and constant.

The correlation between the prices have been calculated and are presented visually in Figure 5.4. Interestingly, the correlation is very positive for all oils over the full period. This positive correlation must be considered as an indication of a relationship between the oils. However, correlation of the prices over Sub-sample 3 shows that the fish oil has become negative correlated with some oils, and is no longer strictly positive correlated with all oils.



(a) Correlation full sample.

(b) Correlation Sub-sample 3

Figure 5.4: The correlation is strong positive for the full sample, whereas it becomes negative for some oils in Sub-Sample 3 due to price divergence. Blue color is positive correlation, and red is negative correlation.

Although, the vegetable oils seem to move together over time we must confirm if the time series are stationary or not. Preliminary information of the data generating processes can be withdrawn from correlograms, as mentioned in the previous chapter. The idea is that non-stationary processes have long memory. The correlation of one period to the value of a previous period yields an impression of the "length" of this memory. Figure A.1 shows that all oils have slow decline in the autocorrelation, hence it indicates long memory and non-stationary processes. However, the time series must be formally tested with the unit roots test.

At the end of this section, we have investigated the distributions of each price series by histograms, presented in Figures A.2. Fish oil, soybean oil, rapeseed oil and coconut oil are skewed to the left. A left (negative) skewed plot indicates that a majority of the prices are above its average price for the full period. Therefore, these plots illustrates that among the mentioned oils, it is more likely to find high prices rather than low, compared to its average price, in the period 2000 to 2015. The other six oils, show weak tendency to behave as normal. Summarizing, there is a question of the underlying distributions can be considered as normal.

5.2 Dickey-Fuller Tests for Unit Roots

Motivated by the correlograms in previous chapter, we will apply the ADF-test to check for presence of unit roots. The procedure was described in Section 4.2. The number of lags in the model were determined with the general to specific rule recommended by Hall (1992), with a $k_{max} = 12$. Some oils showed discouraging results where no significant lag was detected, then the Akaike selection criteria (AIK) was applied. Initially the process was richly specified to include a constant, trend and eleven seasonal dummies, proceeding with more restrictive processes. First step is to run the ADF-test on the variables in *levels*. If the null of non-stationarity can not be rejected, we proceed with the ADF-test on the first order differences of the variables. Then, a rejection of the null hypothesis indicates that the variables are stationary and integrated of order one $\sim I(1)$. All prices and the associated first differences can be seen in Figures A.3 and A.4. The prices are not fluctuating around zero, hence there is evidence of a constant term in the process. All variables seem to be increasing with time, hence there could exist a weak time trend in all of the oils. The ADF-tests have the purpose to better understand how the underlying variable is behaving. Based on visual interpretation of the price series, the ADF-tests are executed with *no constant*, *a constant* and *constant & time trend* by Equations (4.20) to (4.22), respectively. The results from the ADF-tests on the full period, Sub-sample 1, Sub-sample 2 and Sub-sample 3 are reported in Tables 5.2, 5.3, 5.4 and 5.5, respectively. Where the first column gives the variable of interest (both in levels and first order differences, Δ), the second, third and fourth shows the ADF-test results for the respective test given in the header of the column.

For the full period, all oils seem to be integrated of order one based on the results presented in Table 5.2. The null is not rejected in any variables in level, except for corn oil and coconut oil which showed ambiguous results when applying both a trend and a constant at the 5 % significance level. The ADF-test is sensitive to the number of lags, which could have caused the rejection of the null at 5 % significance level for the two oils. The first differences of all oils are rejected at 1 % significance level, hence over the full period all variables are considered as integrated of order one.

The sub-samples can be inspected visually from Figure A.5. Neither sub-samples have a clear trend, or a perfectly constant mean through the period. Thus, its difficult to characterize

Table 5.2: ADF-test results for the full sample period, January 2000 to December 2015. The chosen lags are given in the parentheses. The (1%, 5%, 10%)-critical values are (-3.99, -3.43, -3.13), (-3.46, -2.88, -2.57) and (-2.58, -1.95, -1.62) for “Constant & Trend”, “Constant” and “No Constant”, respectively.

Variable	**Constant & Trend**	**Constant**	**No Constant**
Soybean oil	-1.8086 (10)	-1.9273 (1)	0.6762 (8)
Δ Soybean oil	-8.8710** (1)	-5.3370** (8)	-5.7160** (7)
Groundnut oil	-2.6926 (8)	-2.7395 (6)	0.3451 (7)
Δ Groundnut oil	-6.7126** (1)	-4.2122** (7)	-4.1640** (6)
Sunflower oil	-2.7996 (8)	-2.5902 (7)	0.3544 (3)
Δ Sunflower oil	-8.4828** (1)	-5.1459** (3)	-6.0089** (2)
Rapeseed oil	-1.9123 (3)	-2.0056 (3)	0.5445 (3)
Δ Rapeseed oil	-8.0699** (1)	-5.0890** (3)	-5.8114** (2)
Corn oil	-3.5788* (3)	-2.3207 (1)	0.2061 (1)
Δ Corn oil	-5.7493** (7)	-5.8861** (5)	-7.3019** (0)
Palm oil	-1.7386 (6)	-1.8063 (5)	0.3691 (5)
Δ Palm oil	-9.3729** (1)	-5.7133** (5)	-5.8411** (4)
Palm kernel oil	-2.9070 (5)	-1.9328 (5)	0.3225 (5)
Δ Palm kernel oil	-7.8946** (1)	-5.2978** (3)	-5.8616** (4)
Coconut oil	-3.5046* (5)	-1.7964 (3)	0.3422 (3)
Δ Coconut oil	-4.9202** (3)	-4.9372** (3)	-5.6297** (2)
Fish oil	-3.2034 (2)	-1.7190 (11)	1.4721 (12)
Δ Fish oil	-7.7558** (1)	-7.6811** (1)	-8.9144** (0)
Linseed oil	-2.4564 (3)	-2.0527 (1)	0.4685 (8)
Δ Linseed oil	-7.9764** (1)	-7.9124** (1)	-10.0114** (0)

* indicates test statistics at 5 % significance level. ** indicates test statistics at 1 % significance level.

the sample period as purely a trend process, or constant process. From the ADF-test results for Sub-sample 1, presented in Table 5.3, only corn and groundnut oil can be considered as an I(1) variable at 10 % significance level, when we apply both a trend and constant in the test. Possible explanation is the lack of trend in the given period, hence the given results. Another explanations could be a presence of parameter instability, which is less likely than the first explanation for this time period. The results may indicate that the oils are trend-stationary, but we progress with the assumption of the variables having a stochastic trend for this period (until otherwise is proven). With only an intercept in the ADF-equation for Sub-sample 1 all oils except groundnut oil are integrated of order one in this period. For the pure random walk processes, all variables are integrated of order one. Thus, we can apply cointegration tests for this sub period by including only a constant in the process.

Table 5.3: ADF-test results for Sub-sample 1, January 2003 to December 2008. The chosen lags are given in the parentheses. The (1%, 5%, 10%)-critical values are (-4.04, -3.45, -3.15), (-3.51, -2.89, -2.58) and (-2.58, -1.95, -1.62) for "Constant & Trend", "Constant" and "No Constant", respectively.

Variable	**Constant & Trend**	**Constant**	**No Constant**
Soybean oil	-1.8855 (1)	-1.8683 (1)	-0.5351 (8)
Δ Soybean oil	-3.0438 (7)	-4.8021** (0)	-3.2073** (7)
Groundnut oil	-2.3306 (1)	-2.3024 (1)	-0.1371 (7)
Δ Groundnut oil	-2.0164 (6)	-2.6989 (0)	-2.4563* (6)
Sunflower oil	-2.7281 (3)	-2.3713 (3)	0.5761 (11)
Δ Sunflower oil	-2.5102 (10)	-3.6045** (0)	-3.6329** (0)
Rapeseed oil	-2.5149 (3)	-1.7297 (1)	0.2631 (9)
Δ Rapeseed oil	-1.8351 (8)	-4.3201** (0)	-2.3224* (8)
Corn oil	-2.1526 (1)	-1.9719 (9)	0.6967 (8)
Δ Corn oil	-3.2150 (7)	-3.4452* (7)	-3.1604** (0)
Palm oil	-2.1078 (1)	-1.6923 (1)	0.7438 (7)
Δ Palm oil	-2.6330 (6)	-4.6467** (0)	-2.8230** (6)
Palm kernel oil	-2.4861 (3)	-1.6195 (8)	0.7256 (8)
Δ Palm kernel oil	-2.9922 (7)	-4.9648** (0)	-3.0880** (7)
Coconut oil	-2.0418 (1)	-1.6964 (1)	1.2535 (8)
Δ Coconut oil	-3.2060 (7)	-5.1568** (0)	-3.0952** (7)
Fish oil	2.7706 (2)	-1.6493 (1)	1.5831 (11)
Δ Fish oil	-2.8001 (10)	-4.3107** (0)	-4.3163** (0)
Linseed oil	-1.8627 (1)	-1.7685 (1)	0.1841 (1)
Δ Linseed oil	-2.50435 (12)	-5.8580** (0)	-5.8928** (0)

* indicates test statistics at 5 % significance level. ** indicates test statistics at 1 % significance level.

In Table 5.4 the ADF-test results are presented for Sub-sample 2. The test fails to reject the null of unit root of the variables in level for any of the tests, while the null is rejected for all first differences of the variables. Therefore, all variables will be treated as I(1) in this sub-sample. The same conclusion is taken with respect to Sub-sample 3 (see Table 5.5). By the means of ADF-tests, we have confirmed that all variables are integrated of order one with an unrestricted constant. Hence, all oils can be examined further by the Johansen cointegration approach.

5.3 The Johansen Cointegration Test

The bivariate Johansen test is first presented, followed by the multivariate test of the system. For the bivariate and full system, tests have been performed on the full sample period, Sub-sample 1, Sub-sample 2 and Sub-sample 3. Both the bivariate and multivariate test are conducted with a

Table 5.4: ADF-test results for Sub-sample 2, January 2009 to December 2015. The chosen lags are given in the parentheses. The (1%, 5%, 10%)-critical values are (-4.04, -3.45, -3.15), (-3.51, -2.89, -2.58) and (-2.60, -1.95, -1.62) for “Constant & Trend”, “Constant” and “No Constant”, respectively.

Variable	**Constant & Trend**	**Constant**	**No Constant**
Soybean oil	-1.7416 (2)	-1.6773 (1)	0.0436 (2)
Δ Soybean oil	-7.5335** (1)	-6.9342** (1)	-6.9840** (1)
Groundnut oil	-1.7969 (6)	-1.8402 (6)	0.1372 (6)
Δ Groundnut oil	-6.0008** (1)	-8.6187** (0)	-2.5956** (5)
Sunflower oil	-2.1284 (2)	-1.6496 (10)	0.1966 (2)
Δ Sunflower oil	-7.2351** (1)	-6.8926** (1)	-6.9305** (1)
Rapeseed oil	-2.4281 (2)	-1.6800 (1)	0.2914 (2)
Δ Rapeseed oil	-7.4133** (1)	-6.7696** (1)	-6.8068** (1)
Corn oil	-2.3961 (8)	-1.9112 (1)	0.2075 (8)
Δ Corn oil	-3.7510* (7)	-5.1446** (1)	-3.3983** (7)
Palm oil	-2.0113 (0)	-1.5845 (1)	-0.1141 (2)
Δ Palm oil	-7.8615** (0)	-6.4503** (1)	-7.4420** (0)
Palm kernel oil	-2.4910 (3)	-2.4763 (3)	-0.0927 (3)
Δ Palm kernel oil	-4.4687** (5)	-3.4584** (2)	-3.4912** (2)
Coconut oil	-2.2856 (3)	-2.3436 (3)	0.2367 (3)
Δ Coconut oil	-3.5123* (2)	-3.5173** (2)	-3.5232** (2)
Fish oil	-1.8734 (1)	-2.1288 (1)	0.8627 (1)
Δ Fish oil	-7.0338** (1)	-7.1305** (0)	-6.6383** (1)
Linseed oil	-2.2897 (1)	-2.4171 (1)	0.0586 (2)
Δ Linseed oil	-7.54305** (1)	-7.0208** (1)	-7.0647** (1)

* indicates test statistics at 5 % significance level. ** indicates test statistics at 1 % significance level.

restricted intercept and 11 monthly, centered dummies with Equation (4.27). The procedure follows the theory outlined in section 4.3.2. The intercept can be interpreted as the price difference between two oils in the long-run, and will (hopefully) correspond to the difference in price type and transportation costs (FOB vs. CIF). Misspecification tests have showed problems with the underlying distribution of the variables. Different lags were applied to encounter for potential autocorrelation. For most oils 2 or 3 lags were sufficient, except for groundnut oil, palm kernel oil and coconut oil which required more lags to remove potential serial correlation. The diagnostic tests are summarized in Table B.14, where there is clearly a problem with the normality assumption. When too many lags are included, the test results can be weak. Test results indicate that this may be the case. Hence, the bivariate cointegration results when these three oils are included must not be evaluated without any other information (see section 5.4). The residuals

Table 5.5: ADF-test results for Sub-sample 3, June 2011 to December 2015. The chosen lags are given in the parentheses. The (1%, 5%, 10%)-critical values are (-4.04, -3.45, -3.15), (-3.51, -2.89, -2.58) and (-2.60, -1.95, -1.62) for "Constant & Trend", "Constant" and "No Constant", respectively.

Variable	**Constant & Trend**	**Constant**	**No Constant**
Soybean oil	-3.1375 (5)	0.08444 (5)	-1.7597 (2)
Δ Soybean oil	-4.4862** (9)	-4.7765** (3)	-5.6161** (0)
Groundnut oil	-2.3169 (9)	-0.9997 (7)	-1.2053 (7)
Δ Groundnut oil	-5.0311** (1)	-2.3822 (6)	-2.0763* (6)
Sunflower oil	-1.4366 (3)	-1.0482 (2)	-1.6037 (2)
Δ Sunflower oil	-6.7480** (1)	-6.8065** (1)	-6.5284** (1)
Rapeseed oil	-2.4541 (1)	-0.1848 (11)	-1.9912 (3)
Δ Rapeseed oil	-5.8741** (1)	-5.8566** (1)	-3.7859** (2)
Corn oil	-2.2599 (1)	-1.6242 (1)	-1.0189 (1)
Δ Corn oil	-4.6977** (1)	-5.0338** (0)	-4.9559** (0)
Palm oil	-2.3292 (1)	-0.4671 (1)	-1.6087 (1)
Δ Palm oil	-5.7451** (1)	-6.7759** (0)	-6.4916** (0)
Palm kernel oil	-2.5236 (3)	-2.5902 (6)	-1.0872 (6)
Δ Palm kernel oil	-3.7589* (5)	-3.7000** (5)	-3.5516** (5)
Coconut oil	-2.1503 (1)	-2.4824 (0)	-0.6362 (1)
Δ Coconut oil	-3.4652* (3)	-6.3098** (0)	-6.3303** (0)
Fish oil	-1.9839 (1)	-2.1546 (1)	0.1389 (2)
Δ Fish oil	-4.4830** (1)	-5.5373** (0)	-4.4357** (1)
Linseed oil	-3.1677 (1)	-0.8412 (2)	-1.4763 (2)
Δ Linseed oil	-6.1324** (1)	-6.1475** (1)	-5.9031** (1)

* indicates test statistics at 5 % significance level. ** indicates test statistics at 1 % significance level.

from the tests showed deviation from the assumption of normality, which weakens the Johansen cointegration test results where normality is assumed. Hence, for the bivariate tests where the test result was near the critical value of 5 %, we proceeded as the oils were cointegrated.

From January 2000 until December 2015 there have been some periods of abnormal prices, e.g. the months surrounding the financial crisis. Instead of including dummy variables to encounter these periods, we assume that by splitting the period up in three sub-periods will give us enough and correct information to come up with reasonable conclusions. In the following a rejection of the null of no cointegration refers to the 5 % significance level in both max- and trace-test, if no further comments are added.

5.3.1 Pairwise Test of Market Integration

From the ADF-test results, we concluded that all variables may be applicable in cointegration tests. Ten time series with 192 observations each is a demanding data set when working with multivariate cointegration. Most likely we will not find *n-1* cointegration vectors, and the multivariate cointegration test results can be difficult to interpret. Therefore, we initially performed bivariate cointegration tests with the objective to determine specific relationship between oil-pairs. Hence, the test evaluates if the rank of Π equals zero or one, which suggest that there is no cointegration vector or one cointegration vector, respectively, for the two oils of interest. Since the variables are integrated of order one, a rank equals to two is not expected. Both the trace- and maximum eigenvalue test is evaluated by Equations (4.31) and (4.33), respectively, with the associated hypothesis from Equations (4.32) and (4.34).

In theory, a system of *n* variables can result in *n-1* cointegration vectors. If this occurs, there must be $(n^2 - n)/2$ cointegrated pairs in the system (Asche et al., 1999). If the system does not result in $n-1$ cointegration vectors, the number of cointegrated pairs will become less than $(n^2 - n)/2$. Initially, all oils are assumed to be cointegrated as we test all possible cointegration pairs in our system. With 10 variables, there are 45 possible pairs to investigate. The Johansen frame work is applied for the bivariate tests. Hence, the problem of normalization on one of the two variables is redundant.

In Table 5.6 the results of the pairwise cointegration tests for the full period are presented. The two first columns give the variables of interest, followed by the maximum eigenvalue test results and the trace test results. In the last column the test of the LOP is given, where we reject the LOP if the p-value is less than 0.05. It must be emphasized that the LOP is rejected if there is no cointegration vector between the two variables even though it is supported by the test statistics. The bivariate LOP is tested by imposing a restriction on $\beta' = (1, -1)$.

Fish oil is only integrated with palm kernel oil and corn oil. This is discouraging with respect to fish oil being integrated with the global vegetable oil market. Due to the shift in price levels in 2011, it is reasonable to assume fish oil as no longer integrated with the other oils in the long run. Clearly, the quality difference of the vegetable oils and fish oil causes the prices to behave different. Later, we split the data set into sub-periods which can help to explain how the fish oil price has developed over time with respect to the vegetable oils.

Soybean oil is one of the most produced vegetable oils globally, and the degree of market integration is strong for the full period. Soybean is cointegrated with other large producing oils such as palm oil, corn oil (no cointegration vector by the max-test, but the null of no cointegration vector is rejected by the trace test) and rapeseed oil. The null hypothesis of cointegration between soybean oil and the coconut oil, fish oil and linseed is rejected even at the 10 % critical value. Among the soybean pairs which have one cointegration vector, all pairs fail to reject the hypothesis of the LOP, except for corn oil which rejects it at the 5 % significance. The results are in compliance with the price development in Figures 2.7 and 2.8.

Sunflower oil strongly rejects the null of no cointegration with corn oil and soybean oil at 5 % significance for the full period, while it only rejects the null at 10 % significance from the trace-test with rapeseed oil, palm oil and palm kernel oil. If there is a true cointegration relationship (which is questionable at only 10 % significance level), the LOP is not rejected for any of the oils.

When evaluating the trace-test results of rapeseed oil, the null of no cointegration vector is rejected for soybean oil, corn oil, palm oil (10 % significance level), palm kernel oil and coconut oil (10 % significance level). However, the max test shows contradicting results. The trace-test is weighted more than the max-test, since it is considered as more robust against non-normal residuals, which could be the cause of no cointegration vectors from the max-test (Tveterås, 2000). The null of no cointegration between fish oil and rapeseed oil is failed to reject. As they both are ingredients in fish feed, we expected them to be at least close to cointegration. However, the quality, production volumes and different price movement in the last five years are significant, and could be argued as explanations for the result. The LOP is not rejected between rapeseed oil and the other oils. One reason for this that palm oil has approximately twice the global production volume as rapeseed oil (see Table 2.3), and can be argued to control the oil prices globally. How production volumes equals market power is beyond the scope of this thesis, but it is encouraging to think of a price leader as an oil with large volumes, which somewhat dictates the inferior oils by its own price movement.

Corn oil is similar to palm oil and palm kernel oil in quality, and the null of no cointegration is rejected for both tests (palm kernel at 10 % significance). However, the LOP is rejected at the 5 % level. The pairs of corn oil and fish oil/rapeseed oil rejects the null of no cointegration, and accepts the LOP hypothesis. The bivariate test for the full period is summarized in

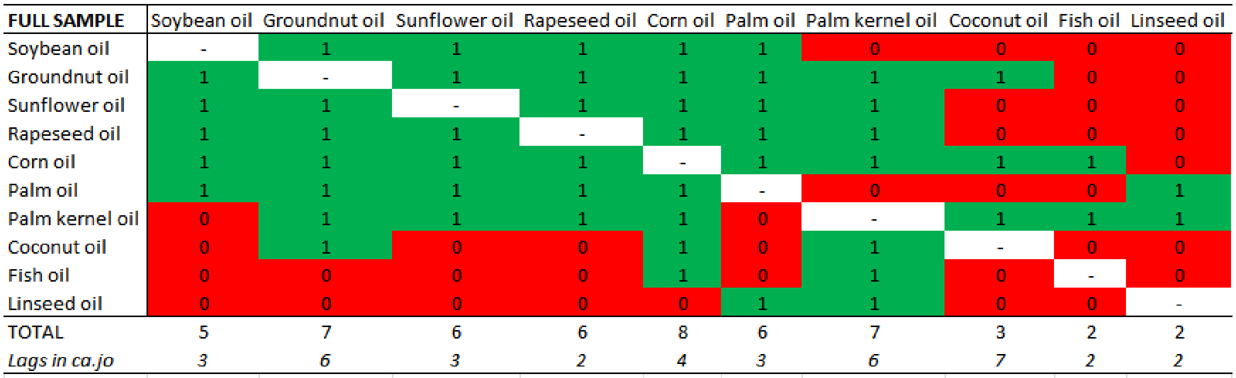
Table 5.6: Bivariate Johansen tests for the full period.

		Max Test		**Trace Test**		**LOP**
Variable 1	**Variable 2**	r==0	r<=1	r == 0	r<=1	p-value
Soybean oil	Groundnut oil	26.98**	3.61	30.58**	3.61	0.06
Soybean oil	Sunflower oil	15.46	4.64	20.10*	4.64	0.11
Soybean oil	Rapeseed oil	17.54*	4.99	22.53*	4.99	0.22
Soybean oil	Corn oil	24.34**	4.05	28.39**	4.05	0.02
Soybean oil	Palm oil	20.55**	3.05	23.60*	3.05	0.50
Soybean oil	Palm kernel oil	11.67	4.46	16.12	4.46	0.34
Soybean oil	Coconut oil	8.38	2.32	10.70	2.32	0.15
Soybean oil	Fish oil	6.60	4.34	10.93	4.34	0.93
Soybean oil	Linseed oil	11.78	3.93	15.71	3.93	0.78
Groundnut oil	Sunflower oil	20.49**	4.87	25.36**	4.87	0.45
Groundnut oil	Rapeseed oil	19.53*	5.37	24.91**	5.37	0.98
Groundnut oil	Corn oil	27.74**	4.27	32.01**	4.27	0.34
Groundnut oil	Palm oil	23.96**	3.52	27.48**	3.52	0.31
Groundnut oil	Palm kernel oil	16.99*	4.24	21.23*	4.24	0.35
Groundnut oil	Coconut oil	19.90	5.00	24.90	5.00	0.65
Groundnut oil	Fish oil	10.68	7.18	17.86	7.18	0.00
Groundnut oil	Linseed oil	10.31	3.67	13.98	3.67	0.60
Sunflower oil	Rapeseed oil	12.89	4.98	17.86	4.98	0.28
Sunflower oil	Corn oil	21.86**	6.29	28.15**	6.29	0.78
Sunflower oil	Palm oil	15.78*	3.52	19.30	3.52	0.03
Sunflower oil	Palm kernel oil	13.00	4.88	17.88	4.88	0.10
Sunflower oil	Coconut oil	11.16	3.92	15.08	3.92	0.06
Sunflower oil	Fish oil	8.12	5.93	14.05	5.93	0.40
Sunflower oil	Linseed oil	11.69	5.50	17.09	5.50	0.46
Rapeseed oil	Corn oil	16.31*	6.66	22.97*	6.66	0.19
Rapeseed oil	Palm oil	11.70	6.41	18.12	6.41	0.15
Rapeseed oil	Palm kernel oil	12.27	8.35	20.62*	8.35	0.19
Rapeseed oil	Coconut oil	9.55	7.93	17.48	7.93	0.21
Rapeseed oil	Fish oil	6.87	2.21	9.08	2.21	0.34
Rapeseed oil	Linseed oil	9.67	4.81	14.48	4.81	0.88
Corn oil	Palm oil	22.80	3.12	25.92	3.12	0.01
Corn oil	Palm kernel oil	14.78	5.20	19.98	5.20	0.04
Corn oil	Coconut oil	13.98	13.75	18.65	4.67	0.02
Corn oil	Fish oil	11.67	9.64	21.31*	9.64	0.27
Corn oil	Linseed oil	12.35	4.62	16.67	4.62	0.49
Palm oil	Palm kernel oil	12.11	3.06	15.17	3.06	0.24
Palm oil	Coconut oil	6.73	2.87	9.60	2.87	0.30
Palm oil	Fish oil	10.02	3.71	13.73	3.71	0.66
Palm oil	Linseed oil	13.31	3.69	17.00	3.69	0.58
Palm kernel oil	Coconut oil	16.40*	3.13	19.54	3.13	0.09
Palm kernel oil	Fish oil	14.46	6.56	21.02*	6.56	0.62
Palm kernel oil	Linseed oil	14.87	2.99	17.86	2.99	0.33
Coconut oil	Fish oil	11.37	3.48	14.85	3.48	0.97
Coconut oil	Linseed oil	10.60	2.23	12.84	2.23	0.14
Fish oil	Linseed oil	5.81	5.50	11.30	5.50	0.58

Critical values for the Max-test is 15.7 for H_0: r==0 and 9.2 for H_0: r<=1. Critical values for the Trace-test is 20.0 for H_0: r==0 and 9.2 for H_0: r<=1. ** indicates 1 % significance level, * indicates 5 % significance level.

Table 5.7, where a green cell indicates that the null of no cointegration was rejected at the 10 % significance (hence, it is not strong evidence of cointegration) and a red cell shows the pairs of no cointegration. There is a clear relationship between some of the vegetable oils, whereas fish oil is not considered as a part of the global vegetable oil market based on the results from the bivariate cointegration tests over the full period. Previously, fish oil was argued to diverge from the other vegetable oils at around June 2011. As the period between 2011 to 2015 represents approximately 1/3 of the period, the lack of cointegration with fish oil could be related to this event of divergence in price. Thus, dividing the full period into sub samples could shed further light on the relationship among the oils, especially when compared to the outcome from the full period.

Table 5.7: Summary of the full period bivariate Johansen test. Green cell indicates cointegration, whereas red means no cointegration. There is symmetry around the diagonal, $a_{ij} = a_{ji}$.



FULL SAMPLE	Soybean oil	Groundnut oil	Sunflower oil	Rapeseed oil	Corn oil	Palm oil	Palm kernel oil	Coconut oil	Fish oil	Linseed oil
Soybean oil	-	1	1	1	1	1	0	0	0	0
Groundnut oil	1	-	1	1	1	1	1	1	0	0
Sunflower oil	1	1	-	1	1	1	1	0	0	0
Rapeseed oil	1	1	1	-	1	1	1	0	0	0
Corn oil	1	1	1	1	-	1	1	1	1	0
Palm oil	1	1	1	1	1	-	0	0	0	1
Palm kernel oil	0	1	1	1	1	0	-	1	1	1
Coconut oil	0	1	0	0	1	0	1	-	0	0
Fish oil	0	0	0	0	1	0	1	0	-	0
Linseed oil	0	0	0	0	0	1	1	0	0	-
TOTAL	5	7	6	6	8	6	7	3	2	2
Lags in ca.jo	*3*	*6*	*3*	*2*	*4*	*3*	*6*	*7*	*2*	*2*

In Tables B.7 to B.9 the results for the three sub samples are presented in same context as Table 5.6. Remember, the price developments are plotted in Figures A.5a to A.5c. At first glance Sub-sample 1 appear to have a common trend. This is confirmed by the cointegration tests. Soybean oil is characterized as cheap compared to the other oils in this period (see Table B.5). Especially the drop in price around 2004/05 can be considered as the origin of the low relative prices. The null of no cointegration is rejected for soybean oil in pair with groundnut oil, sunflower oil, corn oil (max-test at 5 %), palm oil and palm kernel oil (only trace-test). There is evidence of integration between these oils. Compared to the full period, palm kernel oil had no cointegration relationship with soybean oil, whereas it does now. This result suffers from presence of autocorrelation, which have resulted in a spurious regression. Palm oil and soybean oil, with a relative price of 1.15 with soybean as the premium oil, follows the same trend through this period. Hence, the result of one cointegration vector is reliable. Same can be said for corn oil. In pair with rapeseed oil, soybean was cointegrated over the full sample. However, in this sub period the null hypothesis is rejected even at 10 % significance level. This indicates that either soybean or rapeseed oil has found a "new" market in this period (the latter is shown to be the case). The LOP hypothesis is accepted for all pairs where one cointegration vector is accepted.

Sunflower oil, which was cointegrated with six oils over the full period, now only appears as cointegrated (weakly) with corn and fish oil. In addition the LOP for sunflower/corn oil is rejected at 1 % significance level.

Another interesting event in the period of Sub-sample 1, is how rapeseed oil has become cointegrated with fish oil. There is strong statistical evidence of co-movement between fish oil and rapeseed oil in the period between 2003-2009. The relationship between palm kernel oil and fish oil has evaporated. There is statistical evidence of corn oil not being cointegrated with fish oil, which can be explained by the high prices of palm kernel oil in the period. Ideally, the intercept should encountered the price differential of transportation included in the contracts, such that the results may be an outcome of spurious regression.

To summarize Sub-sample 1, we create a corresponding table as Table 5.7, but subtracts the values from Sub-sample 1 ($r = 1$ or 0, for cointegration and no cointegration, respectively) by the corresponding cointegration vector (1 or 0 from Table 5.7). Hence, the possible outcomes are given as:

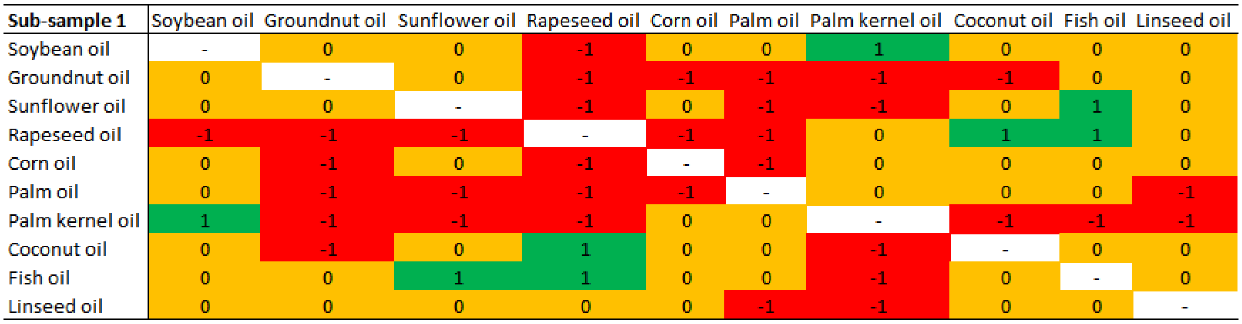
(a) $r_{\mathrm{Sub-Full}} = -1$ (red), which indicates that the oil pair is cointegrated over the full period, while not cointegrated in Sub-sample 1.

(b) $r_{\mathrm{Sub-Full}} = 0$ (orange), which indicates that the oil pair is either cointegrated or not cointegrated for both the full period and Sub-sample 1.

(c) $r_{\mathrm{Sub-Full}} = 1$ (green color), which indicates that the oil pair is not cointegrated over the full period, while it is cointegrated in Sub-sample 1

Thus, Sub-sample 1 is presented in Table 5.8 as the difference in cointegration state compared to the full period. Clearly most of the oil pairs have same cointegration relationship as over the full period, while some oils show a de-integration where it is cointegrated over the full period and not over the Sub-sample 1 period. A few oil pairs have gone from no cointegration to cointegrated: soybean/palmkernel (spurious), sunflower/fish and rapeseed/fish. Hence, fish oil was more cointegrated with vegetable oils than over the full period. Therefore, it must have been an event which occurred after 2009 which caused fish oil to loose its relationship with sunflower oil and rapeseed oil, since these cointegration relationships were non-existing over the full period.

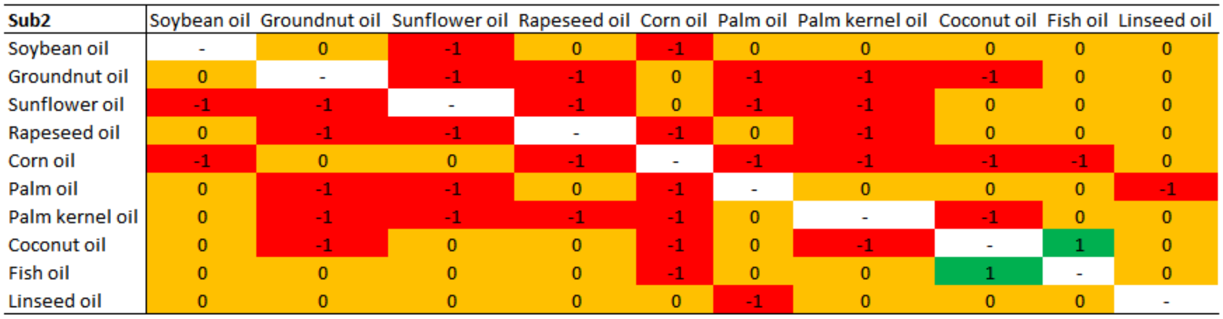
Table 5.8: Summary of the bivariate Johansen test over the Sub-sample 1 period compared to the results from the full period. Color scale is described in the list above.



Sub-sample 1	Soybean oil	Groundnut oil	Sunflower oil	Rapeseed oil	Corn oil	Palm oil	Palm kernel oil	Coconut oil	Fish oil	Linseed oil
Soybean oil	-	0	0	-1	0	0	1	0	0	0
Groundnut oil	0	-	0	-1	-1	-1	-1	-1	0	0
Sunflower oil	0	0	-	-1	0	-1	-1	0	1	0
Rapeseed oil	-1	-1	-1	-	-1	-1	0	1	1	0
Corn oil	0	-1	0	-1	-	-1	0	0	0	0
Palm oil	0	-1	-1	-1	-1	-	0	0	0	-1
Palm kernel oil	1	-1	-1	-1	0	0	-	-1	-1	-1
Coconut oil	0	-1	0	1	0	0	-1	-	0	0
Fish oil	0	0	1	1	0	0	-1	0	-	0
Linseed oil	0	0	0	0	0	-1	-1	0	0	-

The test results over the period from 2009 until the end of the observations are given in Table B.8. Among the interesting results is how soybean oil appears to be cointegrated with rapeseed oil, which was rejected in previous period. Another very important result is the rejection of cointegration between fish oil and rapeseed oil in this period. The differences of pairwise cointegration relationships between Sub-sample 2 and the full period are presented in Table 5.9, with the associated Johansen test results given in Table B.8. The number of lags, which was specified to reduce autocorrelation and ARCH effects, may have caused some pairs to behave as un-cointegrated even though they actually are cointegrated. However, we proceed with explaining the reported results. The only "green" cell in Table 5.9 is between fish oil and coconut oil. The lagged effect is clearly a factor of affections between the two. From Figure A.5b, the price movement between fish oil and coconut oil is similar, but in a different time and scale. Majority of the oil pairs have the same cointegration relationship as over the full period (orange cells). The evidence of no cointegration for most of the oils can be related to large price variation during this period: The price increased in the aftermath of the financial crisis, followed by a decreasing trend (except for fish oil and groundnut oil). The rate at these up- and downward trend occurrences is somewhat different in terms of magnitude and time for each single oil.

Table 5.9: Summary of the bivariate Johansen test over the Sub-sample 2 period compared to the results from the full period. Color scale is the same as for Table 5.8.

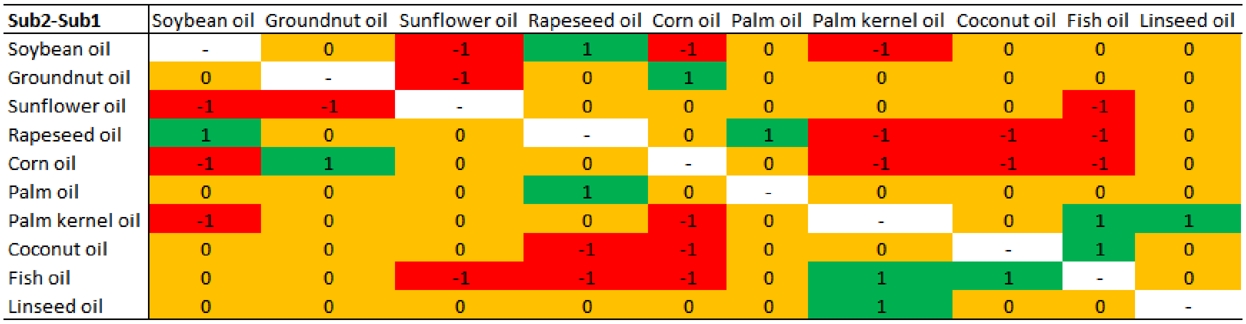


Sub2	Soybean oil	Groundnut oil	Sunflower oil	Rapeseed oil	Corn oil	Palm oil	Palm kernel oil	Coconut oil	Fish oil	Linseed oil
Soybean oil	-	0	-1	0	-1	0	0	0	0	0
Groundnut oil	0	-	-1	-1	0	-1	-1	-1	0	0
Sunflower oil	-1	-1	-	-1	0	-1	-1	0	0	0
Rapeseed oil	0	-1	-1	-	-1	0	-1	0	0	0
Corn oil	-1	0	0	-1	-	-1	-1	-1	-1	0
Palm oil	0	-1	-1	0	-1	-	0	0	0	-1
Palm kernel oil	0	-1	-1	-1	-1	0	-	-1	0	0
Coconut oil	0	-1	0	0	-1	0	-1	-	1	0
Fish oil	0	0	0	0	-1	0	0	1	-	0
Linseed oil	0	0	0	0	0	-1	0	0	0	-

The development of cointegration relationship from the commodity boom period, to the aftermath of the financial crisis is of great interest. This is presented in Table 5.10 as the differences in cointegration relationship between the two periods. A *green* cell indicates that the price relationship was not cointegrated in Sub-sample 1 whereas cointegrated in Sub-sample 2. A *red* cell describes the opposite, the pair was cointegrated in Sub-sample 1 and not in Sub-sample 2. A zero-cell means a constant relationship between the two oils in both periods; either cointegration or no cointegration in both sample periods. Among the major oils there are some interesting trends. Soybean oil is yet again cointegrated with rapeseed oil, while fish oil is no longer cointegrated with rapeseed oil. Rapeseed also becomes cointegated with palm oil in the post financial crisis period. Corn oil shows statistical evidence of loosing cointegration partners in the Sub-sample 2 period. This result may be due to wrongly concluding the non-existence of a relationship in the period of 2009-15, since the full period shows there is a relationship which is supported by the price movements. Therefore, considering these bivariate test results it is important to keep in mind that the small power of the test can suffer from having a small sample size. Overall, the bivariate test for the full period of January 2000 until December 2015, are weighted more than each single sub-period as we fear that the small samples may have contaminated our tests as we manipulated number of lags to achieve white noise residuals – increasing number of lags equals loss of initial information .

Finally, the period from June 2011 until the end of our sample is briefly discussed. We hypothesized that since the price-rally stopped in middle of 2011 and started a decreasing trend for most of the vegetable oils, there could be a potential common trend for most of the oils. In

Table 5.10: Summary of the bivariate Johansen test over the Sub-sample 2 compared to Sub-sample 1. Color scale is the same as for Table 5.8.



Sub2-Sub1	Soybean oil	Groundnut oil	Sunflower oil	Rapeseed oil	Corn oil	Palm oil	Palm kernel oil	Coconut oil	Fish oil	Linseed oil
Soybean oil	-	0	-1	1	-1	0	-1	0	0	0
Groundnut oil	0	-	-1	0	1	0	0	0	0	0
Sunflower oil	-1	-1	-	0	0	0	0	0	-1	0
Rapeseed oil	1	0	0	-	0	1	-1	-1	-1	0
Corn oil	-1	1	0	0	-	0	-1	-1	-1	0
Palm oil	0	0	0	1	0	-	0	0	0	0
Palm kernel oil	-1	0	0	0	-1	0	-	0	1	1
Coconut oil	0	0	0	-1	-1	0	0	-	1	0
Fish oil	0	0	-1	-1	-1	0	1	1	-	0
Linseed oil	0	0	0	0	0	0	1	0	0	-

Table B.9 the results from this period is given. When a less strict rejection criteria, say 10 %, is considered all oils are considered as cointegrated with soybean oil (except for coconut oil and corn oil). This may indicate that the price increase observed in the period 2009-11 may have caused the oils to loose their cointegration relationship. Whereas when the trends decreased, the oils started to behave with somewhat same fluctuation. Hence, if one arbitrary oil price increased for a short period, then it would affect some other oils. The same can be concluded for rapeseed oil. Whereas, fish oil has stabilized at a higher price level, and there is no evidence of cointegration with other oils than soybean oil for this period. Hence, the results are contradicting. Since, soybean oil seems to be cointegrated with most oils, inclusive fish oil, it contradicts the test result of fish oil being cointegrated with soybean oil in this period. This period is deemed as short in terms of cointegration tests, and must not be weighted too much. However, based on price developments and cointegration tests, we can say that fish oil is no longer cointegrated with the vegetable oils.

Finally, there is not clear that there exists a global vegetable oil market. Weak statistical evidence indicates that there could potentially be one market for the most produced oils such as soybean oil, rapeseed oil, sunflower oil and palm oil in the *long run*. Whereas the less produced oils, show little degree of market integration, and the results vary largely when the period is split into shorter time periods. Discouraged about the pairwise results, we proceed by checking the system simultaneously by the multivariate cointegration approach for completeness of our study.

5.3.2 Multivariate Cointegration of the Full System

In the previous section the pairwise cointegration test results have already rejected the law of one price for the full system. Tveterås (2000) explains that when a full bivariate study has been conducted, the multivariate test is somewhat superfluous. However, the tests for multivariate cointegration can give interesting results, which can help interpret the global vegetable oil market. Table 5.11 presents the results of the joint cointegration test for the ten oils over the full sample period. The trace statistics reject null hypothesis of two cointegration vectors at 5 % significance, while it fail to reject the null of three stationary linear combinations. The results from the maximum eigenvalue test suggest the same result, although the null hypothesis of two cointegration vectors is only rejected at 10 % significance level. Hence, it could exist three linear combinations which are stationary. There is therefore statistical evidence of all oils being cointegrated and supports the theory of an integrated market in the long run.

Table 5.11: Multivariate cointegration test results over the full period, Jan. 2000 to Dec. 2015, with 6 lags.

H_0	**Trace-test**		**Max-test**	
	Test statistic	5 % critical value	Test statistic	5 % critical value
r<=9	2.52	9.24	2.52	9.24
r<=8	10.26	19.96	7.74	15.67
r<=7	18.67	34.91	8.41	22.00
r<=6	34.43	53.12	15.77	28.14
r<=5	60.00	76.07	25.57	34.40
r<=4	88.21	102.14	28.21	40.30
r<=3	122.39	131.70	34.18	46.45
r<=2	171.49**	165.58	49.11*	52.00
r<=1	245.72***	202.92	74.23***	57.42
r==0	344.15***	244.15	98.43***	63.57

* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

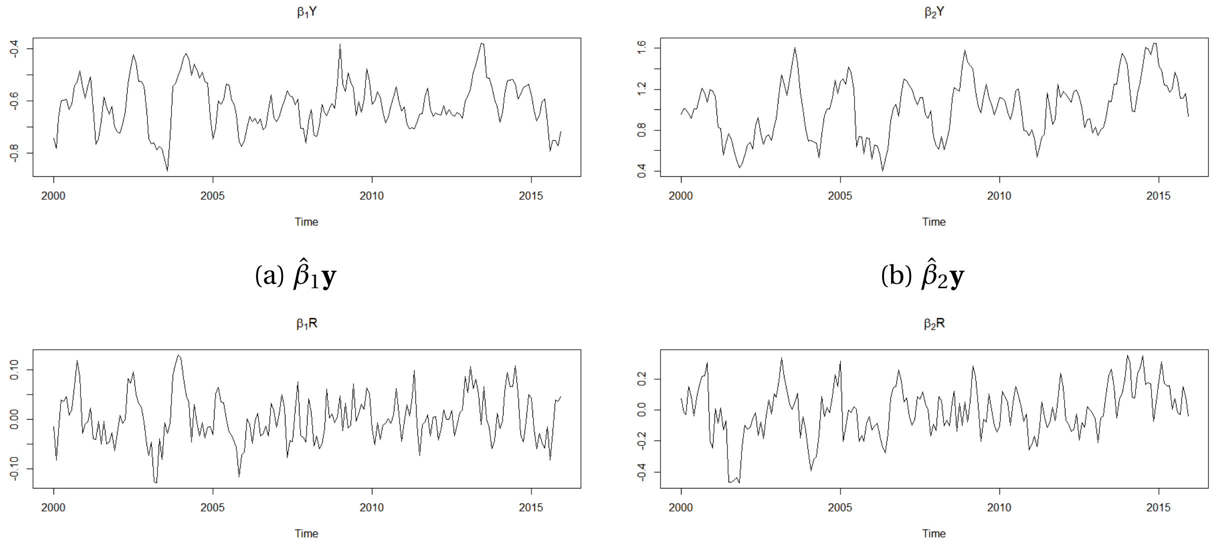
One global oil market was not supported by the bivariate test results. Looking at the coefficients for the different oil in the first cointegration vector given in Table 5.12, we see that fish oil is barely influencing the long run linear relationship with a vector coefficient of -0.005. Hence, fish oil is somewhat excluded from the linear combination which creates a long run linear relationship. This makes sense, based on lack of cointegration between fish oil and the vegetable oils in the long run. The two first cointegration vectors are assumed to create stationary linear combinations with the variables. A plot of $\hat{\beta}_1\mathbf{y}$ and $\hat{\beta}_2\mathbf{y}$ should behave as stationary processes, where $\hat{\beta}_i$ are the estimated cointegration vectors given in Table 5.12 and $\mathbf{y}$ is the 10×1 time series vector. Figures 5.5a and 5.5b present the long-run stationary processes. Short-run influences can affect the linear combinations, and yield false results. Pfaff (2008) states that by multiplying the residuals from the lagged variables in level with the variable vector, we can observe whether the residuals behave as white noise around a zero-mean or not, which is confirmed in Figures 5.5c and 5.5d. Therefore, the estimated cointegration vectors seem to form two stationary processes which support the findings of cointegration between the oils.

Table 5.12: The two cointegration vectors for the full period, normalized upon the first and second variable in the data set, respectively.

Variable	$\hat{\beta}_1$	$\hat{\beta}_2$
Soybean oil	1.000	-1.255
Groundnut oil	-0.262	1.000
Sunflower oil	0.421	-0.927
Rapeseed oil	-0.522	-0.770
Corn oil	-0.590	-0.660
Palm oil	-0.151	2.383
Palm kernel oil	-0.022	-2.054
Coconut oil	-0.060	1.473
Fish oil	-0.005	-0.156
Linseed oil	0.128	1.081
Constant	0.616	-1.056

The jointly cointegration results from the three sub periods of our attention are presented in Tables 5.13 to 5.15. In Sub-sample 1, the trace test suggests five cointegration vectors at 10 % significance level, while the maximum eigenvalue test suggests four. Based on the pairwise coitnegration test, four cointegration vectors seem more reasonable than five. The LOP is tested by imposing the restrictions from Equation (4.37). The null hypothesis of an existing LOP was rejected with a chi-squared statistics of 39.81 and p-value of 0.0001, hence strongly rejection as expected based on the results from our bivariate tests.

The multivariate cointegration test on the observations from 2009 to the end of our sample is given in Table 5.14. The results from the trace and eigenvalue tests are inconsistent. The trace test, which is considered as more robust when the underlying distribution is not normal,



(a) $\hat{\beta}_1\mathbf{y}$

(b) $\hat{\beta}_2\mathbf{y}$

(c) $\hat{\beta}_1\mathbf{R}_1$

(d) $\hat{\beta}_2\mathbf{R}_2$

Figure 5.5: Graphical presentation of the two cointegration vectors from the full period.

Table 5.13: Multivariate cointegration test results of Sub-sample 1 period, Jan. 2003 to Dec. 2008, with 2 lags.

H_0	**Trace-test**		**Max-test**	
	Test statistic	5 % critical value	Test statistic	5 % critical value
r<=9	2.85	8.18	2.85	8.18
r<=8	13.13	17.95	10.28	14.90
r<=7	26.10	31.52	12.97	21.07
r<=6	44.97	48.28	18.87	27.14
r<=5	67.16*	70.60	22.20	33.32
r<=4	105.60**	90.39	38.44*	39.43
r<=3	155.31***	124.25	49.71**	44.91
r<=2	218.01***	157.11	62.70***	51.07
r<=1	306.01***	192.84	88.00***	57.00
r==0	407.61***	232.49	101.6***	62.42

* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

suggests that the rank is five. The maximum eigenvalue test finds only three tests. Based on the variables tested, three cointegration vectors are more in compliance with the results from the bivariate tests for this sample period. However, the oils are considered as cointegrated and form a long-run stationary process. Since many oil pairs rejected the existence of a cointegration relationship in the bivariate test for Sub-sample 1 are in accordance with the presence of six common stochastic trends (number of no-cointegration).

Table 5.14: Multivariate cointegration test results of Sub-sample 2, Jan. 2009 to Dec. 2015, with 2 lags.

H_0	**Trace-test**		**Max-test**	
	Test statistic	5 % critical value	Test statistic	5 % critical value
r<=9	2.36	9.24	0.79	8.18
r<=8	11.31	19.96	8.59	14.90
r<=7	28.38	34.91	15.96	21.07
r<=6	47.33	53.12	18.92	27.14
r<=5	70.21	76.07	22.79	33.32
r<=4	104.73**	102.14	34.22	39.43
r<=3	148.64***	131.70	43.64*	44.91
r<=2	212.18***	165.58	62.43***	51.07
r<=1	279.99***	202.92	67.73***	57.00
r==0	359.31***	244.15	78.99***	62.42

* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

Johansen cointegration test results on the final sub period, from June 2011 to December 2015, are presented in Table 5.15. Sub-sample 3 had more bivariate cointegration pairs than e.g. Sub-sample 1, and the multivaraite test results are in accordance with this. The trace test rejects the hypothesis of having six cointegration vectors, and fails to reject the seventh. The maximum eigenvalue test also suggests seven linear combinations which will cause the variables to behave as stationary in the long-run. Based on the price developments in this period, the result from the multivariate test is more reliable than the bivariate test results. Actually, the two different tests produces somewhat contradicting answers. The prices must be considered as integrated in this period, except for the fish oil and groundnut oil which are traded at premium compared the other oils.

To summarize this section, there is statistical evidence of cointegration between the oils when utilizing the Johansen approach. However, the LOP hypothesis is rejected for all systems. Comparing bivariate and multivariate test, there is somewhat contradicting results. Visually, it appears to be a common trend. However, this observed trend is not necessary stationary, which the biviariate tests showed. Not all pairs have a stationary linear combination between them. When testing for cointegration for all oils as one system, there is presence of more than one

Table 5.15: Multivariate cointegration test results of Sub-sample 3, Jun. 2011 to Dec. 2015, with 4 lags.

H_0	**Trace-test**		**Max-test**	
	Test statistic	5 % critical value	Test statistic	5 % critical value
r<=9	5.19	9.24	3.05	8.18
r<=8	13.99	19.96	6.22	14.90
r<=7	29.15	34.91	14.76	21.07
r<=6	61.57**	53.12	31.49**	27.14
r<=5	97.98***	76.07	34.19***	33.32
r<=4	144.43***	102.14	46.28***	39.43
r<=3	211.17***	131.70	66.06***	44.91
r<=2	295.14***	165.58	83.14***	51.07
r<=1	408.04***	202.92	111.24***	57.00
r==0	557.71***	244.15	148.08***	62.42

* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

cointegration for all periods studied. Hence, we can say that there exists a linear combination of all oils which causes them to behave as integrated in the long-run - therefore we believe that the oils are in one market. For a deeper understanding of which oil affects the other, next section completes our study by investigating the causality and weak exogeneity of each variable.

5.4 Weak Exogeneity – $\alpha_i = 0$

To investigate whether one oil drives the price of other oils, we test how the prices are affected by each other in the long-run. This equal to test each variable in the system for weak exogeneity (lon-run noncausality). Johansen (1991) proposed that weak exogeneity can be tested by imposing restrictions on the adjustment matrix α (where $\Pi = \alpha\beta'$). Hence, we restrict all the row-elements in α to be zero. Thus, the null hypothesis is that $\alpha_{i,1} = \alpha_{i,2} = \ldots = \alpha_{i,9} = 0$, where i is the row associated with the oil being tested as weakly exogenous. All the ten oils are being evaluated. We performed multivariate weak exogeneity test for the full period, and on all sub periods for the bivariate test, investigating long-run effects. The test statistics for the multivariate system are distributed as χ^2 with a 10 %, 5 % and 1 % significance level of 4.61, 5.99 and 9.21, respectively. The test results are presented in Table 5.16. The null hypothesis's of soybean oil, sunflower oil, corn oil, palm oil and linseed oil being exogenous are all rejected. Remember that we had some ambitious results for groundnut oil, palm kernel oil and coconut oil, we will not put much weight on these results. Based on the test results, rapeseed oil and fish oil are both weakly exogenous. The result indicates that fish oil is not influenced by the movement on the other oil prices, but it may affect them. Rapeseed oil also exhibits no affection by the price of the other oils, while it may act as a price leader.

Table 5.16: Test for weak exogeneity over the full period, for all variables.

Potential Exogenous Variable	**Test Statistic**	**p-value**
Soybean oil	30.91***	0.00
Groundnut oil	3.04	0.22
Sunflower oil	7.27**	0.03
Rapeseed oil	1.34	0.51
Corn oil	8.56**	0.01
Palm oil	8.92**	0.01
Palm kernel oil	0.15	0.93
Coconut oil	0.05	0.98
Fish oil	1.06	0.60
Linseed oil	9.70***	0.01

* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

The bivariate cointegration test did not result in all pairs being cointegrated. Hence, the weakly exogeneity test by imposing restriction on the adjustment matrix α, which requires that there is at least one cointegration vector in the system, is not valid for all pairs of oil. However, we performed weak exogeneity tests on all pairs and periods. The results are presented in Table B.10, but must be evaluated with respect to the results of the bivariate test results for the respective period (see Tables B.7 to B.9). The first two columns in Table B.10 gives the variables of interest. For each period, there are two columns denoted as $\alpha_1 = 0$ and $\alpha_2 = 0$. which indicates that we test variable 1 or variable 2 as weakly exogenous, respectively. The null hypothesis is that the variable is exogenous to the other oil. Hence, a p-value less than 0.05 (or 0.10 depending on the criteria evaluated), rejects weak exogeneity. Soybean oil seems to be weakly exogenous towards all oils except palm oil in the full period. In sub-sample 1, fish oil is weakly exogenous to soybean oil. Hence, fish oil has no influence on soybean oil over the full period, while in the commodity boom period, where fish oil increased up to a price level with rapeseed oil, soybean oil prices are better understood by including the fish oil prices. This relationship is lost in the last part of our data. Soybean oil is weakly exogenous to fish oil from 2011 to 2015.

There is statistical evidence of Sunflower oil being weakly exogenous to the other oils, except fish oil over the full period. This is due to the price increase in fish oil between 2003 to 2008, where fish oil price leveled up to the other oils. As fish oil price diverged, sunflower is again found to be exogenous to fish oil. Hence, fish oil and sunflower oil is better predicted by including the lagged values by each other.

Rapeseed oil is weak exogenous to all oils, except for fish oil in the period between January 2009 until the end 2015. However, the opposite is not supported by the statistics. All oils, except palm oil, rejects the null of no influence by rapeseed oil. Hence, rapeseed oil shows evidence of having the greatest influence on the other oil prices. Hence, we have found a potential price leader among our oils. Remember, rapeseed oil showed signs of being cointegrated with the other oils under the bivariate cointegrations tests which now is supported by its strong influence on the other oil prices.

Fish oil has a trend of being weakly exogenous in the last part of our sample, while it is influenced by the other oil price up to 2011. This is in compliance with the price development. From 2011, fish oil diverged from the other oils, and it can be argued that from this point of time fish oil price both exercise and have no influence on the other oil prices, which is observed since the null of no influence is accepted for most pairs.

Test statistics from corn oil-pairs, suggest that corn oil has a degree of affection towards the other oils in the commodity boom period (as the null is rejected). Despite this period, corn oil can be considered as weakly exogenous in the last part and over the full sample period. Pairs including palm oil suggests that palm oil may influence the other prices over the full period, whereas it exhibit weak exogeneity in the commodity boom period.

To summarize the weak exogeneity test, we conclude that the oils show a tendency to shift their characteristics as weakly exogenous depending on the time period. There is a trend of the vegetable oils being better predicted including the fish oil price, and vice versa, in the commodity boom era, while the fish oil price is weakly exogenous in the last part of our sample. Finally, rapeseed oil shows strong statistical evidence of influencing the other oils – hence, it can be considered as a price leader in our system. Further information of the relationship between different oils can be found by a causality test.

5.5 The Toda and Yamamoto Causality Test

Working around the problem of potential no-cointegration pairs, we apply the Toda and Yamamoto (1995) causality test explained in section 4.3.4. In addition to test if one time series Granger-cause another oil, the causality test provides a possibility for cross checking the final results from the cointegration tests. The causality test results are presented in Tables B.11 to B.13 for the Full period, Sub-sample 1 and Sub-sample 2, respectively. Sub-sample 3 is omitted in this part of our study. The null hypothesis tested is that y_i does not Granger-cause y_j, for $i \neq j$. The *p*-value is reported in the rightmost column. Unreported diagnostic tests have been performed on all tests, and serial correlation was removed with same lags as mentioned in our discussion of the bivariate cointegration tests (see section 5.3.1).

Table 5.17 presents the results from the pairwise cointegration tests *and* the causality tests. The variables tested are listed in two first columns of Table 5.17. The *CI vector* columns contain either a *Yes* or *No* to indicate whether there is one cointegration (*CI*) vector or not, respectively, between the oils given in the two leftmost columns. The causality results from Tables B.11 to B.13 are summarized in Table 5.17 in the *Causality* columns. There are four different outcomes; (1) Only y_1 Granger-causes y_2 denoted as $y_1 \rightarrow y_2$, (2) only y_2 Granger-causes y_1 denoted as $y_2 \rightarrow y_1$, (3) y_1 Granger-causes y_2 and y_2 Granger-causes y_1 denoted as $y_1 \leftrightarrow y_2$, and (4) there is no causality between the two variables represented by $y_1 \neq y_2$. Overall the causality seem to match our previous results of bivariate cointegration.

For the full sample period, soybean oil Granger-causes rapeseed oil (slightly), and there was also one cointegration vector between the two. As soybean oil and rapeseed looses its stationary relationship during the commodity boom, the causality effects disappears. From 2009 until the end of our sample period, soybean and rapeseed oil forms one cointegration vector and there is bidirectional causality effects. Soybean oil also Granger-causes linseed oil, whereas there was no cointegration relationship between the two over the full sample. There is statistical evidence of bidirectional Granger-causality between soybean and corn oil. Soybean oil and corn oil were found to form a stationary relationship over the full period and in Sub-sample 2 period, but not during the commodity boom era. Groundnut oil, palm kernel oil and coconut oil, all demand many lags to achieve white noise residuals, and it is believed that these results are ambitious. There is no causality effects between soybean oil and fish oil, as the null hypothesis can not be rejected in neither directions. Soybean and fish oil are not cointegrated in neither periods, however there is statistical evidence of fish oil affecting soybean oil prices in the commodity boom period.

Rapeseed oil Granger-causes corn oil (bidirectional), fish oil, linseed oil, sunflower oil and groundnut oil over the full period. Similar trend is observed during the Sub-sample 1 period. Whereas, rapeseed's strong position as the oil which affects the others is washed out in the Sub-sample 2 period. Rapeseed oil also exhibits more cointegrated with other oils during the commodity boom period, than in any other periods.

Corn oil showed strong causality effects towards palm oil, palm kernel oil, fish oil and linseed oil over the full period. Same trend was observed during 2003 to 2009. In the post-financial crisis period, the price developments of the oils showed similar trends, but at different rates and degree of fluctuations. This difference in price movement, have caused the prices to not be cointegrated, and no causality effects are observed.

Palm oil was only cointegrated with soybean oil, corn oil and groundnut oil over the full total period. During Sub-sample 3, no cointegration vector was suggested by the cointegration tests. This is in compliance with the causality results in Sub-sample 3, where there is no causality between palm oil and the others (expect for 10 % significance of fish oil causing palm oil, and bidirectional causality between palm and corn oil in the sub period 3).

Finally, fish oil did not Granger-cause any other oils but linseed oil during the full period, while statistics suggest that the other way around could be true. However, in the Sub-sample 2 period fish oil seem to Granger-cause soybean oil, groundnut oil, rapeseed oil (bidirectional) and palm kernel oil. In addition, fish oil formed more cointegration pairs during this period. Totally, we may conclude that fish oil had strong affection on most oils between 2003 and 2008. However, as the fish oil price diverged positively from the other oils during the break in June 2011, its power of influence on the other oils seems to disappear based on the lack of evidence of fish oil Granger-causing the other oils.

5.6 Is the Long-Run Relationship Linked to Quality Differences?

We proposed in section 2.3 that the even though the vegetable oils are generally considered as substitutes, the end-users have different preference towards them. Here, only the vegetable oils are evaluated, as we consider fish oil to have a very different quality than all the vegetable oils. In Table 5.18 the three groupings are listed. Firstly, *Group C* is not of evaluated as it is difficult to benchmark a common end-user for these "*independent*" oil.

Group B contains only two oils, which prices are given as CIF contracts, and have already been analyzed by the bivariate cointegration tests. Over the full period the maximum eigenvalue test suggested that coconut oil and palm kernel oil were integrated at 5 % significance. The trace test rejected cointegration at 5 %, but accepted at 10 % significance level. The LOP was weakly accepted for the two oils, and palm kernel was considered as weakly exogenous relative to coconut oil. By including the global production volumes in Table 2.3 into this discussion, this is expected as palm kernel oil production was approximately 7-8 times larger than coconut in 2013. In and Inder (1997) reported that cointegration between palm kernel oil and coconut oil was surprisingly rejected for data from 1976 to 1990. Further In and Inder (1997) concluded that "*...a long-run factor yet to be included in the model is at work*" for coconut oil and palm kernel oil. Thus, by evaluating the most recent price series for the two lauric oils, there is evidence that this long-run relationship has been established.

Finally, *Group A* which consist of the most produced edible oils will be evaluated. Only the long-run relationship over the full sample period is studied, as we have a feeling that the tests over the full period have yielded better results than the sub periods so far in our study. Preliminary diagnostic tests indicates that two lags are sufficient to obtain white noise (Breusch-Godfrey test gave a p-value of 0.192, normality test reported a p-value of 0.189 and ARCH-test resulted in a p-value of 0.190 – which indicates no rejection of the null hypothesis). It is interesting to see how, these variables included in one group only needed two lags to remove serial correlation, while the full system needed six lags.

For the multivariate cointegration test, 11 seasonal dummies and one constant were imple-

Table 5.17: Cointegration (CI) vectors and Granger-causality.

Variable 1	**Variable 2**	**Full**		**Sub 1**		**Sub 2**	
y_1	y_2	*CI Vector*	*Causality*	*CI Vector*	*Causality*	*CI Vector*	*Causality*
Soybean oil	Groundnut oil	Yes	$y_1 \rightarrow y_2$	Yes	$y_1 \rightarrow y_2$	Yes*	$y_1 \rightarrow y_2^*$
Soybean oil	Sunflower oil	Yes	$y_2 \rightarrow y_1$	Yes	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Soybean oil	Rapeseed oil	Yes	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	Yes	$y_1 \leftrightarrow y_2$
Soybean oil	Corn oil	Yes	$y_1 \leftrightarrow y_2$	No	$y_2 \rightarrow y_1$	Yes	$y_1 \leftrightarrow y_2$
Soybean oil	Palm oil	Yes	$y_2 \rightarrow y_1$	Yes	$y_1 \rightarrow y_2^*$	No	$y_1 \neq y_2$
Soybean oil	Palm kernel oil	No	$y_2 \rightarrow y_1$	Yes	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Soybean oil	Coconut oil	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Soybean oil	Fish oil	No	$y_1 \neq y_2$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Soybean oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Groundnut oil	Sunflower oil	Yes	$y_1 \leftrightarrow y_2$	Yes	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$
Groundnut oil	Rapeseed oil	Yes	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Groundnut oil	Corn oil	Yes	$y_2 \rightarrow y_1$	Yes*	$y_2 \rightarrow y_1$	Yes	$y_1 \neq y_2$
Groundnut oil	Palm oil	Yes	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Groundnut oil	Palm kernel oil	Yes	$y_1 \neq y_2$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Groundnut oil	Coconut oil	Yes	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \rightarrow y_2^*$
Groundnut oil	Fish oil	No	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Groundnut oil	Linseed oil	No	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Sunflower oil	Rapeseed oil	No	$y_2 \rightarrow y_1$	No	$y_1 \rightarrow y_2$	Yes	$y_1 \rightarrow y_2^*$
Sunflower oil	Corn oil	Yes	$y_1 \leftrightarrow y_2$	Yes*	$y_1 \rightarrow y_2$	Yes	$y_1 \leftrightarrow y_2$
Sunflower oil	Palm oil	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Sunflower oil	Palm kernel oil	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Sunflower oil	Coconut oil	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Sunflower oil	Fish oil	No	$y_1 \rightarrow y_2$	Yes*	$y_1 \rightarrow y_2$	No	$y_1 \rightarrow y_2^*$
Sunflower oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \rightarrow y_2$	No	$y_1 \rightarrow y_2$
Rapeseed oil	Corn oil	Yes	$y_1 \leftrightarrow y_2$	Yes*	$y_2 \rightarrow y_1^*$	Yes	$y_1 \leftrightarrow y_2$
Rapeseed oil	Palm oil	No	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Rapeseed oil	Palm kernel oil	Yes	$y_2 \rightarrow y_1$	Yes	$y_2 \rightarrow y_1^*$	No	$y_1 \neq y_2$
Rapeseed oil	Coconut oil	No	$y_1 \neq y_2$	Yes	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Rapeseed oil	Fish oil	No	$y_1 \rightarrow y_2$	Yes	$y_1 \leftrightarrow y_2$	No	$y_1 \rightarrow y_2$
Rapeseed oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2$
Corn oil	Palm oil	Yes	$y_1 \leftrightarrow y_2$	No	$y_1 \neq y_2$	No	$y_1 \leftrightarrow y_2$
Corn oil	Palm kernel oil	Yes	$y_1 \leftrightarrow y_2$	Yes	$y_2 \rightarrow y_1$	No	$y_2 \rightarrow y_1$
Corn oil	Coconut oil	No	$y_2 \rightarrow y_1$	Yes	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Corn oil	Fish oil	Yes	$y_1 \rightarrow y_2$	Yes	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Corn oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2^*$
Palm oil	Palm kernel oil	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Palm oil	Coconut oil	No	$y_1 \neq y_2$	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$
Palm oil	Fish oil	No	$y_1 \rightarrow y_2$	No	$y_1 \rightarrow y_2$	No	$y_2 \rightarrow y_1^*$
Palm oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	Yes*	$y_1 \rightarrow y_2^*$
Palm kernel oil	Coconut oil	Yes	$y_1 \neq y_2$	No	$y_1 \neq y_2$	No	$y_2 \rightarrow y_1^*$
Palm kernel oil	Fish oil	Yes	$y_1 \rightarrow y_2$	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$
Palm kernel oil	Linseed oil	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Coconut oil	Fish oil	No	$y_1 \rightarrow y_2$	No	$y_1 \rightarrow y_2$	Yes*	$y_1 \neq y_2$
Coconut oil	Linseed oil	No	$y_1 \rightarrow y_2$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$
Fish oil	Linseed oil	No	$y_2 \rightarrow y_1$	No	$y_1 \neq y_2$	No	$y_1 \neq y_2$

* indicates significant at the 10 % level.

Table 5.18: Vegetable oils divided in three groups based on their quality and end-users.

Group A *General oils*	**Group B** *Lauric oils*	**Group C** "*Independent*"
Soybean Sunflower Rapeseed Palm	Coconut Palm kernel	Groundnut Linseed Corn

mented in the model (same as before). The trace test has been deemed more robust for the oil prices, and is reported in Table 5.19. The *general oils* have two cointegration vectors, which form stationary linear combinations in the long-run. Hence, the *general oils* are considered as cointegrated.

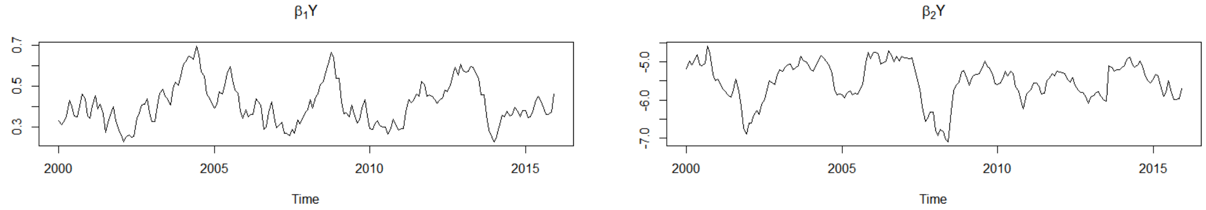
Table 5.19: Multivariate cointegration trace test results for the *general oils*.

H_0	Test statistic	10 % critical value	5 % critical value	1 % critical value
r <= 3	4.52	7.52	9.24	12.97
r <= 2	19.24*	17.85	19.96	24.60
r <= 1	36.23**	32.00	34.91	41.07
r = 0	62.68***	49.65	53.12	60.16

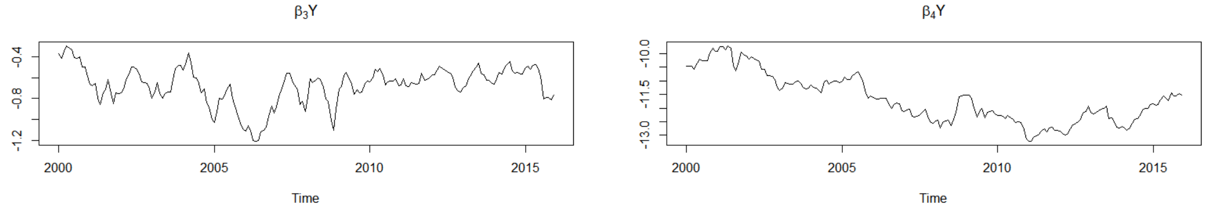
* indicates test statistics at 10 % significance level. ** indicates test statistics at 5 % significance level. *** indicates test statistics at 1 % significance level.

The trace test indicates two cointegration vectors at 5 % significance level, however based on our believe of the oils being highly integrated we plot the processes $\hat{\beta}_i\mathbf{y}$ for $i = 1, 2, 3, 4$ in Figure 5.6. Clearly, the fourth cointegration vector is not stationary as suggested by the multivariate test. However, we can argue that the third cointegration vector actually exists by looking at Figure 5.6c. The unreported disequilibrium process $\hat{\beta}_i\mathbf{R}$ illustrates the same results.

These four oils are considered as integrated, and we believe that there exists one market for soybean, sunflower, rapeseed and palm oil in the long-run. The LOP is tested by imposing the restriction in Equation (4.37), and accepted with an F-statistics of 2.65 and p-value of 0.45. One market for the *general oils* seems reasonable since we found that soybean oil was bi-cointegrated with the three other oils in the group (Table 5.11).



(a) $\hat{\beta}_1\mathbf{y}$ (b) $\hat{\beta}_2\mathbf{y}$



(c) $\hat{\beta}_3\mathbf{y}$ (d) $\hat{\beta}_4\mathbf{y}$

Figure 5.6: Graphical presentation of the four cointegration vectors from the *general oils* system over the full period.

Chapter 6

Discussion and Conclusion

The purpose of this thesis was to elucidate the existence of *one global vegetable oil market.* We had concerns of whether fish oil could be considered as a member of such a global market based on the difference in quality of fish oil compared to vegetable oils. In addition we wanted to test the existence of one global vegetable oil market by cointegration and market integration. In this chapter we will answer these questions based on the results provided in the previous chapter. First, the existence of *one* global vegetable oil market is discussed, followed by a description of fish oil's influence and affection related to the other vegetable oils presented in this thesis.

The first section in this chapter evaluates the study and discusses the results. Complimenting discussions on the limitations of the study and potential improvements will also be outlined in the first section. Finally, the final conclusions will be presented in the second section.

6.1 The Existence of One Global Vegetable Oil Market

Based on the price movements among the vegetable oils, there is a common trend which the oils move along. Each oil expresses a unique fluctuations along with the common trend. This is related to differences in market equilibrium price between the oils through time. However, a deviation between certain oils in the short-run is eliminated in the long-run. Based on the multivariate cointegration tests, there is statistical evidence of the vegetable oils being integrated within one market. Meaning that the vegetable oils have followed each other in the long-run over the period 2000 to 2015, but the strict law of one price was rejected for the system. However, the existence of a long-run stationary relationship between the vegetable oils is contradicted and weakened by the bivariate cointegration test results.

Over the full sample period there was evidence of two cointegration vectors, which formed linearly independent stationary relationships between the oils. Hence, the system was strongly driven by the 8 ($N - r = 10 - 2$) common stochastic trends. However, the existence of linear independent combinations of the variables through the full period are considered as proof of the existence of an integrated market, in the long-run. The influence by each oil towards this market is somewhat presented by the cointegration vector's coefficients and their associated p-values. The fish oil price, which differs most compared to an overall average price trend, is considered to have approximately zero influence on the market price. This seems reasonable as the fish oil price showed great difference in price from June 2011 until the end of our sample. Thus, there was no longer a long-run relationship between fish oil and the vegetable oils at the final part of our sample. By the weak exogeneity results we can confirm this effect from fish oil. Fish oil was considered as weakly exogenous with a p-value of 0.60. In addition, the causality tests concluded that fish oil, over the full sample, did not Granger-cause the majority of the vegetable oils.

The other vegetable oils, showed more promising results of being integrated in one market over the full period. This is backed by the greater number of cointegrated pairs, which was tested with a bivariate Johansen cointegration test. Soybean oil was strongly cointegrated with groundnut oil, rapeseed oil and corn oil, and slightly integrated with sunflower and palm oil. In the long-run the law of one price was accepted for soybean oil and rapeseed oil over the full period, which was not the case when considering the commodity boom era. A majority of the oil-pairs rejected the hypothesis of being cointegrated and therefore the LOP. The few numbers of cointegration pairs is most likely a result of the common stochastic trends discovered with the multivariate tests. These common stochastic trends, which are not independent and stationary, are hard to interpret. External forces such as a small crop for a certain year or increased production costs can affect these trends, and cause the linear combinations to not behave as stationary. Among the vegetable oils, rapeseed oil is considered as the price leader over the full sample due to four reasons; (1) second to soybean oil, rapeseed oil formed the most number of cointegration pairs, (2) rapeseed oil Granger-caused groundnut oil, sunflower oil, corn oil, palm oil, palm kernel oil, fish oil and linseed oil, (3) the pairwise LOP was accepted between rapeseed oil and soybean oil, groundnut oil and corn oil, and (4) rapeseed was evaluated as weakly exogenous relative to all the other oil prices.

By evaluating the three sub periods 2003-2008, 2009-2015 and 2011-2015, we may have reduced the sample power of our data. The results for the sub periods showed less promising results of being integrated. A naturally explanation can be that the cointegration tests evaluated the existence of a *long-run* stationary relationship. If the sample is too short, there is hard to establish any long-run relationships. How the relationship between the short-run and long-run trend for cointegrated variables should have been investigated. It is therefore recommended that such a study is conducted when evaluating cointegrated commodities, especially when there is reasoning to believe the variables being integrated while the test results suggest otherwise.

An important note with the multivariate test for the full system of ten oils is the large number of lags which had to be added to remove any serial correlation in the set. Six lags was included in the model to obtain white noise residuals for the multivariate systems. In and Inder (1997) applied three lags in his study of edible oils over the time period 1976 to 1990. Obviously, the vegetable oil prices have developed since the 1980s, and have different trends today. The vegetable oil prices are considered as more volatile in the time period of 2000-2015 than in the 1980s. Thus, more lags were needed in this study to remove serial correlation. The Johansen cointegration test and the ADF-test are sensitive to the number of lags, and this could have produced wrong conclusions about the degree of cointegration (both in the bivariate and multivariate tests). It is likely that we included too many lags for the multivariate tests. By reducing number of lags to four, the null hypothesis of zero cointegration vectors was accepted for the multivariate system. Four lags created a model which violated the diagnostic tests, thus it would be a spurious conclusion to say that there is only non-stationary relationships between the oils in the long-run. The trade-off between including too many lags over the sample power, may have caused our results to suffer in terms of rejection boundaries. Near cointegrated pairs, may have been wrongly concluded as not integrated. Hence, further knowledge of how the correct lags are chosen must be considered and we can propose some questions for further investigation: which of the lag selection criteria are best when working with commodity prices? For example, is Hall's general-to-specific rule better than AIC? Also, how can we choose a lag which satisfies the diagnostic tests as by avoiding any spurious regression?

However, the multivariate tests for each of the three sub periods suggest that the vegetable oils are cointegrated. Most important the test results indicate that the oils can go from being cointegrated with one oil to not be cointegrated as the market develops, and vice verse. The supply and demand of the oils will behave differently over time due to e.g. difference in supply from the producers or that the demand has declined or increased due to external some factor. Thus, by including and/or estimating a supply and demand function for each oil we could have achieved a more complete study of the global oil market. Therefore we recommend such work to be complemented by this thesis. Another benefit by establishing supply and demand curves, is that we can validate our cointegration results with the supply and demand elasticity for each oil.

Another potential source of error is the acceptance of all oils being integrated of order one. The unit roots where checked by the ADF-test, and we did not regard other tests such as the KPSS-test any consideration, which could have suggested otherwise than the ADF-test. Another limitation is how we decided to model the variables with only an unrestricted constant. The idea was that the constant should replicate the differences in transportation and/or quality, which is believed to have been accomplished. However, there is a trend in our data during certain time periods. Hence, the models could better have represented the real data better by including a trend. Especially a deterministic trend in the aftermath of the financial crisis and when the prices decreased from June 2011, to better illustrate the real trend in the price. Reasoning for not undertake such actions is based on the believe that a simpler model will produce better result (obviously, if it somewhat correctly represent the underlying data). Secondly, we divided the full period into three shorter periods which should have corrected for these potential trends. Thus, we believe that one unrestricted constant should be sufficient to represent the real data, and that any trend-effects can be controlled by investigates shorter sub periods of the full sample period.

Fish oil is considered as a premium source of Omega-3 fatty acids, and a vital ingredient in fish feed. The unique qualities of fish oil makes it different from the edible oils. In the period of 2003 to 2009, fish oil was found to be integrated with rapeseed oil. In the long-run, this was proved not to be the case. During this period the fish oil share in salmon feed was reduced from 31.1 % to 16.6% (Ytrestøl et al., 2015). As the fish oil share dropped, the vegetable oil (primarily rapeseed oil) content in fish feed increased from 22.2 % to 35.5 % (Ytrestøl et al., 2015). In the same period, the relative price ratio between fish oil and rapeseed oil was stable and approximately one. Such a constant price ratio between fish oil and any vegetable oil was only observed for rapeseed oil. Hence, the shift of oil shares in aqua feed caused the two oils to become integrated in the same market – the aqua feed market. From June 2011 the two oils were no longer integrated, even though they still were in the same market as inputs in aqua feed. It is believed that the demand of Omega-3 for human consumption significantly contributed to make fish oil price in premium compared to rapeseed, and therefore also affected the degree of market integration between fish oil and other oils. The lack of integration can be illustrated by the two oils' price movements: In June 2011 fish oil and rapeseed oil were traded at 1470 and 1410 USD/tonnes, respectively. The corresponding prices in December 2015 were 1700 and 800 USD/tonnes. Demand of fish oil is still high in the aqua feed industry, as the production of salmon keeps increasing. The reduction of fish oil share in the feed mixture makes the fish oil supply more sustainable than without such a reduction in fish oil share. It is therefore very important that feed producers proceed in this trend – this can assure fish oil supply to be feasible in the future, and the cost of salmon feed will be reduced. The latter point is of great importance for the salmon farmers.

By dividing the ten oils into three groups, where the end-user has different preferences for the oils, quality differences have been proved to affect the degree of market integration. The *general oils* (soybean, sunflower, rapeseed and palm), which are substitutes within cooking oils, margarine and compounded fats, were found to be integrated over the full period, supported by the results from the bivariate cointegration results. The law of one price was accepted for the group of soybean, sunflower, rapeseed and palm oil. Therefore, we must conclude that product-aggregation is possible for the general oils. The same was concluded for palm kernel and coconut oil, which also are considered as substitutes as input in soaps. Hence, there is statistical evidence that the property of each oil determines which oils each single oil is cointegrated with – hence, which market the respective oil is integrated with is determined by its special attributes.

6.2 Concluding Remarks on the Global Vegetable Oil Market

After conducting a thoroughly analysis of the vegetable oil market and price development, we are in a position to answer the questions outlined in Chapter 1. Our first concern was related to the existence of one global vegetable oil market. The following can be concluded based on the empirical findings:

(*i*) The multivariate Johansen cointegration tests have confirmed that it exists a global market for the vegetable oils. Over the full period 2 cointegration vectors were discovered.

(*ii*) The law of one price was rejected for the multivariate system of all ten oil in the period from 2000 to 2015.

(*iii*) The degree of market integration is deemed as a function of the quality and end-user of each oil. If both are similar, there is reasoning to believe that the oils are integrated and behaves as substitutes.

(*iv*) The market prices of soybean, sunflower, rapeseed and palm oil are not likely to diverge in the long-run as they are considered as substitutes. In a system of these four oils, the LOP was accepted.

(*v*) There is statistical evidence of rapeseed oil being a potential price leader among the vegetable oils.

(*vi*) The lauric oils, coconut and palm kernel, are cointegrated and acts as substitutes over the full period.

(*vii*) The pairwise integration between oils change over time. Hence, it exists shorter time periods within an integrated market where some oils are somewhat behaving different than the others.

Secondly, we questioned how fish oil is related to the vegetable oils, and if fish oil was integrated with any of them. Based on the empirical results we conclude that:

(*i*) Fish oil was integrated with rapeseed oil during the commodity boom era between January 2003 and December 2009. This trend continued until June 2011. During this period the fish oil share in salmon feed reduced simultaneously as the rapeseed oil shared increased.

(*ii*) In the final part of our sample, fish oil was no longer integrated with any vegetable oil. Fish oil has become a premium oil compared to the other vegetable oils.

(*iii*) It is believed that since the fish oil usage in salmon feed has remained fairly stable since 2010, it is believed that the demand of fish oil from the aquacultural is an inferior cause to the high prices observed in the market since 2011. Moreover, the increasing demand of the Omega-3 market is believed to have pushed the fish oil prices into a premium compared with the other vegetable oils.

Hence, we concluded that there is *one global vegetable oil market*, which is driven by the similarity of quality and end-users among the oils. Fish oil was a part of this market during the commodity boom period, but an increasing trend of consuming Omega-3 directly has increased the demand of fish oil and pushed the prices upwards. Shepherd and Bachis (2014) states the same conclusions, and we believe that the prices most likely will be evaluated as highly valuable in the future, given a constant supply of fish oil.

Finally, the substitution away from fish oil as the major oil input in salmon feed will keep the fish oil supply sustainable into the future. However, as the fish oil share in salmon is reduced, the Omega-3 content of the salmon will be reduced as well. If the value of salmon is related to its high Omega-3 content it can be argued that salmon will loose its value as a result of substitution away from fish oil as input in the salmon feed. Hence, there is a trade-off for the salmon producers in the future between (1) having high feed costs due to larger shares of fish oil and potentially higher value of salmon as a primary source of Omega-3 and (2) lower feed costs due to lower fish oil shares in the feed and potentially having a lower market value of salmon. Which of the two will yield the best pay-off? This question is unfortunately beyond the scope of this thesis to answer.

Bibliography

Alexander, C. (2008). *The Wiley Finance Series: Market Risk Analysis, Practical Financial Econometrics*. Wiley, 1st edition.

Ardeni, P. G. (1989). Does the law of one price really holds for commodity prices? *American Journal of Agricultural Economics*, 71(3):661–669.

Asche, F., Bremnes, H., and Wessells, C. R. (1999). Product aggregation, market integration and relationships between prices: An application to world salmon markets. *American Journal of Agricultural Economics*, 81:568–581.

Asche, F., Øglend, A., and Tveteras, S. (2013). Regime shifts in the fish meal/soybean meal price ratio. *Journal of Agricultural Economics*, 64(1):91–111.

Asche, F., Gordon, D. V., and Hannesson, R. (2003). Test for Market Integration and the Law of One Price, The Market for Whitefish in France. *Center for Fisheries Economics*, 82.

Asche, F., Osmundsen, P., and Tveterås, R. (2000). European market integration for gas, volume flexibility and political risk. *SNF-Report*, 30.

Bimbo, A. P. (2013). Sources of omega-3 fatty acids. In Jacobsen, C., Nielsen, N. S., Horn, A. F., and Sørensen, A.-D. M., editors, *Food Enrichment with Omega-3 Fatty Acids*, chapter 2, pages 27–107. Woodhead Publishing.

Blough, S. R. (1992). The relationship between power and level for generic unit root tests in finite samples. *Journal of Applied Econometrics*, 7:295–308.

Breusch, T. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17:334–355.

Cournot, A. A. (1971). *Researches into the Mathematical Principles of the Theory of Wealth.* New York, A. M. Kelly.

Deaton, A. S. and Muellbauer, J. (1980). Economics and consumer behaviour. *Cambridge University Press, New York.*

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autregressive time series with a unit root. *Econometrica*, 49(4):1057–1072.

Durbin, J. and Watson, G. S. (1950). Testing for serial correlation in least squares regression i. *Biometrika*, 37(3-4):409–428.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and, testing. *Econometrica*, 55(2):251–276.

Fabozzi, F. J., Focardi, S. M., and Rachev, S. T. (2014). *Basics of Financial Econometrics, Tools, Concept, and Asset Management Applications.* Wiley, 1st edition.

FAO (2016). Food and Agricultural Organization of the United Nations: Compare data. `http://faostat3.fao.org/compare/E`. [Online; accessed: 2016-09-10].

Goddard, E. W. and Glance, S. (1989). Demand for fats and oils in canada, united states and japan. *Canadian Journal of Agricultural Economics*, 37(4):488–498.

Godfrey, L. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica*, 46:1303–1310.

Godwin, B. K. and Schroeder, T. C. (1991). Co-integration tests and spatial prime linkages in regional cattle markets. *American Journal of Agricultural Economics*, 73(2):452–464.

Gordon, D. V. and Hannesson (1996). On prices of fresh and frozen cod fish in european and u.s. markets. *Marine Resource Economics*, 11(4):223–238.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Econometrica*, 16:150–161.

Granger, C. W. J. (1983). Cointegrated variables and error correction models. *UCSD Discussion Paper*, 83(13).

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrcs. *Journal of Econometrics*, 2:111–120.

Hall, A. (1992). Testing for a unit root in time series using instrumental variable estimators with pretest data based model selection. *Journal of Econometrics*, 54:223–250.

Hicks, J. R. (1936). Value and capital. *London: Oxford University Press.*

In, F. and Inder, B. (1997). Long-run relationships between world vegetable oil prices. *The Australian Journal of Agricultural and Resource Economics*, 41(4):455–470.

ISTA Mielke GmbH (2016). Commodities. `https://www.oilworld.biz/app.php?fid=300&fpar=YToxOntzOjQ6InBjaWQiO3M6MjoiMTIiO30%3D&isSSL=0&aps=0&blub=bf45fac70c6d77a1db7966c4154bb249`. [Online; accessed: 2016-09-01].

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2):231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580.

Johansen, S. and Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration – with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, 52(2):169–210.

Laksefakta (2016). Hva er i foret til laksen? `https://laksefakta.no/hva-spiser-laksen/hva-er-i-foret-til-laksen/`. [Online; accessed: 2016-10-13].

Leontief, W. (1936). Composite commodities and the problem of index numbers. *Econometrica*, 4:39–59.

Lewbel, A. (1996). Aggregation without separability: A generalized composite commodity theorem. *American Economic Relationships*, 86(3):524–561.

Liddle, B. (2009). Long-run relationship among transport demand, income and gasoline price for the us. *Transportation Research Part D*, 14:73–82.

Lloyd, T. (2008). Price leadership in uk food retailing: Time series representation and evidence. In *82nd Annual Conference, March 31 - April 2, 2008, Royal Agricultural College, Cirencester, UK*.

Marine Harvest ASA (2015). *Salmon Farming Handbook 2015*. Marine Harvest ASA.

Marine Harvest ASA (2016). *Salmon Farming Handbook 2016*. Marine Harvest ASA.

Marshall, A. (1947). Principles of economics. *Macmillian*.

Naylor, R. L., Hardy, R. W., Bureau, D., Chiu, A., Elliott, M., Farrell, A. P., Gatlin, D. M., Goldburg, R. J., Hua, K., and Nichols, P. D. (2009). Feeding aquaculture in an era of finite resources. *Proceedings of the National Academy of Sciences*, 106:15103–15110.

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69(6):1519–1554.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6):1361–1401.

Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. Springer, 2nd edition.

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrica*, 75(2):335–346.

RStudio Team (2016). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Shepherd, J. and Bachis, E. (2014). Changing supply and demand for fish oil. *Aquaculture Economics & Management*, 18(4):395–416.

Slade, M. E. (1986). Exogenity test of market boundaries applied to petroleum products. *Journal of Industrial Economics*, 34:291–304.

Statistisk Sentralbyrå (2015). Aquaculture 2015. `https://www.ssb.no/en/jord-skog-jakt-og-fiskeri/statistikker/fiskeoppdrett/aar-forelopige/2016-06-02?fane=tabell#content`. [Online; accessed: 2016-10-1].

Stigler, G. J. (1969). The theory of price. *Macmillian*.

Stigler, G. J. and Sherwin, R. A. (1985). The extent of a market. *Journal of Law and Economics*, 28:555–585.

Tacon, A. G. and Metian, M. (2008). Global overview on the use of fish meal and fish oil in industrially compounded aquafeeds: Trends and future prospects. *Aquaculture*, 285:146–158.

Toda, H. Y. and Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66:225–250.

Tomek, W. G. and Kaiser, H. M. (2014). *Agricultural Product Prices*. Cornell University Press, 5th edition.

Tveterås, S. (2000). The fishmeal market: An integrated part of the vegetable oilmeal market? *Center for Fisheries Economics*, 2/2000.

USDA (2016). Oilseeds: World markets and trade. *Oilseeds*, 11/2016.

Wikipedia (2016). Vegetable Oil. `https://en.wikipedia.org/wiki/Vegetable_oil`. [Online; accessed: 2016-08-20].

Williams, W. G. and Thompson, R. L. (1984). Brazilian soybean policy: the international effects of intervention. *American Journal of Agricultural Economics*, 66(3):488–498.

Woolridge, J. M. (2009). *Introductory Econometrics, A Modern Approach*. South-Western, 4th edition.

World Bank (2008). *Global Economic Prospects 2009*. The World Bank.

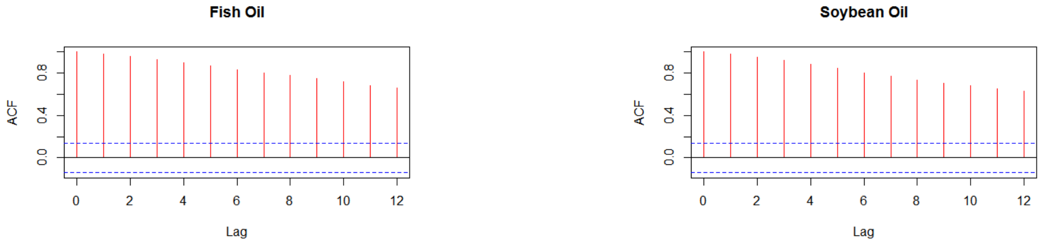
Ytrestøl, T., Aas, T. S., and Åsgård, T. (2015). Utilisation of feed resources in production of atlantic salmon (salmo salar) in norway. *Aquaculture*, 448:365–374.

Yule, G. U. (1926). Why do we sometimes get nonsense-correlation between time series? *Journal of the Royal Statstical Society*, 89(1).

Zivot, E. and Andrews, D. W. K. (1992). Further evidence on the great crash, the oil price shock, and the unit root hypothesis. *Journal of Business and Economic Statistics*, 10(3):251–270.

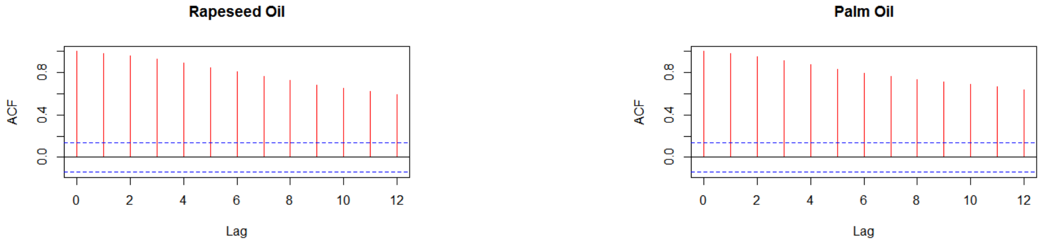
Appendix A

Additional Figures



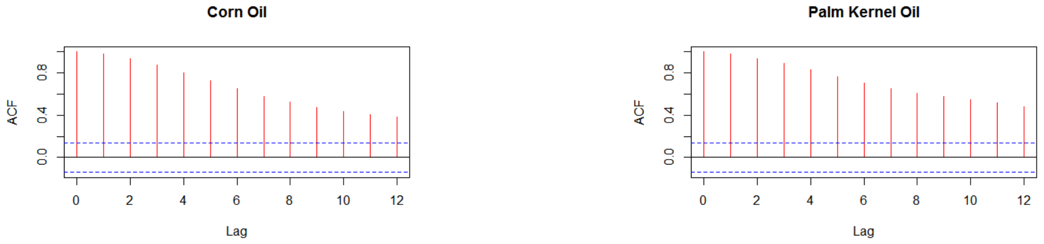
(a) Correlogram for fish oil.

(b) Correlogram for soybean oil



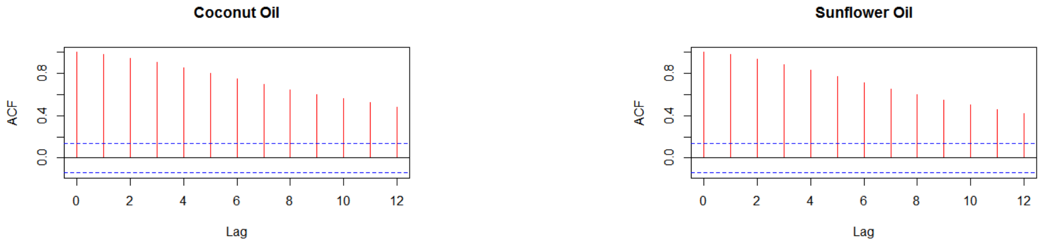
(c) Correlogram for rapeseed oil.

(d) Correlogram for palm oil



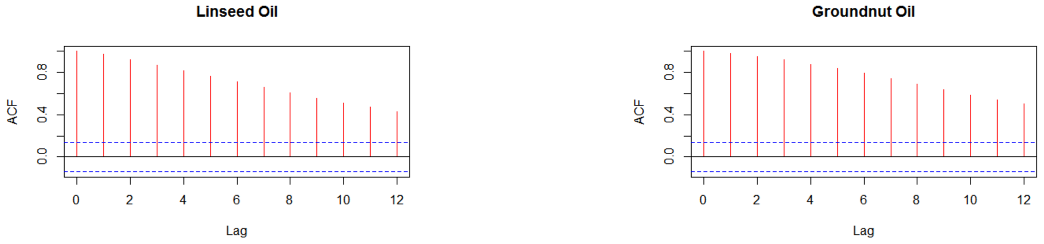
(e) Correlogram for corn oil.

(f) Correlogram for palm kernel oil



(g) Correlogram for coconut oil.

(h) Correlogram for sunflower oil



(i) Correlogram for linseed oil.

(j) Correlogram for groundnut oil

Figure A.1: Correlograms for fish oil and the vegetable oils and their associated price series, with 12 lags.

(a) histogram for fish oil.

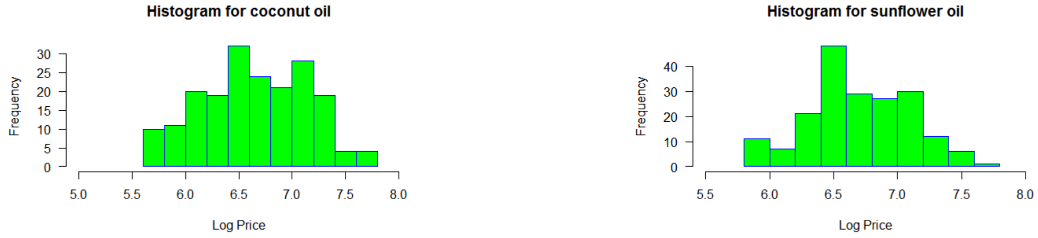
(b) Histogram for soybean oil

(c) Histogram for rapeseed oil.

(d) Histogram for palm oil

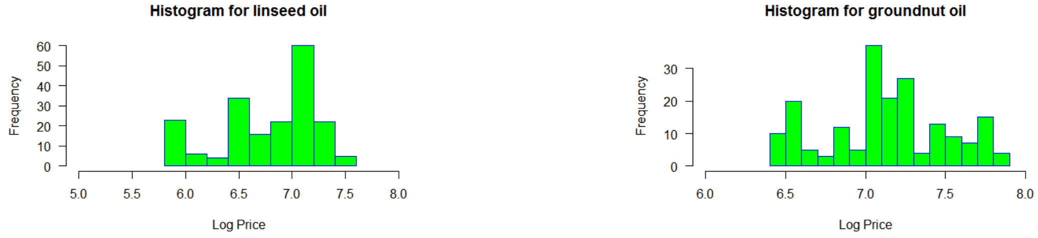
(e) Histogram for corn oil.

(f) Histogram for palm kernel oil



(g) Histogram for coconut oil.

(h) Histogram for sunflower oil



(i) Histogram for linseed oil.

(j) Histogram for groundnut oil

Figure A.2: Histograms of the ten oil price series (logarithmic values).

(a) Fish oil price.

(b) First difference of fish oil price.

(c) Soybean oil price.

(d) First difference of soybean oil price.

(e) Groundnut oil price.

(f) First difference of groundnut oil price.

(g) Rapeseed oil price.

(h) First difference of rapeseed oil price.

(i) Sunflower oil price.

(j) First difference of sunflower oil price.

Figure A.3: The price and first differences of the price for five oils, in logarithmic values.

(a) Corn oil price.

(b) First difference of corn oil price.

(c) Palm oil price.

(d) First difference of palm oil price.

(e) Palm kernel oil price.

(f) First difference of palm kernel oil price.

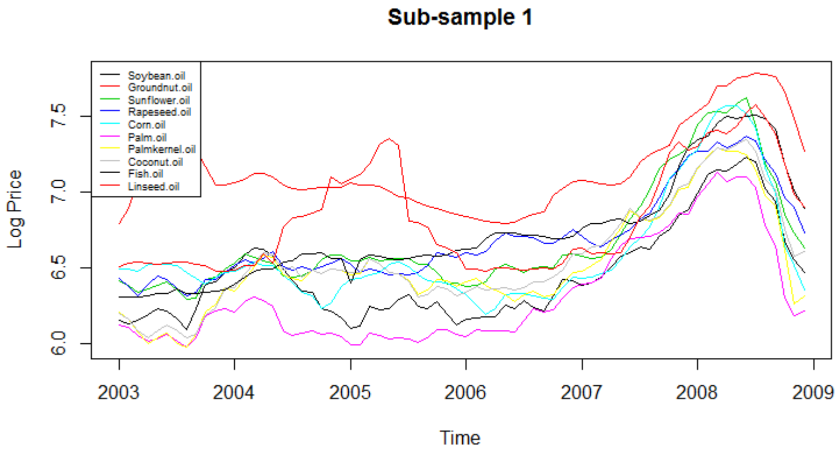
(g) Coconut oil price.

(h) First difference of coconut oil price.

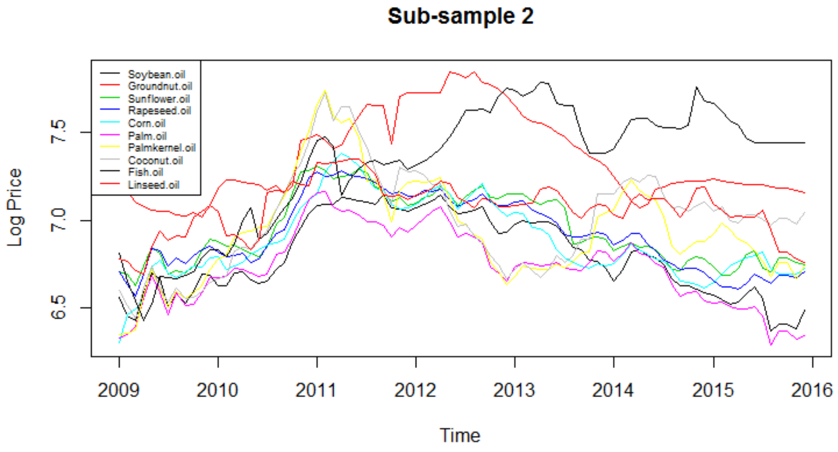
(i) Linseed oil price.

(j) First difference of linseed oil price.

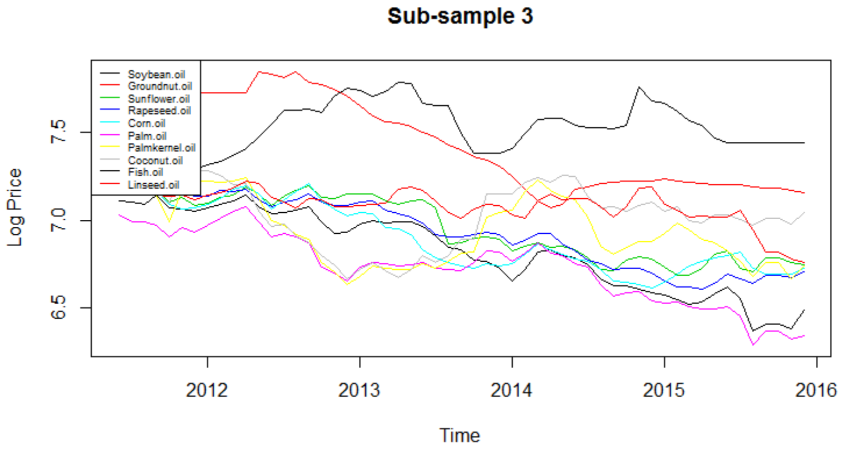
Figure A.4: The price and first differences of the price for five oils, in logarithmic values.



(a) Price development in the period January 2003 to December 2008.



(b) Price development in the period January 2009 to December 20015.



(c) Price development in the period June 2011 to December 2015.

Figure A.5: Price development for Sub-samples 1, 2 and 3.

Appendix B

Additional Tables

Table B.1: Mean (standard deviation) of the soybean oil, groundnut oil and sunflower oil, for the respective time periods given in Table 5.1.

Time period	**Soybean oil**	**Groundnut oil**	**Sunflower oil**
Full sample	717.30 (293.63)	1330.42 (503.76)	877.10 (348.31)
Sub-sample 0	345.33 (61.38)	693.67 (46.49)	490.11 (98.84)
Sub-sample 1	670.53 (245.56)	1315.33 (414.12)	855.63 (388.45)
Sub-sample 2	916.81 (207.48)	1616.25 (421.46)	1061.38 (215.07))
Sub-sample 3	939.11 (208.46)	1741.78 (446.01)	1062.85 (192.04)

Table B.2: Mean (standard deviation) of the rapeseed oil, corn oil, palm oil and palm kernel oil for the respective time periods given in Table 5.1.

Time period	**Rapeseed oil**	**Corn oil**	**Palm oil**	**Palm kernel oil**
Full sample	845.52 (319.25)	814.15 (344.25)	658.88 (277.63)	817.70 (390.62)
Sub-sample 0	411.33 (75.84)	435.72 (85.07)	328.72 (61.15)	389.14 (91.70)
Sub-sample 1	841.05 (281.42)	787.36 (371.48)	590.72 (240.67)	722.31 (272.41)
Sub-sample 2	1035.44 (214.65)	999.31 (232.75)	858.81 (185.70)	1083.16 (353.89)
Sub-sample 3	1036.70 (202.16)	1002.74 (196.56)	850.76 (166.13)	1043.17 (201.88)

Table B.3: Mean (standard deviation) of the coconut oil, fish oil and linseed oil, for the respective time periods given in Table 5.1.

Time period	**Coconut oil**	**Fish oil**	**Linseed oil**
Full sample	856.91 (409.07)	1108.72 (597.22)	967.17 (376.13)
Sub-sample 0	396.47 (89.15)	433.25 (150.27)	433.64 (88.70)
Sub-sample 1	749.29 (287.17)	890.46 (364.05)	978.72 (374.57)
Sub-sample 2	1146.49 (352.59)	1585.3 (478.32)	1185.92 (173.81)
Sub-sample 3	1146.46 (209.80)	1869.65 (264.87)	1197.87 (120.90)

Table B.4: Mean relative prices (standard deviation) for the full sample. Calculated by the **column header**/*row name.*

	Soybean oil	**Rapeseed oil**	**Palm oil**	**Fish oil**
Soybean oil	1.00 (0.00)	1.20 (0.15)	0.92 (0.10)	1.52 (0.52)
Groundnut oil	0.54 (0.09)	0.64 (0.11)	0.50 (0.11)	0.81 (0.30)
Sunflower oil	0.81 (0.11)	0.97 (0.11)	0.75 (0.11)	1.23 (0.45)
Rapeseed oil	0.84 (0.10)	1.00 (0.00)	0.78 (0.11)	1.27 (0.47)
Corn oil	0.88 (0.12)	1.06 (0.16)	0.81 (0.12)	1.34 (0.50)
Palm oil	1.10 (0.12)	1.32 (0.20)	1.00 (0.00)	1.66 (0.57)
Palm kernel oil	0.91 (0.18)	1.09 (0.23)	0.83 (0.13)	1.37 (0.50)
Coconut oil	0.88 (0.18)	1.05 (0.23)	0.80 (0.14)	1.30 (0.45)
Fish oil	0.73 (0.24)	0.87 (0.25)	0.67 (0.22)	1.00 (0.00)
Linseed oil	0.76 (0.14)	0.90 (0.15)	0.69 (0.14)	1.12 (0.37)

Table B.5: Mean relative prices (standard deviation) for Sub-sample 1, January 2003 to December 2008. Calculated by the **column header**/*row name.*

	Soybean oil	**Rapeseed oil**	**Palm oil**	**Fish oil**
Soybean oil	1.00 (0.00)	1.28 (0.18)	0.88 (0.11)	1.33 (0.22)
Groundnut oil	0.51 (0.08)	0.65 (0.12)	0.45 (0.10)	0.68 (0.14)
Sunflower oil	0.81 (0.11)	1.02 (0.13)	0.70 (0.09)	1.06 (0.16)
Rapeseed oil	0.80 (0.11)	1.00 (0.00)	0.70 (0.10)	1.04 (0.10)
Corn oil	0.88 (0.13)	1.12 (0.21)	0.77 (0.12)	1.17 (0.25)
Palm oil	1.16 (0.14)	1.47 (0.21)	1.00 (0.00)	1.53 (0.28)
Palm kernel oil	0.94 (0.15)	1.19 (0.18)	0.82 (0.11)	1.24 (0.22)
Coconut oil	0.91 (0.13)	1.15 (0.15)	0.79 (0.10)	1.19 (0.16)
Fish oil	0.78 (0.15)	0.97 (0.09)	0.67 (0.12)	1.00 (0.00)
Linseed oil	0.72 (0.17)	0.90 (0.20)	0.63 (0.17)	0.93 (0.21)

Table B.6: Mean relative prices (standard deviation) for Sub-sample 2, January 2009 to December 2015. Calculated by the **column header**/*row name*.

	Soybean oil	**Rapeseed oil**	**Palm oil**	**Fish oil**
Soybean oil	1.00 (0.00)	1.14 (0.08)	0.94 (0.09)	1.78 (0.61)
Groundnut oil	0.58 (0.08)	0.66 (0.10)	0.55 (0.11)	1.00 (0.31)
Sunflower oil	0.86 (0.07)	0.98 (0.05)	0.81 (0.09)	1.53 (0.51)
Rapeseed oil	0.88 (0.06)	1.00 (0.00)	0.83 (0.07)	1.57 (0.55)
Corn oil	0.92 (0.10)	1.05 (0.10)	0.87 (0.10)	1.64 (0.56)
Palm oil	1.07 (0.10)	1.21 (0.11)	1.00 (0.00)	1.91 (0.70)
Palm kernel oil	0.89 (0.21)	1.00 (0.23)	0.82 (0.14)	1.57 (0.62)
Coconut oil	0.84 (0.23)	0.95 (0.25)	0.78 (0.17)	1.46 (0.57)
Fish oil	0.63 (0.21)	0.72 (0.25)	0.60 (0.22)	1.00 (0.00)
Linseed oil	0.77 (0.11)	0.87 (0.11)	0.72 (0.10)	1.34 (0.39)

Table B.7: Bivariate Johansen tests for the Sub-sample 1. The LOP is rejected for p-value<0.05, and "-" indicates not valid test results.

		Max Test		**Trace Test**		**LOP**
Variable 1	**Variable 2**	r==0	r<=1	r == 0	r<=1	p-value
Soybean oil	Groundnut oil	41.40**	9.40	50.80**	9.40	0.12
Soybean oil	Sunflower oil	15.46*	6.86	22.33*	6.86	0.71
Soybean oil	Rapeseed oil	4.79	2.57	7.35	2.57	0.43
Soybean oil	Corn oil	16.10*	2.95	19.06	2.95	0.30
Soybean oil	Palm oil	15.70*	5.88	21.55*	5.88	0.09
Soybean oil	Palm kernel oil	13.73	6.42	20.15*	6.42	0.20
Soybean oil	Coconut oil	4.41	3.21	7.62	3.21	0.32
Soybean oil	Fish oil	5.64	3.80	9.43	3.80	0.36
Soybean oil	Linseed oil	5.63	4.26	9.90	4.26	0.97
Groundnut oil	Sunflower oil	24.54**	11.40*	35.94**	11.40*	0.00
Groundnut oil	Rapeseed oil	8.00	6.15	14.15	6.15	0.19
Groundnut oil	Corn oil	11.98	7.46	19.44	7.46	0.78
Groundnut oil	Palm oil	9.97	5.41	15.38	5.41	-
Groundnut oil	Palm kernel oil	9.59	5.90	15.49	5.90	-
Groundnut oil	Coconut oil	10.15	6.78	16.93	6.78	-
Groundnut oil	Fish oil	7.87	3.09	10.96	3.09	-
Groundnut oil	Linseed oil	6.21	3.99	10.19	3.99	-
Sunflower oil	Rapeseed oil	6.15	3.06	9.21	3.06	-
Sunflower oil	Corn oil	14.35	4.71	19.06	4.71	0.01
Sunflower oil	Palm oil	12.63	4.28	16.91	4.28	-
Sunflower oil	Palm kernel oil	11.47	4.91	16.38	4.91	-
Sunflower oil	Coconut oil	8.73	4.09	12.82	4.09	-
Sunflower oil	Fish oil	13.65	4.22	17.87	4.22	0.91
Sunflower oil	Linseed oil	5.73	4.48	10.21	4.48	-
Rapeseed oil	Corn oil	11.08	5.59	16.67	5.59	0.18
Rapeseed oil	Palm oil	8.70	2.77	11.47	2.77	-
Rapeseed oil	Palm kernel oil	12.41	6.87	19.28	6.87	0.06
Rapeseed oil	Coconut oil	12.20	10.20	22.40	10.20	0.89
Rapeseed oil	Fish oil	17.37*	4.09	21.46*	4.09	0.01
Rapeseed oil	Linseed oil	11.71	5.22	16.93	5.22	0.01
Corn oil	Palm oil	12.87	3.28	16.15	3.28	-
Corn oil	Palm kernel oil	20.54**	11.35*	31.90**	11.35*	-
Corn oil	Coconut oil	17.38*	11.36*	28.74**	11.36*	-
Corn oil	Fish oil	16.90*	4.87	21.77*	4.87	0.00
Corn oil	Linseed oil	10.50	4.71	15.21	4.71	0.03
Palm oil	Palm kernel oil	4.62	1.56	6.17	1.56	-
Palm oil	Coconut oil	9.53	3.17	12.71	3.17	-
Palm oil	Fish oil	10.97	3.31	14.28	3.31	-
Palm oil	Linseed oil	6.29	3.98	10.27	3.98	-
Palm kernel oil	Coconut oil	10.97	3.19	14.16	3.19	-
Palm kernel oil	Fish oil	11.14	5.63	16.77	5.63	0.21
Palm kernel oil	Linseed oil	9.59	5.68	15.26	5.68	0.93
Coconut oil	Fish oil	6.76	4.92	11.69	4.92	-
Coconut oil	Linseed oil	8.90	4.36	13.26	4.36	-
Fish oil	Linseed oil	4.71	2.88	7.60	2.88	-

Critical values for the Max-test is 15.7 for H_0: r==0 and 9.2 for H_0: r<=1. Critical values for the Trace-test is 20.0 for H_0: r==0 and 9.2 for H_0: r<=1. ** indicates 1 % significance level, * indicates 5 % siginificance level.

Table B.8: Bivariate Johansen tests for the Sub-sample 2. The LOP is rejected for p-value<0.05, and "-" indicates not valid test results.

		Max Test		**Trace Test**		**LOP**
Variable 1	**Variable 2**	r==0	r<=1	r == 0	r<=1	p-value
Soybean oil	Groundnut oil	15.75*	1.71	17.46	1.71	0.24
Soybean oil	Sunflower oil	13.39	2.32	15.70	2.32	0.15
Soybean oil	Rapeseed oil	23.71**	4.03	27.74**	4.03	0.12
Soybean oil	Corn oil	8.76	6.80	15.55	6.80	-
Soybean oil	Palm oil	13.41	1.05	14.46	1.05	0.18
Soybean oil	Palm kernel oil	8.62	0.77	9.38	0.77	-
Soybean oil	Coconut oil	6.02	1.90	7.92	1.90	-
Soybean oil	Fish oil	11.44	2.44	13.87	2.44	-
Soybean oil	Linseed oil	10.84	2.46	13.30	2.46	-
Groundnut oil	Sunflower oil	12.37	1.50	13.88	1.50	-
Groundnut oil	Rapeseed oil	10.71	1.92	12.64	1.92	-
Groundnut oil	Corn oil	17.40*	3.16	20.56*	3.16	0.18
Groundnut oil	Palm oil	11.59	2.49	14.07	2.49	0.17
Groundnut oil	Palm kernel oil	13.21	1.79	15.00	1.79	0.38
Groundnut oil	Coconut oil	8.57	5.46	14.03	5.46	-
Groundnut oil	Fish oil	7.70	1.26	8.96	1.26	-
Groundnut oil	Linseed oil	15.24	1.11	16.35	1.11	0.00
Sunflower oil	Rapeseed oil	13.26	2.70	15.96	2.70	0.46
Sunflower oil	Corn oil	14.12	4.27	18.39	4.27	0.53
Sunflower oil	Palm oil	4.90	2.58	7.48	2.58	-
Sunflower oil	Palm kernel oil	7.84	1.74	9.58	1.74	-
Sunflower oil	Coconut oil	6.71	1.08	7.80	1.08	-
Sunflower oil	Fish oil	6.78	2.49	9.27	2.49	-
Sunflower oil	Linseed oil	8.55	2.71	11.25	2.71	-
Rapeseed oil	Corn oil	8.82	3.87	12.69	3.87	-
Rapeseed oil	Palm oil	12.92	2.76	15.68	2.76	0.44
Rapeseed oil	Palm kernel oil	6.38	1.09	7.47	1.09	-
Rapeseed oil	Coconut oil	6.64	1.01	7.66	1.01	0.04
Rapeseed oil	Fish oil	14.52	2.70	17.23	2.70	0.04
Rapeseed oil	Linseed oil	7.79	2.38	10.17	2.38	-
Corn oil	Palm oil	5.18	2.04	7.22	2.04	-
Corn oil	Palm kernel oil	9.30	5.17	14.46	5.17	0.04
Corn oil	Coconut oil	6.74	1.41	8.15	1.41	-
Corn oil	Fish oil	9.91	4.51	14.42	4.51	0.09
Corn oil	Linseed oil	7.70	3.90	11.60	3.90	-
Palm oil	Palm kernel oil	8.04	0.67	8.71	0.67	-
Palm oil	Coconut oil	6.56	2.22	8.77	2.22	-
Palm oil	Fish oil	10.71	2.42	13.12	2.42	-
Palm oil	Linseed oil	12.27	2.38	14.65	2.38	-
Palm kernel oil	Coconut oil	11.04	5.59	16.63	5.59	-
Palm kernel oil	Fish oil	14.68	4.05	18.74	4.05	0.01
Palm kernel oil	Linseed oil	13.83	5.14	18.96	5.14	0.01
Coconut oil	Fish oil	12.43	5.98	18.41	5.98	0.11
Coconut oil	Linseed oil	6.46	4.12	10.57	4.12	-
Fish oil	Linseed oil	6.47	4.57	11.04	4.57	-

Critical values for the Max-test is 15.7 for H_0: r==0 and 9.2 for H_0: r<=1. Critical values for the Trace-test is 20.0 for H_0: r==0 and 9.2 for H_0: r<=1. ** indicates 1 % significance level, * indicates 5 % significance level.

Table B.9: Bivariate Johansen tests for the Sub-sample 3. The LOP is rejected for p-value<0.05, and "-" indicates not valid test results.

		Max Test		**Trace Test**		**LOP**
Variable 1	**Variable 2**	r==0	r<=1	r == 0	r<=1	p-value
Soybean oil	Groundnut oil	27.07**	3.33	30.40**	3.33	0.03
Soybean oil	Sunflower oil	12.39	5.25	17.64	5.25	0.04
Soybean oil	Rapeseed oil	15.42*	5.38	20.08*	5.38	0.03
Soybean oil	Corn oil	13.42	2.57	15.99	2.57	0.45
Soybean oil	Palm oil	14.64	3.39	18.02	3.39	0.90
Soybean oil	Palm kernel oil	15.82*	3.87	19.96	3.87	0.34
Soybean oil	Coconut oil	6.90	2.83	9.73	2.83	-
Soybean oil	Fish oil	15.07	6.96	22.02*	6.96	0.15
Soybean oil	Linseed oil	15.89*	1.99	17.88	1.99	0.02
Groundnut oil	Sunflower oil	22.07**	2.05	24.12**	2.05	0.00
Groundnut oil	Rapeseed oil	13.19	2.65	15.85	2.65	0.09
Groundnut oil	Corn oil	8.39	3.67	12.06	3.67	-
Groundnut oil	Palm oil	5.83	2.33	8.15	2.33	-
Groundnut oil	Palm kernel oil	7.74	3.39	11.23	3.49	-
Groundnut oil	Coconut oil	12.37	5.41	17.77	5.41	0.01
Groundnut oil	Fish oil	9.46	2.56	12.01	2.56	-
Groundnut oil	Linseed oil	4.06	3.37	7.43	3.37	-
Sunflower oil	Rapeseed oil	7.31	5.54	12.84	5.54	-
Sunflower oil	Corn oil	14.02	3.41	17.43	3.41	0.01
Sunflower oil	Palm oil	11.26	3.29	14.55	3.29	0.01
Sunflower oil	Palm kernel oil	10.50	5.58	16.08	5.58	0.01
Sunflower oil	Coconut oil	11.63	4.13	15.76	4.13	0.04
Sunflower oil	Fish oil	6.36	4.39	10.75	4.39	-
Sunflower oil	Linseed oil	8.49	1.21	9.70	1.21	0.10
Rapeseed oil	Corn oil	8.40	3.11	11.51	3.11	-
Rapeseed oil	Palm oil	5.59	3.45	9.04	3.45	-
Rapeseed oil	Palm kernel oil	7.19	2.03	9.22	2.03	-
Rapeseed oil	Coconut oil	8.94	3.03	11.98	3.03	-
Rapeseed oil	Fish oil	10.30	5.96	16.26	5.96	0.06
Rapeseed oil	Linseed oil	16.96*	1.73	18.96	1.73	0.00
Corn oil	Palm oil	5.53	2.21	7.75	2.21	-
Corn oil	Palm kernel oil	6.62	3.58	10.21	3.58	-
Corn oil	Coconut oil	11.41	4.78	16.19	4.78	0.01
Corn oil	Fish oil	11.36	5.68	17.04	5.68	0.02
Corn oil	Linseed oil	5.02	1.41	6.43	1.41	-
Palm oil	Palm kernel oil	8.17	3.56	11.73	3.56	-
Palm oil	Coconut oil	7.10	3.33	10.43	3.33	-
Palm oil	Fish oil	12.73	4.78	17.51	4.78	0.21
Palm oil	Linseed oil	5.90	3.98	9.88	3.98	-
Palm kernel oil	Coconut oil	6.93	1.12	8.84	1.12	-
Palm kernel oil	Fish oil	7.48	4.29	11.77	4.29	-
Palm kernel oil	Linseed oil	4.55	2.51	7.06	2.51	-
Coconut oil	Fish oil	10.81	5.11	15.92	5.11	0.02
Coconut oil	Linseed oil	8.16	1.32	9.48	1.32	-
Fish oil	Linseed oil	12.01	2.05	14.06	2.05	0.17
Fish oil	Linseed oil	5.81	5.50	11.30	5.50	0.58

Critical values for the Max-test is 15.7 for H_0: r==0 and 9.2 for H_0: r<=1. Critical values for the Trace-test is 20.0 for H_0: r==0 and 9.2 for H_0: r<=1. ** indicates 1 % significance level, * indicates 5 % significance level.

Table B.10: Weak exogeneity results for all periods. A p-value less than 0.05 indicates that it is not exogenous.

		Full period		**Sub-sample 1**		**Sub-sample 2**		**Sub-sample 3**	
Variable 1	**Variable 2**	$\alpha_1 = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\alpha_2 = 0$	$\alpha_1 = 0$	$\alpha_2 = 0$
Soybean oil	Groundnut oil	0.705	0.000	0.590	0.002	0.301	0.001	0.005	0.000
Soybean oil	Sunflower oil	0.054	0.418	0.146	0.905	0.025	0.854	0.353	0.154
Soybean oil	Rapeseed oil	0.041	0.862	0.134	0.432	0.000	0.024	0.000	0.099
Soybean oil	Corn oil	0.383	0.000	0.944	0.022	0.618	0.422	0.061	0.462
Soybean oil	Palm oil	0.002	0.473	0.001	0.006	0.031	0.964	0.017	0.013
Soybean oil	Palm kernel oil	0.592	0.053	0.016	0.114	0.246	0.099	0.003	0.269
Soybean oil	Coconut oil	0.291	0.002	0.072	0.559	0.846	0.046	0.422	0.044
Soybean oil	Fish oil	0.674	0.007	0.027	0.932	0.043	0.069	0.490	0.109
Soybean oil	Linseed oil	0.407	0.016	0.196	0.297	0.761	0.006	0.001	0.017
Groundnut oil	Sunflower oil	0.020	0.226	0.006	0.810	0.003	0.434	0.861	0.000
Groundnut oil	Rapeseed oil	0.008	0.872	0.277	0.302	0.005	0.389	0.740	0.018
Groundnut oil	Corn oil	0.000	0.049	0.294	0.083	0.000	0.466	0.133	0.089
Groundnut oil	Palm oil	0.000	0.436	0.000	0.483	0.004	0.614	0.003	0.995
Groundnut oil	Palm kernel oil	0.002	0.041	0.034	0.777	0.003	0.302	0.002	0.177
Groundnut oil	Coconut oil	0.028	0.020	0.077	0.981	0.253	0.206	0.024	0.176
Groundnut oil	Fish oil	0.616	0.006	0.038	0.327	0.444	0.018	0.110	0.118
Groundnut oil	Linseed oil	0.064	0.242	0.155	0.907	0.001	0.075	0.268	0.193
Sunflower oil	Rapeseed oil	0.021	0.882	0.102	0.255	0.148	0.369	0.238	0.743
Sunflower oil	Corn oil	0.252	0.000	0.117	0.002	0.036	0.283	0.001	0.107
Sunflower oil	Palm oil	0.024	0.674	0.019	0.725	0.376	0.734	0.062	0.002
Sunflower oil	Palm kernel oil	0.740	0.154	0.232	0.160	0.424	0.074	0.096	0.126
Sunflower oil	Coconut oil	0.717	0.050	0.218	0.403	0.662	0.020	0.040	0.037
Sunflower oil	Fish oil	0.020	0.001	0.254	0.013	0.123	0.139	0.148	0.102
Sunflower oil	Linseed oil	0.817	0.237	0.639	0.305	0.928	0.027	0.092	0.492
Rapeseed oil	Corn oil	0.281	0.001	0.299	0.020	0.863	0.091	0.028	0.375
Rapeseed oil	Palm oil	0.311	0.458	0.019	0.289	0.411	0.093	0.000	0.000
Rapeseed oil	Palm kernel oil	0.609	0.027	0.222	0.021	0.611	0.099	0.141	0.000
Rapeseed oil	Coconut oil	0.679	0.008	0.741	0.291	0.205	0.018	0.175	0.367
Rapeseed oil	Fish oil	0.189	0.005	0.749	0.001	0.028	0.136	0.211	0.346
Rapeseed oil	Linseed oil	0.737	0.022	0.018	0.267	0.184	0.021	0.000	0.037
Corn oil	Palm oil	0.000	0.047	0.002	0.046	0.078	0.269	0.043	0.008
Corn oil	Palm kernel oil	0.132	0.648	0.005	0.034	0.136	0.280	0.067	0.076
Corn oil	Coconut oil	0.497	0.394	0.016	0.120	0.856	0.122	0.721	0.011
Corn oil	Fish oil	0.137	0.001	0.001	0.999	0.076	0.039	0.194	0.161
Corn oil	Linseed oil	0.079	0.181	0.019	0.942	0.580	0.054	0.000	0.016
Palm oil	Palm kernel oil	0.035	0.002	0.003	0.001	0.580	0.031	0.178	0.684
Palm oil	Coconut oil	0.022	0.001	0.451	0.049	0.132	0.040	0.704	0.069
Palm oil	Fish oil	0.277	0.000	0.831	0.010	0.011	0.118	0.438	0.037
Palm oil	Linseed oil	0.853	0.003	0.490	0.156	0.563	0.002	0.018	0.005
Palm kernel oil	Coconut oil	0.105	0.034	0.007	0.026	0.045	0.090	0.212	0.148
Palm kernel oil	Fish oil	0.709	0.008	0.202	0.139	0.550	0.017	0.003	0.004
Palm kernel oil	Linseed oil	0.050	0.031	0.886	0.049	0.983	0.008	0.091	0.469
Coconut oil	Fish oil	0.376	0.037	0.968	0.198	0.529	0.016	0.466	0.088
Coconut oil	Linseed oil	0.032	0.133	0.160	0.065	0.187	0.257	0.003	0.517
Fish oil	Linseed oil	0.021	0.769	0.478	0.210	0.177	0.535	0.022	0.149

Table B.11: Pairwise causality test over the full period, January 2000 to December 2015.

Variable 1 y_1	**Variable 2** y_2	y_2 **does not cause** y_1	y_1 **does not cause** y_2
Soybean oil	Groundnut oil	0.828	0.000
Soybean oil	Sunflower oil	0.016	0.989
Soybean oil	Rapeseed oil	0.129	0.066
Soybean oil	Corn oil	0.000	0.005
Soybean oil	Palm oil	0.041	0.334
Soybean oil	Palm kernel oil	0.001	0.830
Soybean oil	Coconut oil	0.017	0.853
Soybean oil	Fish oil	0.575	0.178
Soybean oil	Linseed oil	0.769	0.010
Groundnut oil	Sunflower oil	0.000	0.067
Groundnut oil	Rapeseed oil	0.000	0.358
Groundnut oil	Corn oil	0.000	0.609
Groundnut oil	Palm oil	0.001	0.074
Groundnut oil	Palm kernel oil	0.014	0.250
Groundnut oil	Coconut oil	0.006	0.131
Groundnut oil	Fish oil	0.003	0.045
Groundnut oil	Linseed oil	0.000	0.124
Sunflower oil	Rapeseed oil	0.958	0.001
Sunflower oil	Corn oil	0.006	0.000
Sunflower oil	Palm oil	0.334	0.061
Sunflower oil	Palm kernel oil	0.190	0.273
Sunflower oil	Coconut oil	0.215	0.260
Sunflower oil	Fish oil	0.796	0.000
Sunflower oil	Linseed oil	0.561	0.000
Rapeseed oil	Corn oil	0.001	0.006
Rapeseed oil	Palm oil	0.003	0.899
Rapeseed oil	Palm kernel oil	0.012	0.269
Rapeseed oil	Coconut oil	0.052	0.530
Rapeseed oil	Fish oil	0.820	0.000
Rapeseed oil	Linseed oil	0.398	0.000
Corn oil	Palm oil	0.004	0.025
Corn oil	Palm kernel oil	0.001	0.043
Corn oil	Coconut oil	0.004	0.392
Corn oil	Fish oil	0.723	0.000
Corn oil	Linseed oil	0.720	0.001
Palm oil	Palm kernel oil	0.406	0.309
Palm oil	Coconut oil	0.504	0.688
Palm oil	Fish oil	0.849	0.005
Palm oil	Linseed oil	0.851	0.003
Palm kernel oil	Coconut oil	0.168	0.287
Palm kernel oil	Fish oil	0.071	0.007
Palm kernel oil	Linseed oil	0.971	0.000
Coconut oil	Fish oil	0.327	0.001
Coconut oil	Linseed oil	0.577	0.001
Fish oil	Linseed oil	0.008	0.652

Table B.12: Pairwise causality test over Sub-sample 1, January 2003 to December 2008.

Variable 1 y_1	**Variable 2** y_2	y_2 **does not cause** y_1	y_1 **does not cause** y_2
Soybean oil	Groundnut oil	0.292	0.000
Soybean oil	Sunflower oil	0.010	0.849
Soybean oil	Rapeseed oil	0.820	0.106
Soybean oil	Corn oil	0.019	0.436
Soybean oil	Palm oil	0.354	0.568
Soybean oil	Palm kernel oil	0.106	0.162
Soybean oil	Coconut oil	0.454	0.064
Soybean oil	Fish oil	0.128	0.046
Soybean oil	Linseed oil	0.410	0.282
Groundnut oil	Sunflower oil	0.000	0.067
Groundnut oil	Rapeseed oil	0.000	0.092
Groundnut oil	Corn oil	0.000	0.449
Groundnut oil	Palm oil	0.000	0.145
Groundnut oil	Palm kernel oil	0.000	0.280
Groundnut oil	Coconut oil	0.000	0.097
Groundnut oil	Fish oil	0.000	0.235
Groundnut oil	Linseed oil	0.016	0.285
Sunflower oil	Rapeseed oil	0.626	0.000
Sunflower oil	Corn oil	0.422	0.000
Sunflower oil	Palm oil	0.118	0.154
Sunflower oil	Palm kernel oil	0.204	0.008
Sunflower oil	Coconut oil	0.101	0.001
Sunflower oil	Fish oil	0.238	0.000
Sunflower oil	Linseed oil	0.968	0.003
Rapeseed oil	Corn oil	0.135	0.342
Rapeseed oil	Palm oil	0.009	0.858
Rapeseed oil	Palm kernel oil	0.092	0.119
Rapeseed oil	Coconut oil	0.344	0.112
Rapeseed oil	Fish oil	0.056	0.000
Rapeseed oil	Linseed oil	0.247	0.087
Corn oil	Palm oil	0.269	0.365
Corn oil	Palm kernel oil	0.574	0.012
Corn oil	Coconut oil	0.725	0.029
Corn oil	Fish oil	0.288	0.000
Corn oil	Linseed oil	0.729	0.068
Palm oil	Palm kernel oil	0.245	0.000
Palm oil	Coconut oil	0.314	0.006
Palm oil	Fish oil	0.910	0.007
Palm oil	Linseed oil	0.600	0.168
Palm kernel oil	Coconut oil	0.297	0.822
Palm kernel oil	Fish oil	0.011	0.055
Palm kernel oil	Linseed oil	0.502	0.066
Coconut oil	Fish oil	0.037	0.008
Coconut oil	Linseed oil	0.414	0.070
Fish oil	Linseed oil	0.603	0.000

Table B.13: Pairwise causality test over Sub-sample 2, January 2009 to December 2015.

Variable 1 y_1	**Variable 2** y_2	y_2 **does not cause** y_1	y_1 **does not cause** y_2
Soybean oil	Groundnut oil	0.655	0.060
Soybean oil	Sunflower oil	0.184	0.829
Soybean oil	Rapeseed oil	0.002	0.039
Soybean oil	Corn oil	0.088	0.015
Soybean oil	Palm oil	0.439	0.374
Soybean oil	Palm kernel oil	0.062	0.235
Soybean oil	Coconut oil	0.285	0.568
Soybean oil	Fish oil	0.970	0.068
Soybean oil	Linseed oil	0.582	0.003
Groundnut oil	Sunflower oil	0.032	0.091
Groundnut oil	Rapeseed oil	0.243	0.716
Groundnut oil	Corn oil	0.433	0.758
Groundnut oil	Palm oil	0.713	0.557
Groundnut oil	Palm kernel oil	0.278	0.518
Groundnut oil	Coconut oil	0.618	0.067
Groundnut oil	Fish oil	0.447	0.284
Groundnut oil	Linseed oil	0.021	0.620
Sunflower oil	Rapeseed oil	0.096	0.543
Sunflower oil	Corn oil	0.035	0.047
Sunflower oil	Palm oil	0.975	0.287
Sunflower oil	Palm kernel oil	0.700	0.216
Sunflower oil	Coconut oil	0.390	0.529
Sunflower oil	Fish oil	0.914	0.049
Sunflower oil	Linseed oil	0.800	0.046
Rapeseed oil	Corn oil	0.068	0.015
Rapeseed oil	Palm oil	0.865	0.176
Rapeseed oil	Palm kernel oil	0.955	0.620
Rapeseed oil	Coconut oil	0.941	0.571
Rapeseed oil	Fish oil	0.375	0.019
Rapeseed oil	Linseed oil	0.117	0.013
Corn oil	Palm oil	0.051	0.066
Corn oil	Palm kernel oil	0.002	0.430
Corn oil	Coconut oil	0.133	0.268
Corn oil	Fish oil	0.820	0.188
Corn oil	Linseed oil	0.307	0.087
Palm oil	Palm kernel oil	0.492	0.968
Palm oil	Coconut oil	0.626	0.995
Palm oil	Fish oil	0.058	0.121
Palm oil	Linseed oil	0.569	0.035
Palm kernel oil	Coconut oil	0.062	0.127
Palm kernel oil	Fish oil	0.331	0.610
Palm kernel oil	Linseed oil	0.753	0.123
Coconut oil	Fish oil	0.625	0.441
Coconut oil	Linseed oil	0.710	0.176
Fish oil	Linseed oil	0.785	0.485

Table B.14: Diagnostic tests of pairwise oils, given as the *p*-value.

		Full period				Sub-sample 1				Sub-sample 2				Sub-sample 3			
Variable 1	**Variable 2**	**Lags**	**BG**	**JB**	**ARCH**	**Lags**	**BG**	**JB**	**ARCH**	**Lags**	**BG**	**JB**	**ARCH**	**Lags**	**BG**	**JB**	**ARCH**
Soybean.oil	Groundnut.oil	3	0.388	0.000	0.001	2	0.140	0.012	0.401	2	0.798	0.000	0.658	5	0.000	0.000	0.726
Soybean.oil	Sunflower.oil	2	0.160	0.000	0.003	2	0.622	0.000	0.197	2	0.097	0.000	0.997	2	0.394	0.000	0.889
Soybean.oil	Rapeseed.oil	5	0.505	0.000	0.001	4	0.676	0.860	0.140	3	0.032	0.006	0.262	3	0.087	0.889	0.256
Soybean.oil	Corn.oil	3	0.921	0.000	0.000	3	0.397	0.648	0.734	2	0.183	0.001	0.539	2	0.719	0.000	0.140
Soybean.oil	Palm.oil	2	0.226	0.445	0.000	4	0.064	0.585	0.482	5	0.014	0.778	0.627	6	0.843	0.684	0.863
Soybean.oil	Palmkernel.oil	6	0.430	0.001	0.044	5	0.510	0.043	0.003	4	0.036	0.481	0.306	5	0.031	0.444	0.268
Soybean.oil	Coconut.oil	5	0.269	0.000	0.669	4	0.631	0.917	0.162	4	0.125	0.000	0.929	2	0.325	0.000	0.990
Soybean.oil	Fish.oil	2	0.337	0.000	0.082	5	0.751	0.000	0.827	2	0.700	0.000	0.950	2	0.706	0.000	0.817
Soybean.oil	Linseed.oil	2	0.745	0.000	0.029	5	0.884	0.000	0.564	2	0.115	0.000	0.127	3	0.016	0.574	0.091
Groundnut.oil	Sunflower.oil	4	0.053	0.000	0.066	2	0.290	0.000	0.068	2	0.729	0.000	0.878	4	0.000	0.000	0.994
Groundnut.oil	Rapeseed.oil	4	0.037	0.000	0.001	4	0.091	0.482	0.761	2	0.525	0.000	0.241	4	0.001	0.000	0.623
Groundnut.oil	Corn.oil	2	0.661	0.000	0.000	3	0.023	0.005	0.274	2	0.796	0.000	0.206	4	0.000	0.000	0.690
Groundnut.oil	Palm.oil	3	0.007	0.000	0.197	2	0.074	0.000	0.637	2	0.605	0.000	0.572	6	0.006	0.000	0.581
Groundnut.oil	Palmkernel.oil	5	0.477	0.000	0.088	2	0.077	0.000	0.068	5	0.136	0.000	0.187	6	0.077	0.000	0.889
Groundnut.oil	Coconut.oil	5	0.322	0.000	0.336	3	0.170	0.082	0.220	2	0.170	0.000	0.720	2	0.776	0.000	0.524
Groundnut.oil	Fish.oil	2	0.023	0.000	0.000	2	0.518	0.000	0.283	2	0.278	0.000	0.069	4	0.001	0.000	0.893
Groundnut.oil	Linseed.oil	3	0.795	0.000	0.209	2	0.632	0.000	0.904	3	0.415	0.000	0.008	4	0.000	0.000	0.803
Sunflower.oil	Rapeseed.oil	4	0.366	0.000	0.006	4	0.525	0.145	0.457	2	0.068	0.055	0.343	2	0.289	0.394	0.804
Sunflower.oil	Corn.oil	3	0.713	0.000	0.000	3	0.634	0.043	0.107	2	0.301	0.528	0.490	3	0.181	0.025	0.548
Sunflower.oil	Palm.oil	2	0.140	0.000	0.000	2	0.735	0.000	0.001	2	0.229	0.009	0.983	4	0.123	0.000	0.989
Sunflower.oil	Palmkernel.oil	5	0.192	0.000	0.000	2	0.424	0.000	0.000	4	0.014	0.061	0.504	4	0.024	0.000	0.913
Sunflower.oil	Coconut.oil	5	0.052	0.000	0.000	2	0.685	0.001	0.023	5	0.111	0.000	0.307	4	0.049	0.000	0.881
Sunflower.oil	Fish.oil	5	0.394	0.000	0.038	2	0.699	0.000	0.010	2	0.407	0.000	0.997	4	0.246	0.000	0.999
Sunflower.oil	Linseed.oil	4	0.588	0.000	0.360	2	0.654	0.000	0.887	2	0.102	0.000	0.370	2	0.017	0.000	0.678
Rapeseed.oil	Corn.oil	3	0.962	0.000	0.000	3	0.747	0.620	0.114	2	0.107	0.020	0.030	2	0.868	0.501	0.771
Rapeseed.oil	Palm.oil	5	0.190	0.000	0.000	5	0.841	0.000	0.129	6	0.026	0.651	0.644	6	0.715	0.686	0.939
Rapeseed.oil	Palmkernel.oil	5	0.870	0.006	0.000	5	0.934	0.217	0.007	4	0.059	0.459	0.635	6	0.137	0.744	0.184
Rapeseed.oil	Coconut.oil	5	0.982	0.000	0.039	5	0.959	0.182	0.030	2	0.041	0.002	0.554	4	0.082	0.000	0.987
Rapeseed.oil	Fish.oil	2	0.594	0.000	0.000	2	0.937	0.000	0.030	2	0.265	0.000	0.104	2	0.718	0.000	0.668
Rapeseed.oil	Linseed.oil	2	0.816	0.000	0.012	5	0.865	0.000	0.984	2	0.024	0.000	0.000	3	0.025	0.880	0.633
Corn.oil	Palm.oil	3	0.233	0.000	0.000	3	0.257	0.010	0.081	2	0.417	0.969	0.392	3	0.518	0.432	0.512
Corn.oil	Palmkernel.oil	5	0.493	0.000	0.000	5	0.064	0.090	0.029	5	0.017	0.891	0.717	3	0.007	0.001	0.997
Corn.oil	Coconut.oil	5	0.405	0.000	0.000	5	0.168	0.051	0.160	4	0.018	0.008	0.985	2	0.693	0.000	0.928
Corn.oil	Fish.oil	5	0.681	0.000	0.000	5	0.410	0.000	0.191	2	0.318	0.000	0.596	2	0.745	0.000	0.968
Corn.oil	Linseed.oil	2	0.709	0.000	0.060	3	0.561	0.000	0.751	2	0.426	0.000	0.667	6	0.163	0.750	0.909
Palm.oil	Palmkernel.oil	5	0.179	0.000	0.000	5	0.058	0.000	0.001	4	0.042	0.475	0.799	5	0.066	0.474	0.906
Palm.oil	Coconut.oil	5	0.048	0.000	0.003	2	0.606	0.000	0.124	5	0.320	0.029	0.566	2	0.680	0.000	1.000
Palm.oil	Fish.oil	5	0.960	0.000	0.156	2	0.917	0.000	0.085	2	0.560	0.000	0.127	2	0.566	0.000	0.752
Palm.oil	Linseed.oil	2	0.358	0.000	0.109	2	0.954	0.000	0.889	2	0.253	0.000	0.017	3	0.172	0.566	0.562
Palmkernel.oil	Coconut.oil	5	0.322	0.000	0.000	6	0.458	0.003	0.008	4	0.003	0.505	0.008	4	0.095	0.521	0.881
Palmkernel.oil	Fish.oil	5	0.925	0.000	0.403	2	0.369	0.000	0.001	4	0.366	0.000	0.205	6	0.267	0.000	0.825
Palmkernel.oil	Linseed.oil	4	0.801	0.000	0.132	2	0.849	0.000	0.003	4	0.055	0.000	0.400	4	0.056	0.202	0.993
Coconut.oil	Fish.oil	5	0.882	0.000	0.362	2	0.550	0.000	0.003	4	0.738	0.000	0.845	5	0.003	0.000	0.666
Coconut.oil	Linseed.oil	5	0.992	0.000	0.391	2	0.856	0.000	0.044	2	0.055	0.000	0.754	6	0.656	0.088	0.988
Fish.oil	Linseed.oil	2	0.753	0.000	0.023	2	0.224	0.000	0.000	2	0.284	0.000	0.886	3	0.486	0.000	0.834

Lags is chosen by the function *VARselect()* in Rstudio, **BG** is the Breusch-Godfrey test, **JB** is the Jarque-Bera test, **ARCH** is the LM test.